GH70 dextransucases : Insights on the molecular determinants of dextran molar mass control

Marion Claverie

Université Fédérale
Toulouse Midi-Pyrénées

THÈSE

En vue de l'obtention du

DOCTORAT DE L'UNIVERSITÉ DE TOULOUSE

Délivré par :

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

Présentée et soutenue par :
Marion Claverie

le mercredi 20 décembre 2017

Titre :

GH70 dextransucrases: Insights on the molecular determinants of dextran molar mass control

---

École doctorale et discipline ou spécialité :
ED SEVAB : Ingénieries microbienne et enzymatique

Unité de recherche :
Laboratory of Biosystems and Chemical Bioengineering (LISBP, UMR 5504)

Directeur/trice(s) de Thèse :

Pr. Magali Remaud-Siméon

Jury :

Pr. Magali Remaud-Siméon (LISBP INSA Toulouse) - Directrice de thèse
Dr. Bernd Nidetzky (Inst. of Biotechnology & Biochemical Engineering, Graz) - Rapporteur
Dr. Mirjam Czjzek (LBI2M Station Biologique de Roscoff) - Rapporteur
Pr. Harry Gilbert (Inst. Cell & Molecular Biosciences, MS Newcastle) - Examinateur
Pr. Pierre Monsan (TWB, LISBP INSA Toulouse) - Président du jury
Dr. Claire Moulis (LISBP INSA Toulouse) - Encadrante

Last name: CLAVERIE **First name**: Marion

Title: GH70 dextransucrases: Insights on the molecular determinants of dextran molar mass control

Specialty: Ecological, Veterinary, Agronomic Sciences and Bioengineering, Field: Enzymatic and microbial engineering.

Year: 2017 **Number of pages**: 232

Glucansucrases (GS) from glycoside hydrolase family 70 (GH70) are α-transglucosylases produced by lactic acid bacteria. From sucrose, an economical and abundant agro resource, they catalyze the polymerization of glucosyl residues. Depending on the enzyme specificity, α-glucans vary in terms of size, types of glucosidic bonds and degree of branching and have found multiple industrial applications mainly related to their molar mass (MM). However synthesizing polymers of controlled size with average MM ranging from 1 kg/mol to several millions g/mol and low polydispersity using one single enzyme remains challenging. Indeed, the molecular mechanisms underpinning the control of polymer size have been scarcely explored. To tackle this question, two GSs producing dextran (glucan composed of a majority of α-(1,6) linkages) were selected, and their mode of action explored via biochemical and structural analyses coupled to mutagenesis. The first enzyme selected, called DSR-M synthesizes only low molar mass (LMM) dextran (28 kg/mol) exclusively composed of α-(1→6) linkages without any trace of HMM dextran (10^5 to 10^8 g/mol). In contrast, DSR-OK (second model), produces the highest MM dextran (>10^9 g/mol) described to date.

Several 3D crystallographic structures of a truncated form of DSR-M (DSR-MΔ2), either free or in complex with its substrate or product (isomaltotetraose) in the domain V or in the active site were solved. Such complexes were never obtained before. Noteworthy, one structure encompassed the most complete domain V reported to date. Analyses of these structures coupled to dextran synthesis monitoring, showed that the LMM dextran specificity of DSR-MΔ2 is explained by a distributive elongation mode due to the weak affinity of its two sugar binding pockets in the domain V which interact with the growing dextran chains and allow the synthesis of dextran longer than 16 kg/mol. $^{15}N^{1}H$ NMR analyses (HSQC), for the first time performed with such a big protein, further revealed the crucial role of aromatic residues in the catalytic domain for the production of dextran from 2 to 16 kg/mol. In comparison, synthesis of HMM dextran by DSR-OK was shown to be mainly due to the sugar binding pockets of its domain V, ensuring much stronger interactions with growing dextran chains. The role of these pockets was evidenced for both enzymes, their functionality proposed to be linked to the presence of one aromatic stacking residue. Their positioning along domain V relatively to the active site is also important to promote efficient binding. All these findings highlight the cooperation between domain V and the catalytic domain for dextran elongation, offer new perspectives to acquire a deeper knowledge on this interplay and open promising strategies for GH70 enzyme engineering aiming at modulating glucan size.

Keywords: Glucansucrase, transglucosylase, GH70, dextransucrase, dextran, glucan binding domain, processivity, distributivity

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

Laboratory: Laboratory of Biosystems and Chemical Bioengineering (UMR CNRS 5504, UMR INRA 792), INSA, Toulouse

NOM: CLAVERIE **Prénom**: Marion

Titre: Dextransucrases de la famille GH70 : investigations sur les déterminants moléculaires du contrôle de la masse molaire des dextranes produits.

Spécialité: Sciences Ecologiques, Vétérinaires, Agronomiques et Bioingénieuries, Filière : Ingénieurie microbienne et enzymatique.

Année: 2017 **Nombre de pages**: 232

Les glucane-saccharases (GS) de la famille GH70 sont des enzymes produites par certaines bactéries lactiques. A partir de saccharose, substrat renouvelable et peu coûteux, elles sont capables de catalyser la synthèse d'α-glucanes, homopolysaccharides dont les propriétés diffèrent suivant la spécificité de l'enzyme (taille, type de liaisons α-osidiques, degrés de branchement). Les glucanes contenant une très grande majorité de liaisons α-(1,6), appelés dextranes, présentent de nombreuses applications industrielles qui dépendent principalement de leur taille. Cependant, la synthèse directe de dextranes de taille contrôlée (de 1 à plusieurs millions de kg/mol) avec une faible polydispersité et en utilisant une seule enzyme n'est encore pas envisageable. En effet, les mécanismes moléculaires mis en jeu pour le contrôle de la taille des polymères produits n'ont encore été que peu explorés. Dans ce contexte, deux GSs ont été sélectionnées. La première, DSR-M synthétise uniquement des dextranes de faible masse molaire (MM) (28 kg/mol) exclusivement composés de liaisons α-(1,6). *A contrario*, le second modèle, DSR-OK produit le plus long dextrane décrit à ce jour (>10^9 g/mol). La caractérisation biochimique et structurale ainsi que la construction de mutants ont permis l'exploration du mode d'action de ces deux candidats.

Plusieurs structures 3D de DSR-MΔ2 (forme tronquée de DSR-M) - sans ou en complexe avec son substrat ou ses produits (isomaltotetraose) - ont été résolues. C'est la première fois que de tels complexes sont décrits et l'une de ces structures présente le domaine V le plus complet décrit à ce jour. L'analyse de ces structures couplée au suivi cinétique de la synthèse du polymère ont montré que la spécificité de DSR-M pour la synthèse de dextranes courts s'explique par un mode d'élongation distributif dû à la faible affinité de deux poches à sucre de son domaine V envers la chaîne en cours de synthèse. Des analyses RMN ($^{15}N^{1}H$ – HSQC) – jamais réalisées auparavant sur une protéine si grosse – ont également étayé l'importance de la présence de résidus aromatiques dans le domaine catalytique pour la synthèse de dextranes supérieurs à 2 kg/mol. En comparaison, la synthèse de dextranes de haute MM par DSR-OK est principalement due au plus grand nombre de poches à sucre de son domaine V, permettant d'assurer une meilleure interaction avec la chaîne en cours d'élongation. L'implication de ces poches dans la détermination de la taille du dextrane a été montrée pour les deux candidats. Leur fonctionnalité est fortement liée à la présence d'un résidu aromatique de stacking, et leur répartition le long du domaine V a aussi une influence. L'ensemble de ces résultats démontre la coopération du domaine V avec le domaine catalytique pour l'élongation des dextranes, tout en offrant de nouvelles perspectives pour approfondir la compréhension de ce mécanisme. Ils offrent également des stratégies prometteuses pour l'ingénierie d'enzyme de la famille des GH70 pour la modulation de la taille des glucanes.

Mots Clés : Glucane-saccharase, dextrane-saccharase, GH70, transglucosylase, dextrane, domaine de liaison au glucane, processivité, distributivité

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

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

Publications

Investigations on the determinants responsible for low molar mass dextran formation by DSR-M dextransucrase. Claverie M., Cioci G., Vuillemin M., Monties N., Roblin P., Lippens G., Remaud-Simeon M. and Moulis C. *ACS Catal.*, 2017, 7, 7106–7119

Futile cycle engineering of the DSR-M dextransucrase modifies the resulting polymer length. Claverie M., Cioci G., Guionnet M., Schörghuber J., Lichtenecker R., Moulis C., Remaud-Simeon M. and Lippens G. *In preparation*

High molar mass dextran synthesis by DSR-OK dextransucrase from *Oenococcus kitaharae* DSM 17330, role of the domain V? Claverie M., Cioci G., Bondy P., Vuillemin M., Esque J., Magali Remaud-Simeon and Claire Moulis. *In preparation*

Characterization of the First α-(1→3) Branching Sucrases of the GH70 Family. Vuillemin M., Claverie M., Brison Y., Séverac E., Bondy P., Morel S., Monsan P., Remaud-Simeon M. and Moulis C. (2016). *J. Biol. Chem.* 291, 7687–7702.

A dextran with unique rheological properties produced by the dextransucrase from *Oenococcus kitaharae* DSM 17330. Vuillemin M., Grimaud F., Claverie M., Rolland-Sabaté A., Garnier C., Lucas P., Monsan P., Dols-Lafargue M., Remaud-Siméon M., and Moulis C. (2017). *Carbohydr. Polym.* 179, 10-18.

Oral communications

Discovery of new α-transglucosylases from *Leuconostoc citreum* NRRL B-1299 and NRRL B -742 for the synthesis of tailor-made α-glucans. Claverie M., Vuillemin M., Passerini D., Severac E., Grimaud F., Monsan P., Morel S., Remaud-Simeon M., Moulis C. *11th Carbohydrate Bioengineering Meeting 2015, May 10-13, Espoo (Finland)*

Biochemical characterization of a new GH-70 enzyme from *Leuconostoc citreum* NRRL B-1299. Claverie M., Vuillemin M., Monsan P., Cioci G., Moulis C., Remaud-Simeon M. *6th Symposium on the Alpha-Amylase Family (ALAMY_6) 2016, September 11-15, Smolenice (Slovakia)*

Posters

Biochemical characterization of a new GH-70 enzyme from *Leuconostoc citreum* NRRL B-1299. Claverie M., Vuillemin M., Severac E., Monsan P., Cioci G., Remaud-Simeon M., Moulis C. *11th Carbohydrate Bioengineering Meeting 2015, may 10-13, Espoo (Finland)*

Characterization of a new GH-70 enzyme from *Leuconostoc citreum* NRRL B-1299 producing low molar mass dextrans. Claverie M., Vuillemin M., Severac E., Monsan P., Cioci G., Remaud-Simeon M., Moulis C. *Biotrans 2015, July 26-31, Vienna_Austria*

Structure/Function relationships study of two novel GH70 enzymes: role of their GBD. Claverie M., Vuillemin M., Cioci G., Moulis C., Remaud-Simeon M. *DLS-CCP4 Data Collection and Structure Solution Workshop2015, December 1-9, Oxford (England)*

Patents

Protein with dextran-saccharase activity, and uses. Vuillemin M., Claverie M., Séverac E., Fontagné-Faucher C., Monsan P., Moulis C., Remaud-Siméon M. (2014). US20170247669.

Very high molar mass dextrans. Vuillemin M., Claverie M., Sabate A., Garnier C., Lucas P., Dols M., Grimaud F., Monsan P., Moulis C., Remaud-Siméon M. (2014). US2017145120.

In memory of my Grandfather

‘’La science, mon garçon, est faite d'erreurs, mais d'erreurs qu'il est bon de commettre, car elles mènent peu à peu à la vérité.’’

Jules Verne, Voyage au centre de la Terre

“There is only one thing that makes a dream impossible to achieve: the fear of failure.”

Paulo Coelho, The Alchemist

‘’Dans la vie, rien n'est à craindre, tout est à comprendre.’’

Marie Curie

Il n'est jamais facile de se dire que c'est terminé et il est vrai qu'en sciences ça ne l'est jamais… Plus on avance plus le chemin s'allonge, il y a parfois des pauses, des changements de routes, quelques demi-tours, des carrefours sans panneau d'indication ; mais chaque pas en engendre un autre et parfois des récompenses ou autres points de ravitaillement vous donne l'envie de toujours poursuivre… Si j'en suis arrivée là aujourd'hui c'est bien grâce aux gens qui m'entourent et même si quelques mots sont peu de choses en retour de ce que j'ai reçu, je souhaite les partager pour les remercier.

Claire, merci de m'avoir fait confiance pour ce stage de cinquième année et de m'avoir ainsi accueillie dans cette grande équipe de l'EAD1… Magali, toi, la maman de cette équipe, toujours à l'écoute et si compréhensive. Tu sais partager ta passion pour ton travail tout en nous faisons comprendre que le reste est encore plus important. Merci à toutes les deux pour tout le temps que vous m'avez donné et pour toutes les discussions scientifiques ou pas que l'on a pu avoir. Vous êtes des exemples et j'espère avoir été à la hauteur dans ce solide trio. Trio ou plutôt quatuor avec toi Marlène ! Encadrante de choc lors de ce fameux stage où tout a commencé, mais surtout grande amie, confidente et modèle. Ça n'a pas toujours été facile de prendre ta suite mais j'espère que j'ai bien pris soin de tes bébés enzymes qui, encore aujourd'hui, continuent à doucement nous dévoiler leurs secrets.

Gianluca, mon fidèle complice ! Tu as fourni énormément d'énergie pour qu'on arrive enfin à résoudre cette structure. Merci d'avoir eu la patience dont je manquais parfois et même s'il reste un sentiment de « on y était presque » pour DSR-OK, je te remercie de m'avoir transmis quelques brides de ton savoir-faire parfois si intuitif. Merci aussi pour ta bonne humeur et ta sympathie.

Merci à toi Guy, pour m'avoir initiée à ton outils préféré et pour me faire confiance pour les mois à venir. Pierre Roblin, merci pour toutes ces longues heures au synchrotron et pour toute l'aide par la suite, merci aussi à tous les swingueurs pour leur sympathie ! J'espère que vous vous souviendrais de mes efforts pour vous ramener le soleil des montagnes béarnaises dans votre SOLEIL à vous (Pense bête à moi-même : ne plus oublier la crème solaire au ski !).

Merci à l'équipe cristallo de l'IPBS pour m'avoir accueillie chez vous pour utiliser vos si « fancy robots » ! Plateforme ICEO et PICT en général, heureusement que l'on vous a !

A tous mes comparses de l'équipe Enzynov, un grand merci pour tout. C'est un bonheur de travailler dans cette ambiance qui nous est propre et j'espère qu'elle perdurera encore bien longtemps.

Floflo et Etienne (ou plutôt, Grimoto et Severacus Gromulus) mes deux frérots, un peu papa aussi parfois (mais non vous n'êtes pas vieux !). Merci pour votre aide et votre disponibilité sans faille

pendant nos longues heures devant ces foutues HPLC ou encore à SOLEIL privés de sommeil ! Mais n'oublie pas Etienne, PCCM et surtout reste comme tu es #unpeufauxculmaishonnête. Et toi Floflo, à quand la prochaine choré façon Mickael ? Ma petite Popo, petite sœur du labo. Je te remercie tellement pour tout ce que tu as fait pour m'aider. Cette thèse t'appartient en grande partie ! Nos discussions m'ont souvent redonné du souffle et je pense que c'était réciproque ;)… Tu es juste parfaite, surtout reste toi ! Moumou, the petit coquinou, toujours le cœur sur la main ! Petit message : quand tu seras PDG de ta boîte, tu penseras à moi hein ?! Booboops, Angie, mes folles chéries, merci pour tous les moments que l'on a partagés, vous êtes toujours au top ! Maeva, merci pour tes coups de main avec toutes ces constructions et n'oublie pas de m'envoyer des photos de tes créations ;) ! Nelly, j'espère que l'I6 ne te donnera plus de cauchemar, merci pour tout ! Lolo, la cinétique ça envoie ! Jérem, j'espère ne pas t'avoir rendu allergique au GH70, pas faciles à modéliser ces bestioles… merci ! Thibault, Gaël, Anaïs, mes trois stagiaires, merci pour votre aide !

Alex, Pablito, les Helsinkiers forever ! Merci pour votre soutien de choc pendant ce congrès presque improvisé et pour votre amitié qui m'est si précieuse. Toi Alex, nos longues discussions devant mon portail alors que l'on voulait rentrer tôt et toi Pablito, tu as toujours les bons mots (même s'ils ne sont pas toujours dans le bon ordre !). Julius, oreille attentive et générosité immense merci pour les moments que l'on a partagés, ne laisse personne te changer.

Coco, Eli ! Les plus jolies mamies ! Toi Eli, merci pour tes coups de mains qui sont toujours tombés à point et ton sourire si communicatif ! Ma petite Coco, j'espère que ta cheville ne ré-enflera pas mais je ne peux pas le dire autrement : tu es géniale ! Merci pour tout (la liste est longue) et je compte sur toi pour la prochaine soirée couture !

Tous les autres EAD1 actuels et anciens, la Sto, Marina, Manon C, Mathilde, Nada, Louise, Emilie, Yoann, Alizée, Bastien, Pablin, Greg, Sandra, Julian, Gaby, Florence, Haiyang, Yannick, Nathalie C, les modélos, David G (#soleil), Sandrine, et tout le reste de la liste… Merci à vous tous.

Merci également aux membres de mon jury de thèse pour l'intérêt qu'ils ont porté à ces travaux et pour les riches échanges lors de ma soutenance. Mirjam Czjzek, un merci particulier pour m'avoir accueillie lors de ma toute première formation à Roscoff et pour tes encouragements lors de nos rencontres tout au long de ces quatre ans.

A toute l'équipe du Bad, Laëtitia, Steph, Nathalie, Vincent, Simon, Evrard et les autres merci de m'avoir permis de me défouler dans les moments compliqués et merci de me faire autant rire même si je suis mauvaise perdante ;) !

Alisson, ma coloc' adorée ! Merci de m'avoir supportée pendant ces deux ans… vivre avec toi était génial ! Pleins de beaux souvenirs et encore de nombreux à venir…

Merci à toi Maman d'avoir toujours été là ,même si souvent je devais te répondre que « non je n'ai encore rien trouvé » ;). Et toi ma kiki sœur, vrai pilier dans ma vie, j'admire ta force et ta créativité. Merci à toutes les deux d'être si présentes et de m'avoir permis de combattre mes nombreux doutes. Papa, merci pour ton soutien même si tu n'as pas toujours compris mes choix, j'espère te rendre fier… Babeth, Bernard merci de faire partie de ma famille et d'être toujours présents dans les moments importants. Gaëlle, américo-béarnaise pleine de joie et de folie, merci pour tous les petits moments (oui tu es loin quand même) que l'on a partagés pendant ces années et merci pour cette belle pause américaine qui m'a donné plein d'énergie après ma première année. Un énorme merci à tous les membres de ma famille pour leur soutien et leur implication avec une attention particulière pour toi Bon-Papa, toi qui avais toujours mille idées pour faire avancer ce monde, sans le savoir tu es sûrement celui qui m'a le plus compris, tu es parti trop tôt pour lire ces mots mais j'espère que de là-haut ils te donneront quand même le sourire. Je vous aime.

Et puis toi Julien, pas facile pour toi de comprendre pourquoi cela demande tant d'efforts et de temps mais malgré tout tu t'es toujours intéressé et surtout tu m'as toujours soutenue. Merci pour ta patience et le bonheur que tu m'apportes chaque jour…

TABLE OF CONTENT

Introduction

Carbohydrates in the form of monosaccharides, oligosaccharides or polysaccharides are essential components of living systems. The diversity of carbohydrate structures is colossal due to the abundance of stereoisomers, variety of assemblies and possible decorations encountered in nature. This vast array of structures is related to a wide panel of biological functions. Carbohydrates can be used as energy source, carbon reserve, structural compounds, mediators of intra- and intercellular recognition... The variations of these molecules in terms of biological and physico-chemical properties are also extensively exploited in food, feed, health, material, chemical and bioenergy industries, with the advantage of being usually biodegradable and biocompatible.

This abundance of structures is related to the multiple Carbohydrate-Active enZymes (CAZymes) involved in their breakdown (glycoside hydrolases, polysaccharide lyases, carbohydrate esterases) or their assembly (glycosyltransferases). For over than 25 years, these CAZymes have been classified in families depending on their sequence and structure similarities (CAZy database, http://www.cazy.org/). Among them, the Glycoside-Hydrolase (GH) group counts currently 145 different families composed of glycoside-hydrolases and transglycosylases. With the GH13 and GH77, the GH70 family belongs to the GH-H clan. This family comprises α-tranglucosylases divided in three subgroups: glucansucrases, branching-sucrases and α-glucanotransferases. Except for the last subgroup, these enzymes use sucrose as substrate to catalyze the transfer of the glucosyl unit onto the hydroxyl group of an acceptor, resulting, depending on the nature of the acceptor, in the formation of α-glucans, glucooligosaccharides or glucoconjugates. In contrast with other polysaccharide synthesizing enzymes, GH70 members are very attractive since they do not require the use of costly activated sugars. Moreover, sucrose is a cheap and abundant renewable resource. Its production from sugar cane or beet exceeded 170 million tons in 2016 and its price is currently around 320€/ton (London Stock Exchange).

Glucansucrases (GSs) constitute the major subgroup of enzymes in the GH70 family. Different types of α-glucans can be produced by GSs and dextransucrases catalyze specifically the synthesis of dextrans, an α-glucan containing more than 50% of α-(1→6) linkages. Among the existing dextrans, the highly linear dextran with more than 95% of α-(1→6) linkages was the first microbial polysaccharide industrially produced in the 40s for pharmaceutical applications. Since that time, linear dextrans have found many other applications in cosmetics, agri-food and feed, fine and bulk chemistry or oil recovery fields. However, dextran uses largely depend on their physico-chemical properties related to their structure and more particularly their size. Indeed, most of the applications concern products of controlled molar mass (10, 40, 70 kg/mol), still today obtained after acid hydrolysis and solvent fractionation of the native polymers. Being able to control the final dextran

size directly from the synthesis would hence be a significant advantage from a biotechnological point of view to develop more efficient and eco-friendly processes.

In the past decades, considerable progress has been made in the characterization of the structure-function relationships of GH70 α-transglucosylases and dextran-producing enzymes in particular. Notably, various amino acid sequences allowing the prediction of osidic linkage specificity of these biocatalysts have been identified. But less is known on the determinants governing the size of the produced polymers.

Investigating this aspect was the central objective of this PhD thesis conducted in the Enzynov subgroup of the CIMEs team (LISBP Toulouse). In order to enlarge the repertoire of GH70 family enzymes and investigate their structure-function relationships with the goal of providing optimized enzymes for tailor-made products, the Enzynov team has recently identified, in the genomes of *Leuconostoc citreum* NRRL B-1299 and *Oenococcus kitaharae* DSM 17330, two genes of special interest to address the question of the polymer size control by dextransucrases. Indeed, the first one (DSR-M from *L. citreum* 1299 strain) naturally synthesizes only a low molar mass (LMM) linear dextran whereas the second one (DSR-OK from *O. kitaharae* DSM17330) produces the highest molar mass (HMM) dextran described to date. These enzymes were chosen as enzyme models for our study and were extensively biochemically and structurally characterized, with the key objective of identifying structural determinants involved in chain length specificity.

The first chapter of this manuscript is a literature review presenting the general context of the present study. Notably, the structures, properties and applications of α-glucans produced from sucrose by lactic acid bacteria -and more particularly of dextrans- are described in a first part while the enzymes responsible for their synthesis is the topic of a second part. Finally, the main known characteristics of the two enzymatic model of study, namely DSR-M and DSR-OK, are reviewed in the third part.

The principal results obtained during the project will then be presented in three following chapters in the form of three scientific articles.

Chapters II and III are dedicated to the characterization of DSR-M, the LMM polymerase candidate. A full description of DSR-M catalytic mechanism supported by structural data and investigations on the function of its domain V are discussed in the first article. The second describes more specifically an NMR study that was used to investigate more particularly the role of a tryptophan residue located in the catalytic center of the enzyme and decipher its function in the elongation process.

Finally, Chapter IV is centered on the second model of study (DSR-OK, the HMM polymerase). A deep characterization of this enzyme and in particular of the role of its domain V is addressed. The comparison of the control of dextran chain elongation by each candidate is discussed in the "conclusion and perspectives" section.

Chapter I

GH70 α-transglucosylases: attractive tools for the production of bio-sourced tailored polysaccharides, a literature review

Context

Polysaccharides are polymers of osidic units from natural origin (vegetal, animal, algal or microbiological) and represent the source of much of the biomass on the planet (Prestegard et al., 2015). They can be classified in two main categories: i) storage polysaccharides which are a storage form of energy, for example starch in plants or glycogen in animals and humans ii) structural polysaccharides such as cellulose in plants and chitin in the shells of crustaceans. The most common monosaccharides that appear as parts of polysaccharides are glucose, fructose, galactose and mannose. The world market is still dominated by plant (eg. cellulose, starch, arabic gum) and algal (eg. agar, alginate, carrageenan) polysaccharides; however, microbial polysaccharides are raising interest for several years. Depending on their chemical composition, their size and their structure, polysaccharides display a large panel of physico-chemical properties and are used in various application sectors. For instance, they can be used as thickener, stabilizer or gelling agents in agri-food and pharmaceutical industries. Some of them also present biological activities and can be used in medical applications. Among microbial polysaccharides, we can find exopolysaccharides (EPS), extracellular polysaccharides produced by bacteria, archaea, algae or filamentous fungi. EPS can differ in monomer composition, substituent decoration, degree and type of branching as well as molar mass. EPS are referred to as homopolysaccharides when they are composed of only one single monomer type, or hetero-polysaccharides when different monosaccharides are present in the polymer. Thanks to their functional properties, some of them can also be of industrial interest (Moscovici, 2015). Xanthan gum, Gellan, Alginate, Levan, Hyaluronan, Succinoglycan and Glucans are the most studied and commercially used EPS (Freitas et al., 2011). The latest are classified in two main groups: β- glucans (e.g cellulose) and α-glucans (dextran, mutan, alternan and reuteran).

α-glucans produced from sucrose by lactic acid bacteria (LAB) belonging to the genera *Leuconostoc, Streptococcus, Weissella, Lactobacillus, Pediococcus* and *Oenococcus* are neutral homopolysaccharides of α-D-glucosyl units (Monsan et al., 2001; Shukla and Goyal, 2013; Vuillemin et al., 2017). Their properties and applications will be described in the first part of this literature review. They are synthesized by α-transglucosylases from the family 70 of the glycoside-hydrolases (GH70) comprising glucansucrases, branching-sucrases and α-4,6-glucanotransferases. Description of these enzymes and the actual knowledges on their structure-function relationships will be the object of the second part. Finally, the two enzyme models of the study will be presented in a third and last part of this review.

I. α-glucans produced from sucrose by lactic acid bacteria

I.1. Structural diversity

Depending on the arrangement and particularly on the type of their osidic linkages (α-(1→2), α-(1→3), α-(1→4) and/or α-(1→6)), these homopolymers present a large diversity of structures and can be weakly or highly branched. They usually display a high molar mass (HMM from 10^5 to 10^9 g/mol). According to the nature of their predominant osidic linkage, LAB α-glucans are classified in four main categories (Figure 1).

- Dextrans contain more than 50% of α-(1→6) linkages in their main chain and are synthesized by transglucosylases named dextransucrases. Depending on the productive enzyme, dextran molecules may possess variable branching ratios through α-(1→2), α-(1→3), and/or α-(1→4) bonds as well as different molar masses from 10^5 to 10^9 g/mol (Jeanes et al., 1954; Kobayashi and Matsuda, 1977). To our knowledge, dextran produced by *Leuconostoc mesenteroides* NRRL B-512F is still today the most industrially used and is very linear with 95% of α-(1→6) and 5% of α-(1→3) linkages (Larm et al., 1971; Kuge et al., 1987). In contrast, the dextran produced by *Leuconostoc citreum* B-742 (fraction S) consists of a linear chain of α-(1→6) linked D-glucopyranosyl units, but each of them is substituted with a single glucosyl unit through α-(1→3) bond, thus resulting in a highly branched "comb-like" polymer (Seymour et al., 1979a, 1979b; Côté and Robyt, 1983). Most recently, Vuillemin et al. proposed that this original dextran is produced thanks to the combined action of an enzymatic duo comprising both a dextransucrase and a branching-sucrase (Vuillemin et al., 2016) (see paragraph, II.3.1. Brief overview of the branching-sucrases, p46). Considering its high content in rather rare α-(1→2) linkages (30%), the dextran produced by *Leuconostoc citreum* NRRL B-1299 has also been widely studied (Bourne et al., 1972; Kobayashi and Matsuda, 1977; Seymour et al., 1977) and is also mainly produced by a duo of enzymes (Passerini et al., 2015).
- Mutans, produced by mutansucrases, are composed of more than 50% of α-(1→3) glucosidic linkages and can be branched through α-(1→2), α-(1→4), and/or α-(1→6) linkages (Hamada and Slade, 1980). Their molar mass is equivalent to that of dextran and they are essentially synthesized by Streptoccocal enzymes from the buccal flora. Mutans are implicated in the cariogenesis by participating in dental plaque formation which favors bacterial proliferation and then dental decay. For that reason, these α-glucans and the enzymes responsible for their production have been largely studied (Koga et al., 2002; Leme et al., 2006; Pleszczyńska et al., 2010).

- Alternans, produced by alternansucrases, are composed of alternating α-(1→3) and α-(1→6) linkages in the main chain, and also contains low level of α-(1→3) linked branches (Côté and Robyt, 1982). It has been suggested that some segments of consecutively α-(1→6) linked glucose are present in the main chain. This could allow dextranase cleavage within contiguous α-(1→6) linkages yielding to smaller products potentially usable as Arabic gum substitutes (Leathers et al., 2009). Up to now, although databases contain numerous putative alternansucrase encoding genes, only a few strains synthesizing alternans have been characterized, such as *Leuconostoc mesenteroides* NRRL B-1501 and NRRL B-1498 as well as *Leuconostoc citreum* NRRL B-1355 and LBAE C11 (Jeanes et al., 1954; Seymour and Knapp, 1980; Amari et al., 2015).
- Reuterans, produced by reuteransucrases, are mainly composed of α-(1→4) linked glucosyl units but also display different amounts of α-(1→6) linkages and α-(1→6) branches. *Lactobacillus reuteri* 121, 35-5 and ATCC 55730 (Kralj et al., 2002, 2005a) were described as reuteran producers. The polymer produced by the strains 121 and 35-5 present the particularity to be composed of an alternation of α-(1→4) and α-(1→6) linkages in the main chain (Dobruchowska et al., 2013; van Leeuwen et al., 2008a).

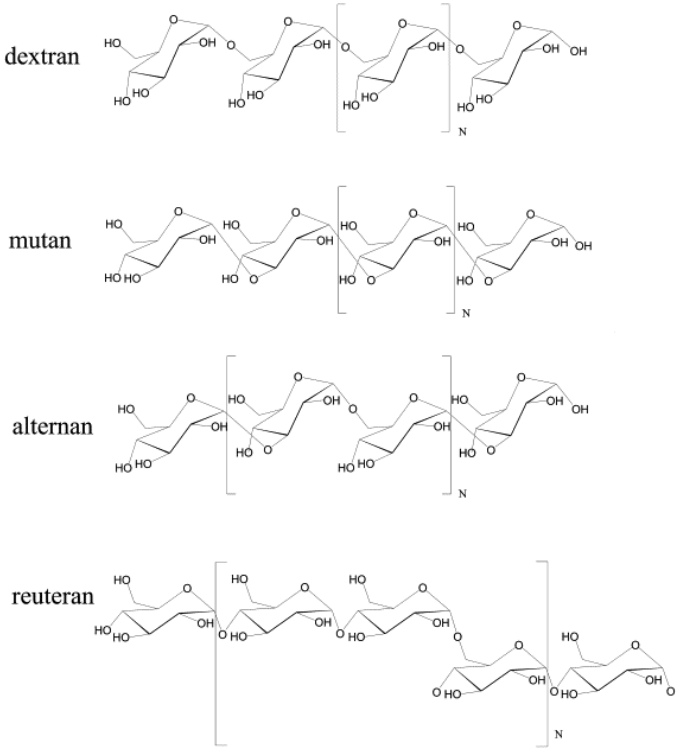


Figure 1: **Schematic representation of α-glucan structures.** From (Monsan et al., 2010).

All those α-glucans are produced by α-transglucosylases belonging to the family 70 of the Glycoside Hydrolases and will be presented in the section: II. α-transglucosylases of the GH70 family: structure/function relationships and understanding of their catalytic mechanism, p44. As my project was related to the study of dextransucrases, the following paragraphs will more specifically focus on dextran.

I.2. Dextran applications

Thanks to their physico-chemical properties, native dextrans, partially hydrolyzed dextrans and their derivatives have wide applications in food, fine and bulk chemistry, medicine, pharmacy and cosmetics sectors (Badel et al., 2011; Monsan et al., 2001; Naessens et al., 2005; Suresh Kumar et al., 2007).

I.2.1. Pharmaceutical and medical sectors

The resistance of dextrans to high temperature and sterilization, its low antigenicity (when α-(1→6) linkage content is over 95%), its high stability and solubility in water (Robyt, 1985) added to the facts that it causes limited side effects when administrated in human, the fact that it can be chemically modified and reticulated to metallic ions and/or hemoglobin proteins (Takemoto and Liebhaber, 1962) were among the most advantageous properties for pharmaceutical and medical usages.

During the World War II, the first industrial productions of dextran were dedicated to the preparation of clinical dextrans used as a blood volume expander by the armed forces (cf. I.3. Dextran production processes, p41). Clinical dextrans typically consisted of fractions with a molar mass comprised between 50 and 100 kg/mol. Indeed, Native dextran with a molar mass over 10^8 g/mol is unsuitable for human biomedical uses as it can interfere with the normal blood coagulation process while lower mass dextrans are rapidly eliminated by the organism and therefore therapeutically ineffective (De Belder, 2003; Hellman et al., 1955; Naessens et al., 2005). Nowadays, clinical fractions of 40 or 70 kg/mol are available and used to replace moderate blood losses. Contrary to other crystalloid molecules which rapidly penetrate vascular membranes, dextrans present the advantage to be longer retained in the blood circulation to maintain water in the blood flux (Robyt, 1985). In addition to the application as plasma volume expander, dextran 40 is also used for blood flow improvement since it reduces blood viscosity and inhibits erythrocyte aggregation. As injection of dextran solution to patient may induce anaphylactic reactions (dextran-induced anaphylactoid reactions (DIARs)) which generally are antibody mediated), small volume of dextran 1 (1 kg/mol) are now administered prior to the Dextran 40 or 70 infusion (De Belder, 2003; Hedin and Richter, 1982; Naessens et al., 2005). Dextrans 40 and 70 also possess antithrombotic properties and

provide a prophylactic treatment for deep venous thrombosis and post-operative fatal pulmonary emboli (Foster et al., 1966; Lambie et al., 1970).

Other pharmaceutical applications have also been developed (see the website www.dextran.net). For instance, it can be used in:

- blood cell separation, separation of red blood cells from other cells and components of the blood,
- organ preservation, a multitude of different organs and tissues are being stored in dextran solution for increased longevity, and dextrans are also used in preparation prior to organ transplantation,
- cryopreservation, in combination with DMSO, glycerol etc. Dextrans can be used to cryo-preserve biological samples,
- drug lyophilization procedures, dextran is often used as excipient during lyophilization as a bulking agent and/or a collapse temperature modifier for preservation of small molecules, proteins or vaccines, which are unstable in aqueous medium and/or thermolabile (Baheti et al., 2010),
- ophthalmic solutions, due to its lubricating nature, dextran is used in eye drops against dryness or as tear-replacement but also in eye solutions containing medicating components,
- vaccine preparations as a carrier, a backbone and/or as antigen stabilizer

Dextran derivatives such as iron-dextrans can also be used in case of severe anemia (Auerbach et al., 1988 and http://www.infed.com/). Numerous studies on iron-dextran applications have been published in the recent years and report beneficial effects of dextran-based treatments (in particular with low molar mass dextrans) (Cho et al., 2013; Deitering et al., 2016; Lee et al., 2017; Pollock and Muduma, 2017).

Concerning dextran-sulfate, it is mainly used for its anticoagulant properties closed to that of heparin, or to treat arteriosclerosis in oral medicine. It has also been studied for antiviral applications (De Belder, 2003). In particular, it might prevent viral adhesion on host cells notably in the case of sexually transmitted diseases such as HIV or herpes (Ghosh et al., 2009; Hershline, 2004; Piret et al., 2000).

Numerous recent studies proposed the use of dextran as medicine vectors for drug or probe delivery (Dai et al., 2014; Sun and Mao, 2012; Zhou et al., 2013). Certain dextran derivatives (copolymers of dextran from 1.5 to 70 kg/mol, with other hydrophobic polymers) are able to auto-assemble, resulting in spherical aggregates allowing drug capture (Sun and Chu, 2011; Wasiak et al., 2016) (Figure 2). This is particularly the case of dextran poly (ε-caprolactone) (Villemson et al., 2006) which

is a biocompatible and biodegradable amphiphile allowing the encapsulation of different active molecules such as amoxicilline (antiobiotic, (Saldías et al., 2015)), doxorubicin (anti-cancer drug, (Li et al., 2013)), indomethacin (anti-inflammatory effect, (Lee et al., 2015)), and hence their delivery by diffusion or erosion of the polymeric capsule. In the same idea, microcapsules of dextran/chitosan have been developed for polyphenol delivery presenting anti-inflammatory, anti-viral, anti-oxidant or anti-bacterial properties (Paini et al., 2015). As they allow spatial and temporal delivery control, these encapsulation technics are promising for targeted therapies against cancer for example.

A recent patent proposed that fluorescent dextran may also be used as angiographic dye for diagnosis and evaluation of eye diseases in humans before and during surgeries (Ni et al., 2013).

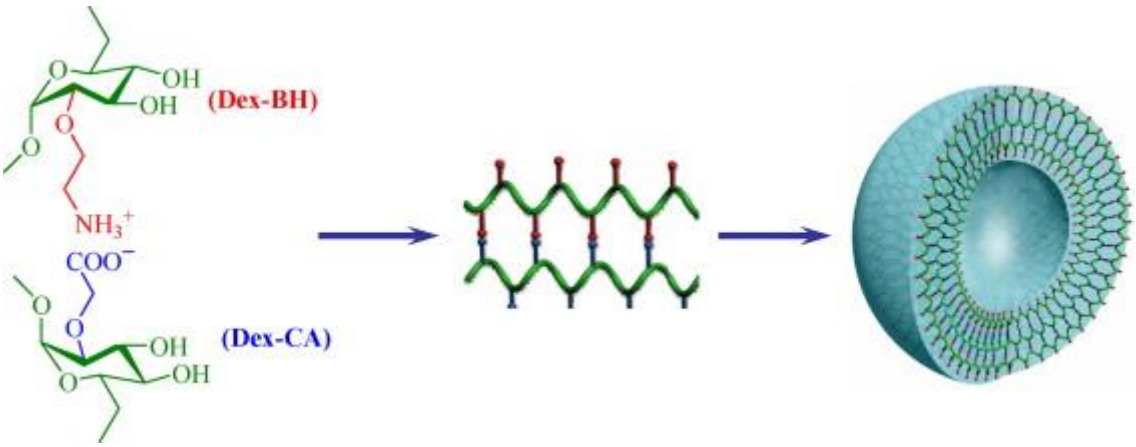


Figure 2: **Schematic illustration of dextran capsule formation.** Dex–BH and Dex–CA can become oppositely charged due to protonation and deprotonation in aqueous solution, respectively; they attract each other through electrostatic interactions to form spherical structures. From (Sun and Chu, 2011).

Finally, new applications in tissue regeneration using dextran-based polymers and scaffolds for controlled release and tissue engineering are in progress (Sun and Mao, 2012). Notably, dextran molecules cross-linked to pullulan form a biocompatible scaffold that could be used for cardiovascular tissue engineering to improve cardiac functions by restoring damaged tissues (Figure 3) (Chaouat et al., 2006; Le Visage et al., 2012; Silva et al., 2015). The use of such scaffold may promote local cellular engraftment and survival.

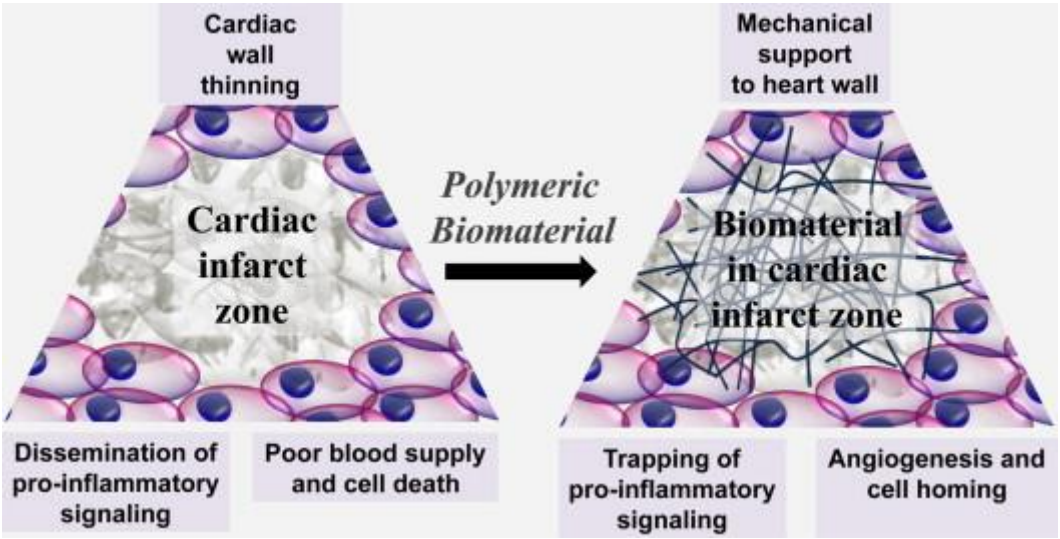


Figure 3: **Processes related to cardiac infarct (left) and the mechanisms of biomaterial interaction with these processes to promote cardiac tissue repair (right).** From (Silva et al., 2015).

Actual and potential dextran applications in the pharmaceutical and medical fields are thus numerous. Another important application field at the border between nutrition and health concerns the use of dextran and more particularly oligodextrans as functional ingredients with prebiotic properties and/or activities against metabolic diseases. This will be further described in the following paragraph.

I.2.2. Prebiotics sector

Oligodextrans, and more particularly isomaltooligosaccharides (IMOS) - glucooligosaccharides (GOS) quasi-exclusievely composed of α-(1→6) linkages - were shown to display prebiotic properties and their commercialization as functional ingredients is permanently increasing (Goffin et al., 2011; Ketabi et al., 2011; Madsen et al., 2017).

I.2.2.1. Definition and market

Since several years, a growing interest in functional food is observed enhancing the market development of such ingredients. As defined by Roberfroid: "A food can be regarded as functional if it is satisfactorily demonstrated to affect beneficially one or more target functions in the body, beyond adequate nutritional effects, in a way which is relevant to either the state of well-being and health or the reduction of the risk of a disease. A food can be made functional by increasing the concentration, adding or improving the bioavailability of a particular component" (Roberfroid, 2002). As they influence intestinal microbiome composition and hence contribute to the host good health, prebiotic molecules are considered as functional food. Prebiotics are bioselective substrates which favor beneficial flora to the detriment of pathogens (Bindels et al., 2015). They should not be confounded with probiotics, which are beneficial living microorganisms and help maintaining healthy gastrointestinal conditions.

Prebiotic intakes hence lead to better global intestinal health, mineral absorption, cholesterol regulation, immunity, as well as prevention and resistance to various diseases. Interestingly, very recently, it has been suggested that gut microbiota play a direct role in brain chemistry balance and therefore maintaining a beneficial balance of gut microbiota could lead to better psychological health and less age-related decline (Messaoudi et al., 2011; Savignac et al., 2013). These newly discovered properties reinforce the popularity of prebiotics, and stimulate studies on the microbiota-gut-brain axis (Burokas et al., 2015, 2017; Desbonnet et al., 2015; Schmidt et al., 2015).

According to a new research report by Global Market Insights, Inc, **the prebiotics market** exceeded USD 3 billion in 2015 and should reach USD 7.5 billion by 2023. It mainly concerns food and beverage applications with an expected production level of 1.1 million tons by 2023 (https://www.gminsights.com/industry-analysis/prebiotics-market).

I.2.2.2. Dextrans and isomaltooligosaccharides as prebiotics

α-(1→6) linkages are less digestible by the organism than the more common α-(1→4) linkages, which are degraded by human α-amylases. Hence, α-(1→6) linked carbohydrates are more susceptible to attain the intestinal tract with a limited degradation and promote the development of beneficial intestinal bacteria such as bifidobacteria, lactobacilli and *Bacteroides*, thereby possessing prebiotic properties (Chen et al., 2001; Crittenden and Playne, 1996; Hu et al., 2013; Kaneko et al., 1994; Olano-Martin et al., 2000).

IMOS, and more generally "IMO" syrups composed of GOS with polymerization degree (DP) varying from 2 to around 10, are the market leader in the dietary carbohydrate sector of functional foods in Japan. The preparations are not exclusively composed of isomaltooligosacharides but also contain glucooligosaccharides with both α-(1→6) and α-(1→4) linkages and to a lesser extent glucooligosaccharides with α-(1→3) or α-(1→2) linkages (Goffin et al., 2011). These compounds are referred as "ALOs" (Anomalously linked oligosaccharides). They present low digestibility by endogenous enzymes and can reach the large intestine where they can be fermented by intestinal beneficial bacteria. Commercial IMO preparations have been shown to exert prebiotics effects and are now incorporated in many commercially available food products including protein/fiber bars, shakes, and other dietary supplements (Ketabi et al., 2011; Likotrafiti et al., 2014; Madsen et al., 2017).

The majority of "IMO" preparations commercially available are produced from plant starch. After starch saccharification, transglucosylation of maltose and maltodextrins of low polymerization degree is performed with one α-transglucosylase of *Aspergillus niger* belonging to GH31 family (Röper and Koch, 1988) resulting in a mixture containing Panose-type oligosaccharides (α-(1→6)

oligoglucan chains terminated by an α-(1→4)-linked glucose), true isomaltooligosaccharides but also kojibiose, nigerose, glucose and residual resistant maltodextrins (Fernández-Arrojo et al., 2007; Gutiérrez-Alonso et al., 2016; Hung et al., 2005; Saburi et al., 2015; Wang et al., 2009). The Hayashibara Corporation in Japan industrially developed this transglucosylation reaction process for "IMO" syrup commercialization (Yoneyama et al., 1992). Other commercialized IMOS products, such as ISOThrive Prebiotic Nectars (available from 2016, ISOThrive, LLC, Healdsburg, Calif. USA, https://www.isothrive.com/) are produced via bacterial fermentation ("fermented" IMO) of sucrose in the presence of a maltose acceptor mediated by a dextransucrase (Madsen et al., 2017). IMOS produced by these two methods are currently available on the commercial market in the USA (Madsen et al., 2017). The prebiotic properties of these compounds also largely depend on their size which hence needs to be tightly controlled (Madsen et al., 2017; Sarbini et al., 2011, 2013).

The α-(1→2) linkages were proved to be more resistant to the action of digestive enzymes and result in a higher total dietary fiber than the α-(1→6) ones (Valette et al., 1993; Djouzi et al., 1995; Sarbini et al., 2011) (Table 1). The products were evaluated and shown to present prebiotic properties, as well as preventive and curative action on metabolic diseases (Valette et al., 1993; Djouzi and Andrieux, 1997; Sarbini et al., 2011; Serino et al., 2012). In addition, α-(1→2) GOS production using enzyme extracts from *L. citreum* NRRL B-1299 was investigated from sucrose and maltose substrates (Paul et al., 1992). These products are commercialized under the trade name BioEcolia® by the Solabia group to be used for the protection and biostimulation of the beneficial microbial skin flora and favor natural skin defenses to the detriment of undesirable flora (http://www.solabia.com). In the same way, α-(1→3) GOS, also produced from sucrose and maltose by *L. citreum* NRRL B-742, were shown to specifically stimulate bifidobacteria, lactobacilli and *Bacteroides* growth without being metabolized by pathogen colonizers (Chung and Day, 2002, 2004).

Dextrans of higher molar masses could also be used as prebiotic agents. This was already suggested by Goulas et al. (Goulas et al., 2004a) and more recently by Das et al. who demonstrated that dextran from *Lb. plantarum* DM5 (86.5% α-(1→6), 13.5% α-(1→3) and of 1.11×10^6 g/mol) was less digestible than already commercialized inulin type prebiotics. Indeed, in vitro studies shew that it can stimulate bifidobacteria and lactobacilli growth and not *Escherichia* bacteria (Das and Goyal, 2014; Das et al., 2014).

The use of oligoalternans as low-glycemic sweeteners in food and beverage was also patented by the Cargill company (Carlson et al., 2006).

Table 1: **Dietary fiber content of oligodextrans containing α-(1→2) linkages.** Adapted from (Sarbini et al., 2011).

Compound	Average MW of backbones (Da)	Degree of α-1,2 branching (%)	Average MW of oligodextrans (Da)	Total dietary fibre (%)	DP < 3 (%)	DP > 3 (%)
Dextran 1 kDa	1000	0	1000	74	1.6	98.4
Dextran 1 kDa + 16 % α-1,2	1000	16	1190	101	0.1	99.9
Dextran 1 kDa + 32 % α-1,2	1000	32	1493	101	0.0	100
Dextran 6 kDa	6000	0	6000	62	0.4	99.6
Dextran 6 kDa + 33 % α-1,2	6000	33	8955	82	0.2	99.8
Dextran 70 kDa	70000	0	70000	88	1.4	98.6
Dextran 70 kDa + 15 % α-1,2	70000	15	82353	na	0.6	99.4
Dextran 70 kDa + 37 % α-1,2	70000	37	111111	102	2.2	97.8

Total dietary fibres of the compounds were determined by Total Dietary Fibre Analysis Method AOAC 2009.01.

All these studies revealed the potential of isomaltooligosaccharides, oligoalternans, α-(1→2) and/or α-(1→3) branched GOS, to be used as functional food ingredients. However, it must be pointed out that clinical studies are still missing and should be initiated to reinforce the claims related to prebiotics properties and action against metabolic diseases.

I.2.3. Agri-food sector

In 2001, the European Union authorized the addition of dextran (containing more than 95% of α-(1→6) linkages and a molar mass higher than $2x10^6$ g/mol) as food ingredient in bakery products, candies or ice creams (Scientific Committee On Food, 2000). In these food applications, dextrans are applied as gelling, viscosifying, thickening, emulsifying and sweetening agents, thanks to their physico-chemical and rheological properties (Kothari et al., 2014). Dextran content should not exceed 5% of the final product mass to fulfil recommendations. In 2002, the incorporation of dextran in sourdough of bakery products was patented and licensed by Puratos (Belgian society). The addition of dextrans improves structure build-up of baked products mainly by affecting the dough viscosity resulting in breads with higher volume, sweetness and crispness (Vandamme et al., 2002).

Table 2: **Food applications of dextran.** Adapted from (Kothari et al., 2014).

Applications	Properties	References
Bakery	Improves freshness, mouthfeel, softness, crumb texture, loaf volume, and shelf life	(Katina et al., 2009)
Confectionary	Improves moisture retention and viscosity and inhibits sugar crystallization and as gelling agents in gum and jelly candies	(Maina et al., 2011)
Fermented dairy products	Increases viscosity and creaminess and reduces syneresis	(Mende et al., 2013)
Ice cream	Cryoprotectant	(Naessens et al., 2005)
Cheese making: reduces-fat cheese	Improves water binding and increases moisture content in the nonfat substances	(Awad et al., 2005)

I.2.4. Analytical chemistry

In 1959, Pharmacia Fine chemicals (Uppsala, Sweden), commercialized cross-linked dextran for the purification and separation of biochemically important macromolecules as proteins, nucleic acids, and polyssacharides. This product obtained by the reaction of an alkaline solution of dextran with epichlorydrin consists in a gel of cross-linked chains which is used as molecular filter (Porath and Flodin, 1959; Flodin and Porath, 1961; Naessens et al., 2005). These hydrophilic gels are insoluble and the size of their pores can be controlled. Commercial cross-linked dextran is now known as Sephadex®. Sephadex derivatives such as carboxymethyl (CM) Sephadex, diethylaminoethyl (DEAE) Sephadex, diethyl (2-hydroxypropyl) aminoethyl (QAE) Sephadex, sulfopropyl (SP) Sephadex, Sephacryl and Superdex, also comprised dextrans and allow molecular separation by ion exchange or hydrophobic interactions for example.

I.2.5. Other applications

Among the numerous actual and potential applications of dextrans, some examples follow:

- High molar mass dextrans are used in oil and alumina recovery. In the Bayer process for the production of alumina from boxite ore, dextran is used as a flocculating agent during the purification stages (Barham and Tippett, 2004; Moody and Rushforth, 1991). In the field of oil recovery, dextran addition enhances the rheological properties of the displacing fluid ,improving oil mobility and leading to enhanced oil recovery (Wever et al., 2011).
- Mercaptodextran, a dextran derivative, could be used in environmental biotechnologies to capture heavy metal ions such as silver, mercuric, cupric and auric ions in polluted environments (Naessens et al., 2005).
- Several patents were deposited by photography companies for the incorporation of dextrans into X-ray and photographic emulsions, allowing lower silver use without quality loss (Naoi et al., 1987; Taguchi, 1999).

- As they are biocompatible, highly stable and display moisturizing properties, dextrans and their derivatives are also used in several formulations in cosmetic industry (Naessens et al., 2005).
- Dextran of *L. mesenteroides* NRRL B-1498 (Fraction L, predominantly α-(1→6) linked glucose, with approximately 3–4% branching through α-(1→3) linkages) protects steel against corrosion by forming film on the surface which likely blocks the electron transfer from the steel to acceptors, in both humid and acidic environments (Geel-Schutten, 2003; Finkenstadt et al., 2011; Ignatova-Ivanova and Ivanov, 2014).

As we have seen, dextran applications are abundant and diverse but mostly dependent of their size. While native dextrans are of interest for their physico-chemical properties and particularly their viscosity for various industrial sectors (petrol recovery, texturing agent in food …), other application fields require smaller fractions with precise sizes. For that reason, different dextran production processes have been proposed and will be presented in what follows.

I.3. Dextran production processes from sucrose

Historically, dextransucrases responsible for dextran synthesis from sucrose were predominantly isolated from strains belonging to *Leuconostoc*, *Streptococcus* and *Lactobacillus* species (Monsan et al., 2001). In the recent years, other dextran producing LAB strains such as *Pediococcus*, *Weisella* and *Oenococcus* sp. were identified and sometimes the recombinant glucansucrases involved in their synthesis characterized (Amari et al., 2013; Bounaix et al., 2010; Shukla and Goyal, 2013; Shukla et al., 2016; Vuillemin et al., 2017).

As mentioned before, the industrial production of dextran started in the early 50s to supply the armed forces in plasma volume expanders (Bixler et al., 1953). The strain *Leuconostoc mesenteroides* NRRL B-512F, isolated in 1943 from a bottle of unpasteurized root beer showing a high viscosity, was selected because it produced a dextran with more than 95% of α-(1→6) linkages and an average size over 10^6 g/mol, that revealed a low level of antigenicity (Jeanes et al., 1948). The process first described by Bixler for clinical dextran production included a first fermentation on sucrose carbon source followed by precipitation, acid hydrolysis and fractionation by methanol precipitation. A full description of this industrial process was made by Bixler et al. (Figure 4) (Bixler et al., 1953).

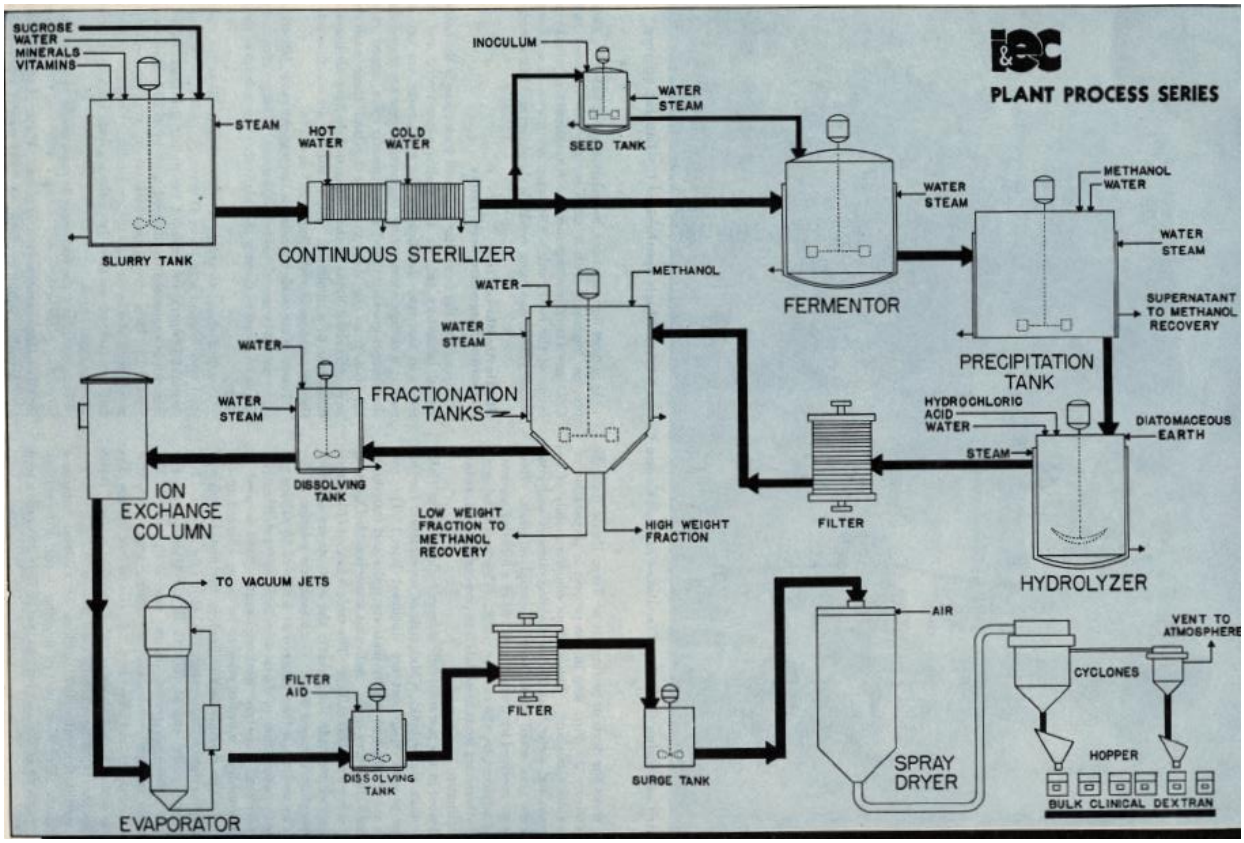


Figure 4: **Flow sheet for the production of clinical dextran.** From (Bixler et al., 1953).

As dextran applications mostly depend on their molar masses, the control of dextran size remains a key issue for industrial production. In 2005 Naessens *et al*. reported that dextran was still produced by fermentation of *L. mesenteroides* strains in media supplemented with sucrose followed by purification and when needed by fractionation steps using organic solvents such as ethanol, acetone, isopropanol for instance (Naessens et al., 2005). However, concerning dextran fractions of specific size, production costs are high and global yields are low (10 to 12%) because of important loss during the fractionation step (Khalikova et al., 2005). Solvent use is also a limit regarding pollution and recuperation issues. For those reasons, alternative processes to produce dextran of controlled size have been proposed:

- To circumvent chemical hydrolysis and solvent fractionation, enzymatic digestion with dextranase (specific enzyme for α-(1→6) linkage cleavage) has been envisaged and was patented in 1958 by Novak and Stoycos (Novak and Stoycos, 1958; Khalikova et al., 2005). *In vitro* synergistic activities of dextransucrase and dextranase for the synthesis of low molar mass dextran have also been reported in numerous studies (Gan et al., 2014; Goulas et al., 2004b; Kubik et al., 2004; Mountzouris et al., 2002; Ölçer and Tanriseven, 2010; Zhang et al., 2013). In the same idea, Kim and Day proposed a co-fermentation of *Leuconostoc mesenteroides* in the presence of sucrose with a constitutive mutant of *Lipomyces starkeyi*

producing a dextranase in order to produce a dextran of $75^{+}_{-}25$ kg/mol in good yields (75%) (Day and Kim, 1992).

- Enzymatic *in vitro* syntheses from cell extract or purified enzymes was also presented as a good alternative (Barker et al., 2007). They result in a better uniformity of the products and so facilitate its purification and the recovery of by-produced fructose. This could also allow a tight control of pH, temperature and enzyme/substrate ratio and thereby a better control of the dextran molar mass. Indeed, molar mass of synthesized dextran was shown to be related to initial sucrose concentration (generally dextran are smaller and more branched at high concentrations but it might depend on the synthesizing enzymes), to temperature and in some cases to enzyme concentration (Braswell and Stern, 1959; Braswell et al., 1962; Kim et al., 2003; Falconer et al., 2011).
- The presence of exogenous molecules (in particular acceptors of osidic nature) in addition to sucrose also enables the polymer size to be considerably reduced. This "acceptor reaction" during which exogenous molecules are preferentially glycosylated in disfavor of the polymer synthesis can also provide interesting glucan fractions. For instance, the synthesis of controlled size dextran (from 1 to 100 kg/mol) using sucrose and maltose was proposed in different studies (Hellman et al., 1955; Paul et al., 1986; Remaud et al., 1991, 1992).
- More recently, progress in structure and function studies of dextransucrases synthesizing high molar mass polymers allowed the engineered of *L. mesenteroides* NRRL B-512F dextransucrase variants producing isomaltooligosaccharides or dextrans of controlled molar mass (of about 10 and 40 kg/mol) in a single-step process (Moulis et al., 2008) (more details in the "II.4.3. Focus on domain V" paragraph and on Figure 19, p63).

Overall, α-glucans and particularly dextrans display a wide spectrum of applications in various sectors including pharmacy, medicines, cosmetics, agri-food, biomaterial and environment. All these applications are directly related to dextran physico-chemical properties, themselves determined by dextran size and 3D-organization in solution. To further develop applications, there is a clear need of investigations on the relation between dextran structure and physico-chemical properties. In addition, progress in the comprehension of dextran synthesis by dextransucrases could also open new perspectives in term of applications. In particular, optimized fermentative or enzymatic production processes using natural or engineered enzymes synthesizing tailor-made dextrans with specific size and structure depending on the targeted applications could be proposed. Such developments directly rely on the identification of the enzymatic factors allowing the control of polymer structure. We will examine the structure and mechanism of glucansucrases in detail in the next chapter and how it is possible to relate specific structures to specific activities.

II. α-transglucosylases of the GH70 family: structure/function relationships and understanding of their catalytic mechanism

II.1. Generalities

Glucansucrases (GSs) are large (140 to 313 kDa) bacterial α-transglucosylases, also named Glucosyltransferases (GTFs), which are classified in the family 70 of the Glycoside-Hydrolases (GH 70) in the Cazy database (Lombard et al., 2014), a family in which three subgroups of α-transglucosylases can be found:

- the glucansucrases, which synthesize polymers from sucrose,
- the branching sucrases, which catalyze branch formation onto linear dextran in the presence of dextran acceptor and sucrose
- the glucanotransferases, inactive on sucrose but able to disproportionate maltodextrins to form new types of carbohydrate polymers

The Family GH70 constitutes with the families GH13 and GH77 the GH-H clan, also called α-amylase super family that comprises mechanistically, structurally and evolutionary related enzymes. They share the same α-retaining mechanism involving an Asp and a Glu residues - acting, respectively, as a nucleophile and acid/base catalyst - and found in a catalytic $(\beta/\alpha)_8$ barrel domain (Davies and Henrissat, 1995; Henrissat, 1991). While the number of novel putative GH70 enzymes increases rapidly thanks to sequencing progress, the percentage of fully biochemically characterized enzymes remains low. Indeed, among the 493 sequences indexed, only 66 enzymes were biochemically characterized and no more than five structures were solved so far (http://www.cazy.org/GH70_characterized.html, Sept 2017).

The GSs and branching sucrases of GH70 family are usually extracellular enzymes reported to date to be exclusively produced by lactic acid bacteria (LAB) from *Leuconostoc*, *Streptococcus*, *Weissella*, *Lactobacillus, Pediococcus or Oenococcus* genera (Leemhuis et al., 2013a; Meng et al., 2016a; Monsan et al., 2010; Moulis et al., 2016; Shukla and Goyal, 2013; Vuillemin et al., 2017). Genomic sequence analyses of GH70 producing LAB revealed that they generally contain several GH70 encoding genes. For example, the inventory of the GH70 enzymes encoded by *Leuconostoc citreum* NRRL B-1299 allowed the identification of six distinct GSs, namely DSRA, DSRB, DSR-E, DSR-DP, DSR-M and BRS-A which are all characterized (Passerini et al., 2015). The various *gs* genes found in a strain usually encode enzymes showing different product specificities. However, extensive physiological characterization of their role has not been undertaken at this time. Thus, whether they are all produced and how is their expression regulated remain unclear. In contrast to GSs, the α-

4,6/α-4,3 glucanotransferases are more widespread in nature and can be found in a wider range of LAB bacteria from *Bacillus*, *Fructobacillus* (fructophilic LAB) genera as well as in non LAB genera such as *Geobacillus* *Exigobacterium* and *Azotobacter* (gammaproteobacteria, gram negative)(http://www.cazy.org/GH70_all.html).

Although, this literature review will mainly focus on the sucrose-active polymerases, some details on the two other subgroups will also be presented. (cf II.3. Particularity of the branching-sucrases and α-4,6/α-4,3 –glucanotransferases, p46).

II.2. Reactions catalyzed by GH70 glucansucrases

GH70 glucansucrases represent the major part of characterized GH70 enzymes active on sucrose. They usually catalyze the synthesis of long α-glucan chains (>10^6 g/mol) containing all the possible glucosidic linkages, namely α-(1→2), α-(1→3), α-(1→4) or α-(1→6). They are named according to the type of polymer they synthesize to which the suffix "ase" is added (cf I. α-glucans produced from sucrose by lactic acid bacteria). Among the variety of glucansucrases, dextransucrases represent more than 65% of the characterized GH70 to date (CAZy Database, Sept 2017).

To catalyse glucan synthesis, glucansucrases first cleave the sucrose osidic linkage through the formation of a β-D-glucosyl enzyme intermediate with a concomitant release of fructose (Figure 5). The D-glucosyl unit can be then transferred onto various acceptor molecules with retention of the α-anomeric configuration (Koshland, 1953; Uitdehaag, 1999). Depending on the nature of the acceptor, different products are obtained. The main ones result from the iterative transfer of the glucosyl unit on growing α-glucan chains, leading to oligosaccharide or α-glucan synthesis. In parallel, transfer reaction can also occur onto water leading to glucose, or fructose resulting in the formation of sucrose isomers (turanose, trehalulose, isomaltulose, and mainly leucrose). In addition, if other hydroxylated molecules are introduced in the reaction medium, they can possibly be glucosylated through the "acceptor reaction" at the expense of the polymer synthesis, the efficiency of the acceptor reaction being directly linked to the acceptor recognition (Koepsell et al., 1953; Mayer, 1987). This reaction also makes α-transglucosylases very attractive biocatalysts for the production of glucoderivatives through the glucosylation of non-natural acceptors such as various sugars (maltose, cellobiose, gentiobiose, lactose, thiooligosaccharides, …), flavonoids, salicin or, with high industrial application potential (Argüello Morales et al., 2001; Bertrand et al., 2006; Hellmuth et al., 2007; Malbert et al., 2014; Meulenbeld et al., 1999; Monsan et al., 2010; Nam et al., 2007).

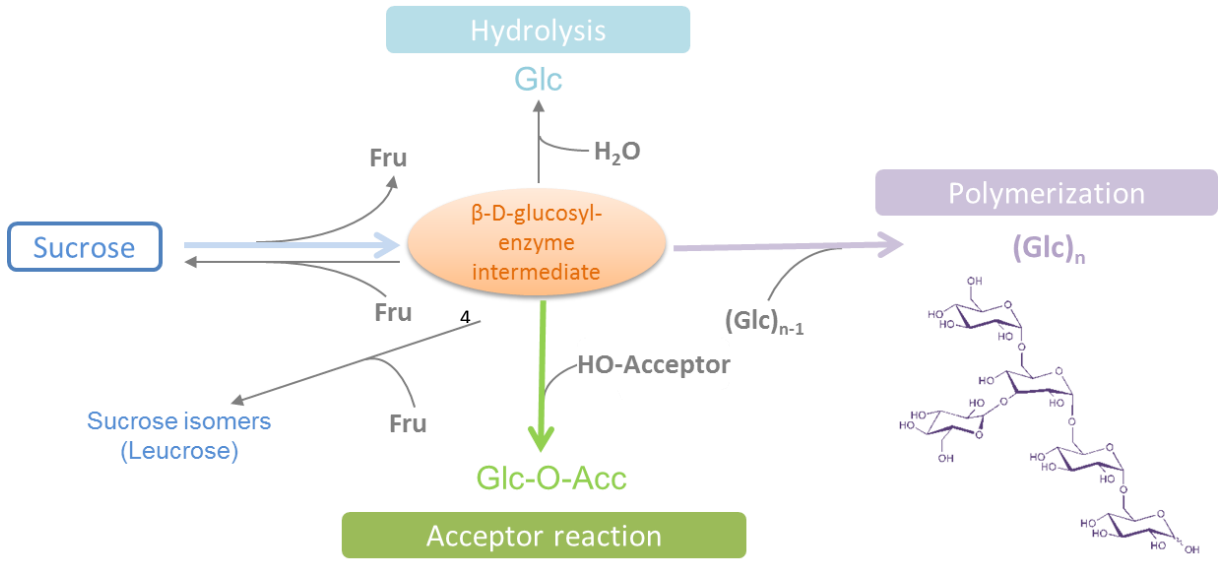


Figure 5: **Main reactions catalyzed by sucrose-utilizing glucansucrases**.

On sucrose only, the polymerization reaction is predominant, but transferase efficiency can vary from an enzyme to another. For example, *Leuconostoc* GSs are excellent polymerases for which the transferase activity is above 85% (Moulis et al., 2006a). In comparison, the hydrolytic activity of GTF-O reuteransucrase from *Lactobacillus reuteri* ATCC 55730 corresponds to more than 50% of the transferred glucosyl units (reuteran synthesis represents only 36%) whereas GTF-A, another reuterasucrase from *Lactobacillus reuteri*, converts 73% of the glucosyl unit from sucrose into polymer (Kralj et al., 2005a; Meng et al., 2016b).

α-transglucosylases also have the ability to catalyze disproportionation reactions (Binder et al., 1983) which consist of the following reaction: $2(glucose)_n \rightarrow (glucose)_{n+1} + (glucose)_{n-1}$. A novel GH70 enzyme category catalyzing only reactions close to the disproportionation mechanism was recently identified and named α-4,6-/α-4,3-glucanotransferase. These enzymes, non-active on sucrose but on maltooligosaccharides constitute a GH70 family subgroup as well as the branching sucrase and will be presented in the following paragraph.

II.3. Particularity of the branching-sucrases and α-4,6/α-4,3 –glucanotransferases

Although glucansucrases are the main representative enzymes of the GH70 family, two new enzyme subgroups were recently identified in this family, namely the branching sucrases and the glucanotransferases. Members of these subgroups present unique catalytic properties compared to glucansucrases.

II.3.1. Brief overview of the branching-sucrases

For a long time, GBD-CD2 was the sole known branching-sucrase of the GH70 family. This is however a non-natural enzyme obtained after domain truncation of the peculiar DSR-E from *L. citreum* NRRL

B-1299, the unique bi-functional GS characterized to date which displays two catalytic domains separated by a glucan-binding domain (GBD) (Figure 6). From sucrose, recombinant DSR-E synthesizes a polymer of α-(1→6) linked glucopyranosyl units with α-(1→2) branches (Bozonnet et al., 2002; Fabre et al., 2005).

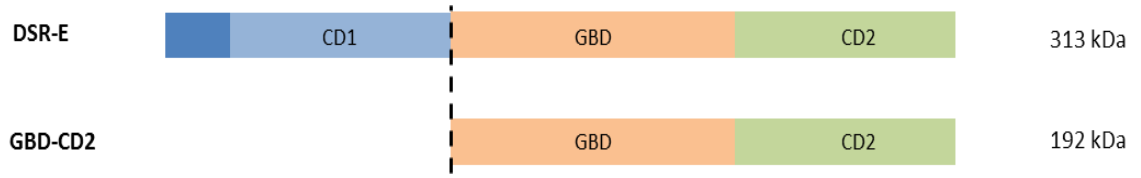


Figure 6: **Schematic structure of DSR-E from *L. citreum* B1299 and GBD-CD2 truncated variant**. CD, Catalytic Domain; GBD, Glucan Binding Domain.

By generating and characterizing two truncated forms, namely CD1-GBD and GBD-CD2, the first catalytic domain was shown to be responsible for the synthesis of α-(1→6) linked α-glucan while GBD-CD2 is only able to catalyze α-(1→2) glucosyl transfer onto linear dextran. From sucrose only, GBD-CD2 acts as a sucrose hydrolase and is incapable of polymerizing the glucosyl units. Indeed, 78% of the available glucosyl units are transferred onto water and 16% onto fructose to produce leucrose (α-D-Glcp-(1→5)-D-fructopyranose), kojibiose (α-D-Glcp-(1→2)-D-Glcp), and traces of maltulose (α-D-Glcp-(1→4)-D-fructofuranose). Nevertheless, in the presence of an exogenous dextran acceptor, this enzyme almost exclusively catalyzes polymer glucosylation through the formation of α-(1→2) linkages (Figure 7). Moreover, it has been shown that the percentage of α-(1→2)-linked glucose can be controlled from 10 to 40% by playing on the initial [sucrose]/[dextran] ratio, thus conferring a valuable potential to this original catalyst for the production of a wide diversity of branched α-glucans (Brison et al., 2010).

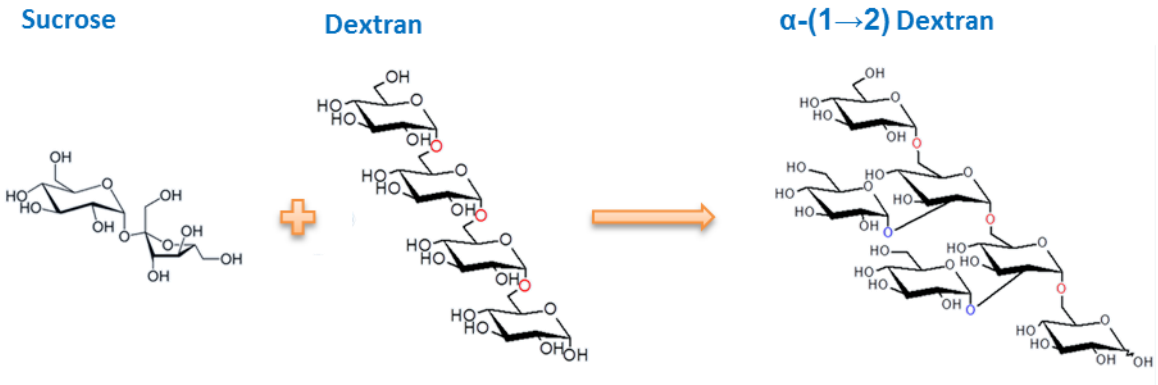


Figure 7: **Dextran glucosylation by the branching-sucrase GBD-CD2**. Oxygen atoms of α-(1→6) glucosidic bonds are in red while α-(1→2) ones are in blue.

Recently, natural branching-sucrase activities were identified in *Leuconostoc* and *Lactobacillus* strains highlighting their more spread occurrence and diversity. Contrary to DSR-E, they display a "classic" primary structure similar to that of polymerases with only one catalytic domain and a GBD (see paragraphII.4.1. Global structure and domain organization of GH70 enzymes, p49) but they were proposed to constitute a new subgroup of the GH70 family as they used two substrates. This group houses now five characterized enzymes, namely, the above described GBD-CD2, BRS-A (Passerini et al., 2015), BRS-B, BRS-C, and BRS-D (Vuillemin et al., 2016). As DSR-E or the truncated version GBD-CD2, BRS-A is originated from *L. citreum* NRRL B-1299 and is responsible for the high content of a-(1→2) linkages (28%) observed for the dextran produced by this strain (Passerini et al., 2015). BRS-B encoding gene was isolated from the *L. citreum* NRRL B-742 genome and the recombinant protein was shown to display an α-(1→3) branching specificity. As for GBD-CD2 or BRS-A, playing on the [sucrose]/[dextran] initial ratio allows a tight control of the polymer branching degree up to 50%, meaning that each glucose of the linear dextran is glucosylated. For that reason, BRS-B was proposed to be responsible of the comb-like hyper-branched dextran of *L. citreum* NRRL B-742 synthesis. BRS-C from *Leuconostoc fallax* KCTC 3537 and BRS-D from *Lactobacillus kunkeei* EFB6 catalyze α-(1→3) and α-(1→2) transglucosylation, respectively. Sequence analyses revealed that these branching-sucrases share common motifs with GH70 GSs but also display distinctive features. This is particularly the case in the +1 binding subsite, where Phe675 of motif II and Ile783, His785, Lys789, and Val795 of motif IV (BRS-B numbering) are only conserved in the branching sucrases (for subsites and motifs description, see paragraph: II.4.2.1. Structure of the catalytic core, p51). These particularities can be used to rapidly sort out branching sucrases from genomic data (Vuillemin et al., 2016).

The determinants responsible for their specificity to branch dextrans and incapacity to catalyze polymer formation have not clearly been identified yet, although the structure of GBD-CD2 has been solved. To this aim, further structure-function studies are necessary. However, as they are more specific for acceptor reaction and are able to catalyze a variety of glucosidic linkage type, branching-sucrases appear to be promising for new functional glucooligosaccharide or glucoconjugate synthesis.

II.3.2. α-4,6- and α-4,3-glucanotransferases

The first GS-like enzyme of this subgroup was found in *L. reuteri* 121 and named GtfB. Being non active on sucrose but only on maltodextrins and starch substrates, GtfB catalyzes α-(1→4)-linkage cleavage and the formation of new α-(1→6)-linkages (see Figure 15, p58). This α-4,6-glucanotransferase activity results in the production of α-glucans with linear chains of α-(1→6) and α-(1→4)-linkages (Kralj et al., 2011). Lately, two other GtfB homologues from *Lactobacillus reuteri* strains DSM 20016 (GtfW) and ML1 (GtfML4) displaying the same donor/acceptor specificity were

also characterized (Leemhuis et al., 2013b). α-4,6-glucanotransferase activity can also be predicted by sequence analysis thanks to signature residues in acceptor binding subsites +1 and +2 differing from those found in GH70 GSs (Leemhuis et al., 2013b). Beside these three biochemically characterized enzymes, 46 putative GtfB-like enzymes are currently found in the GenBank database, and they are almost all from *Lactobacillus* genus (three are found in *Pediococcus* strains). The first α-4,6-glucanotransferase structure, that of GtfB-ΔNΔV from *Lactobacillus reuteri* 121 (PDB: 5JBD), was solved last year (Bai et al., 2017). Thanks to structural and phylogenetic analyses, notably focusing on three loops shaping the active site, the enzymes of this subfamily are suggested to be evolutionary intermediates between GH13 amylosucrases and GH70 sucrose-active enzymes (see Figure 11, p54).

GtfC of *Exiguobacterium sibiricum* 255–15 (Gangoiti et al., 2016a) and GtfD of *Azotobacter chroococcum* NCIMB 8003 (Gangoiti et al., 2016b) are also both inactive on sucrose but display α-4,6-glucanotransferase activity that resulted in the synthesis of isomalto/malto oligosaccharides (IMMO) and of a reuteran-like α-glucan, respectively. The domain organization of GtfC and GtfD resembles that of GH13 family enzymes with a conserved order of motifs I-II-III and IV (contrary to GH70s) and an absence of domain V (see paragraph: II.4.2.1. Structure of the catalytic core, p51, for description of these motifs). Thus, GtfC and D are also proposed to be representative of two other GH70 subfamilies.

Most recently, genome sequencing of *Lactobacillus fermentum* NCC 2970 resulted in the discovery of a novel GtfB-like enzyme presenting divergence in the conserved motifs II and IV located in donor/acceptor binding subsites. This enzyme was further shown to possess an α-4,3-glucanotransferase activity meaning that it cleaves α-(1→4)-linkages and synthesizes new α-(1→3)-linkages (Gangoiti et al., 2017). All these recent findings - due to progress of sequencing- highlight the diversity existing in GH70 family in terms of substrate and linkage specificity.

II.4. Structures of GH70 family enzymes and understanding of their mechanism

II.4.1. Global structure and domain organization of GH70 enzymes

Until the first 3D-structure characterization, the primary structures of GH70 enzymes were traditionally represented in three different domains linearly organized along the polypeptidic chain and preceded by a signal peptide (Figure 8). The sequence was proposed to be composed of i) a poorly conserved N-terminal variable region, ii) a catalytic domain of around a thousand of very conserved residues and iii) a C-terminal domain comprising repeated units proposed to be likely involved in glucan binding (see paragraph, II.4.3.1. Structural organization of domain V, p58).

Figure 8: **General primary structure of GS from LAB as depicted before the first X-ray 3D structure resolution of a GH70 glucansucrase.** SP signal peptide, VR variable region, CD catalytic domain, GBD glucan binding domain.

Since 2010, attempts to solve the tridimensional structure of GH70 enzymes remained unsuccessful probably due to the large molar mass of these enzymes added to an important flexibility rendering protein purification and crystallization very challenging. When the first 3D-structure was solved, the previously proposed linear organization had to be corrected. Indeed, the structure of GTF180-ΔN from *L. reuteri* (PDB: 3KLK) revealed an atypical U-shape fold resulting in an organization in five domains, namely the domains A, B, C, IV and V (Vujicic-Zagar et al., 2010a). Domain C is the sole domain composed of a continuous polypeptide fragment – forming an eight-stranded β-sheet with a Greek key motif – and constitutes the base of the "U" but its precise role remains unknown. Domains A, B, IV and V consist in a tangle of both N- and C-terminal segments (Figure 9). While domains IV and V are unique to GH70 enzymes, domains A, B and C are common with GH13 family members. The active site is localized in a cavity of domain A at the interface with domain B (Figure 9). Domain A adopts a $(\beta/\alpha)_8$ barrel fold circularly permuted compared to that of the GH13 enzymes (cf. next paragraph). Domain B is folded in five-stranded β-sheets and some of its loops participate in the catalytic cleft shaping. A calcium binding site is found at its interface with domain A at around 10 Å of the nucleophilic aspartate. Domain IV connects domains B and V and displays a particular fold showing no similarity to any other characterized protein structures. As the connection between domains IV and V is made of two relatively long unstructured polypeptide chains, Ito et al. (Ito et al., 2011) proposed that domain IV may serve as a "hinge" allowing domain V movement near or away from the catalytic core. This was further supported by domain V flexibility analyses (Pijning et al., 2014)(see paragraph, II.4.3.1. Structural organization of domain V, p58). The secondary structure elements organization of domain V is very similar from an enzyme to another and their role in glucan binding was recently structurally confirmed (Brison et al., 2016) (cf section "II.4.3. Focus on domain V", p58 for detailed description).

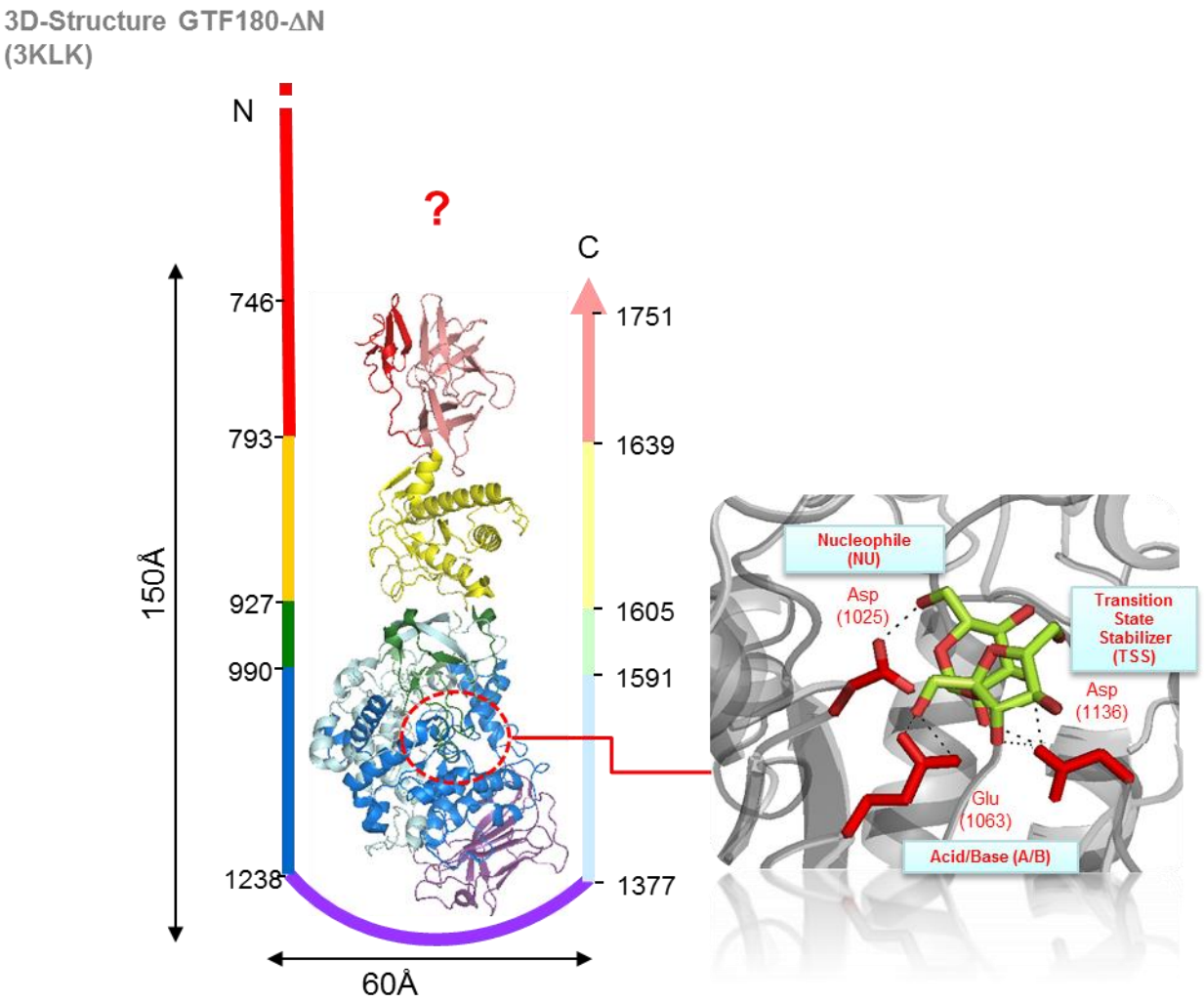


Figure 9: **Three-dimensional structure of GTF-180ΔN, the first solved structure in GH70 family showing a five domain organization.** Red, domain V; yellow, domain IV; green, domain B; blue, domain A; purple, domain C. A zoom on the catalytic site with the catalytic triad in complex with sucrose (PDB: 3HZ3) is also shown.

Since that time, four other 3D-structures of recombinant truncated GH70 enzymes sharing the same features were solved: the reuteransucrase GTFA-ΔN from *L. reuteri* 121 (Pijning et al., 2012), the mutansucrase GTF-SI from *Streptococcus mutans* (Ito et al., 2011), the branching sucrase ΔN123-GBD-CD2 engineered from *L. citreum* NRRLB-1299 DSR-E (Brison et al., 2012) and most recently, the α-4,6-glucanotransferase, Gtf-B-ΔNΔV from *L. reuteri* 121 (Bai et al., 2017).

II.4.2. Catalytic core structure and reactional mechanism

II.4.2.1. Structure of the catalytic core

The catalytic core comprising domains A, B and C of GH70 members is very similar to that of GH13 family. However, except for GtfC and GtfD (see p48), the GH70 $(\beta/\alpha)_8$ barrel fold of domain A is generally circularly permuted compared to that of GH13 family (MacGregor et al., 1996; Mooser and Wong, 1988). This $(\beta/\alpha)_8$ barrel consists of eight parallel β-strands (β1-β8) forming the inner β-barrel

and alternating with eight α-helices (α1-α8) surrounding the β-strands. Due to the circular permutation, from N- to C-terminus, GH70 $(\beta/\alpha)_8$ barrel starts with helix α3 of family GH13 followed by Nter-β4-α4-β5-α5-β6-α6-β7-α7-β8-α8-β1-α1-β2-α2-β3-Cter (Figure 10).

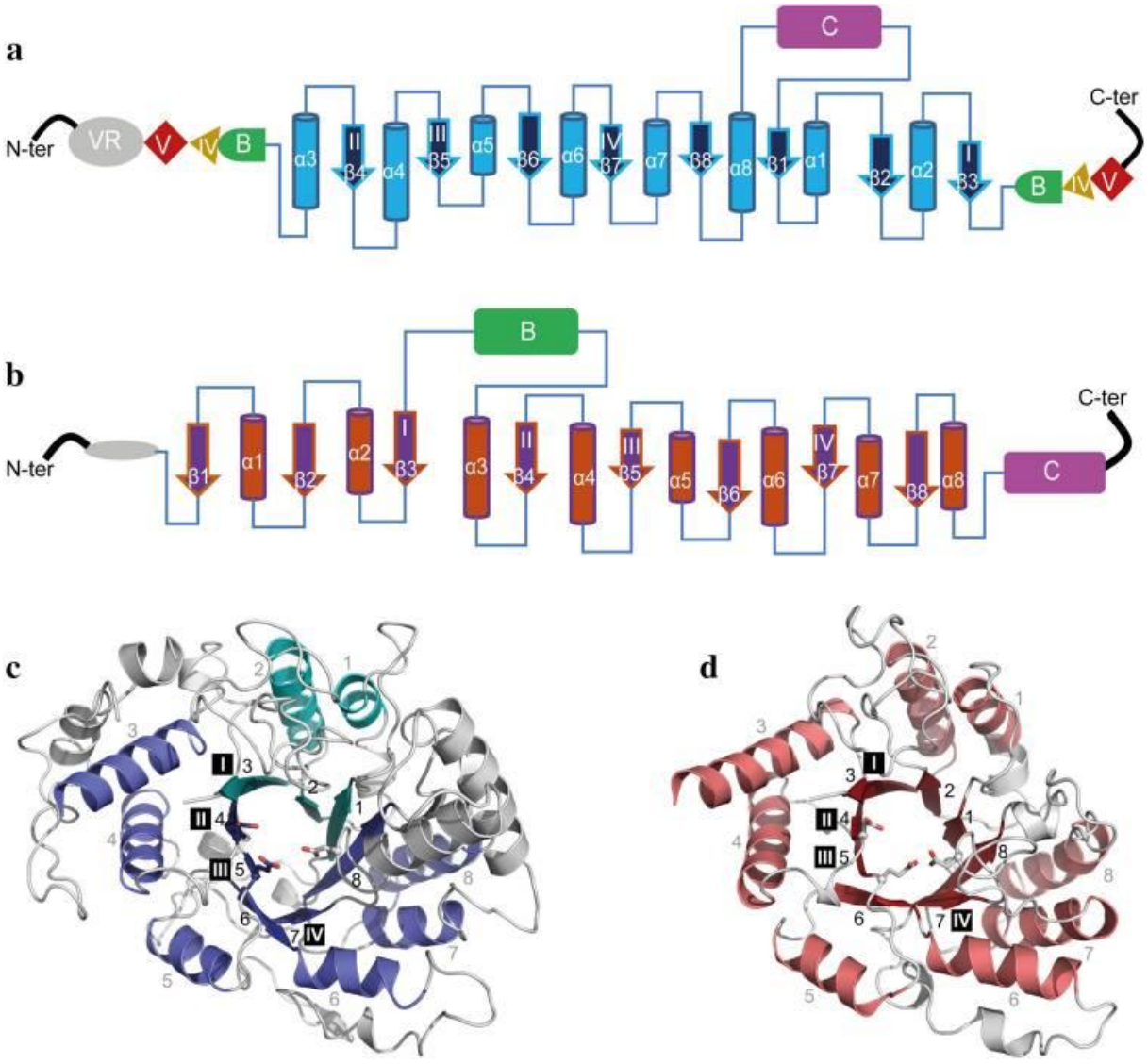


Figure 10: **Topology diagrams models of family GH70 GSs with a circularly permutated $(\beta/\alpha)_8$ barrel (a) and the family GH13 α-amylase $(\beta/\alpha)_8$ barrel (b)**. Cylinders represent α-helices and arrows represent β-strands. The equivalent α-helices and β-strands in GH70 GSs and GH13 α-amylases are numbered the same. The different domains in GH70 and GH13 enzymes are indicated. The four conserved sequence motifs (I–IV) which are located in β-strands 3, 4, 5, and 7, respectively, are indicated within the β-strand. The structure of the catalytic domain in the GH70 GSs representative GTF180-ΔN (c, PDB: 3KLK) of *L. reuteri* 180 and in the GH13 representative α-amylase of *Bacillus licheniformis* (d, PDB: 1BPL). From (Meng et al., 2016a).

Accordingly the conserved motifs characteristic of GH13 enzymes are also permuted in GH70 proteins, with motif I located at C-terminal to motifs II, III and IV. All are located in domain A. Residues of the catalytic triad, namely, the nucleophile Asp1025, the acid/base catalyst Glu2063 and the transition state stabilizer Asp1136 (GTF-180 numbering) are located in the conserved motifs II, III

and IV, respectively (Figure 9 and Figure 10). Thanks to previous works on structure-function relationships, GH70 specificity can be predicted by analysis of these "signature" motifs (Table 3). Indeed, it has been shown that amino acids upstream and downstream the catalytic residues are involved in the linkage type specificity (see paragraph, II.4.4.2. Linkage specificity, p67).

Table 3: **Sequence alignment of functional conserved motifs I–IV of GH70 enzymes.** The numbering order of the motifs (II, III, IV and I) refers to the motifs originally defined in the family GH13.

		Motif II		Motif III		Motif IV		Motif I	Specificity
GTF-I [Sd]	449	SIRVDAVDNVD	486	HVSIVEAWSDN	559	FARAHDSEVQDLIRD	931	ADWVPDQ	
GTF-I [Ss]	443	SIRVDAVDNVD	480	HVSIVEAWSDN	553	FARAHDSEVQDIIRD	925	ADWVPDQ	α-(1→3)
GTF-B [Sm]	1011	SIRVDAVDNVD	1048	HLSILEAWSDN	1120	FIRAHDSEVQDLIAD	1488	ADWVPDQ	
GTF-A [Lr]	1020	SVRVDAPDNID	1056	HINILEDWNHA	1128	FVRAHDNNSQDQIQN	1508	ADWVPDQ	α-(1→4) / α-(1→6)
GTF-O [Lr]	1020	SVRVDAPDNID	1056	HINILEDWNSS	1128	FIRAHDNNSQDQIQN	1508	ADWVPDQ	
ASR [Lm]	631	GIRVDAVDNVD	668	HLSILEDWNGK	762	FVRAHDYDAQDPIRK	1168	ADWVPDQ	α-(1→6) / α-(1→3)
GTF-W [Lr]	748	GFRVDAADNID	785	HLVYNEGYHSG	568	FVTNHDQR-KNVINQ	1216	EDLVMNQ	α-(4→6)
GTF-ML4 [Lr]	1012	GFRVDAADNID	1049	HLSYNEGYHSG	1121	FVTNHDQR-KNLINR	1479	EDIVMNQ	
DSR-CB4 [Lc]	526	GIRVDAVDNVD	563	HLSILEDWSHN	636	FVRAHDSEVQTVIAQ	1001	ADWVPDQ	
DSR-C [Lm]	498	GIRVDAVDNVD	535	HLSILEDWSHN	608	FVRAHDSEVQTVIAQ	973	ADWVPDQ	α-(1→6)
DSR-S [Lm]	547	GIRVDAVDNVD	584	HLSILEDWSHN	657	FVRAHDSEVQTVIAQ	1023	ADWVPDQ	
GTF-180 [Lr]	1021	GIRVDAVDNVD	1058	HINILEDWGWD	1131	FVRAHDSNAQDQIRQ	1503	ADWVPDQ	
DSR-E (CD1)	520	GYRVDAVDNVD	557	HISILEDWDNN	630	FIRAHDSEVQTVIAQ	1010	NDWVPDQ	α-(1→6)
subsite -1 / subsite +1		NU		AB		TSS			

Structurally, the catalytic domains of the GH70 structures solved so far are very similar to each other. However, some distinctions are observable notably on the loops delineating the catalytic cavity (loops, A1, A2 and B). Although differences are present between the sucrose-active enzymes, it is particularly the case for the α-4,6-glucanotransferases in which loop A2 is similar to its equivalent in GH13 enzymes while loop A1 is closest to GH70 members (Figure 11). The structural rearrangement of these loops is likely responsible of the evolution of catalytic activities from GH13 to GH70 enzymes and was proposed to be a consequence of diet evolution (Bai et al., 2017).

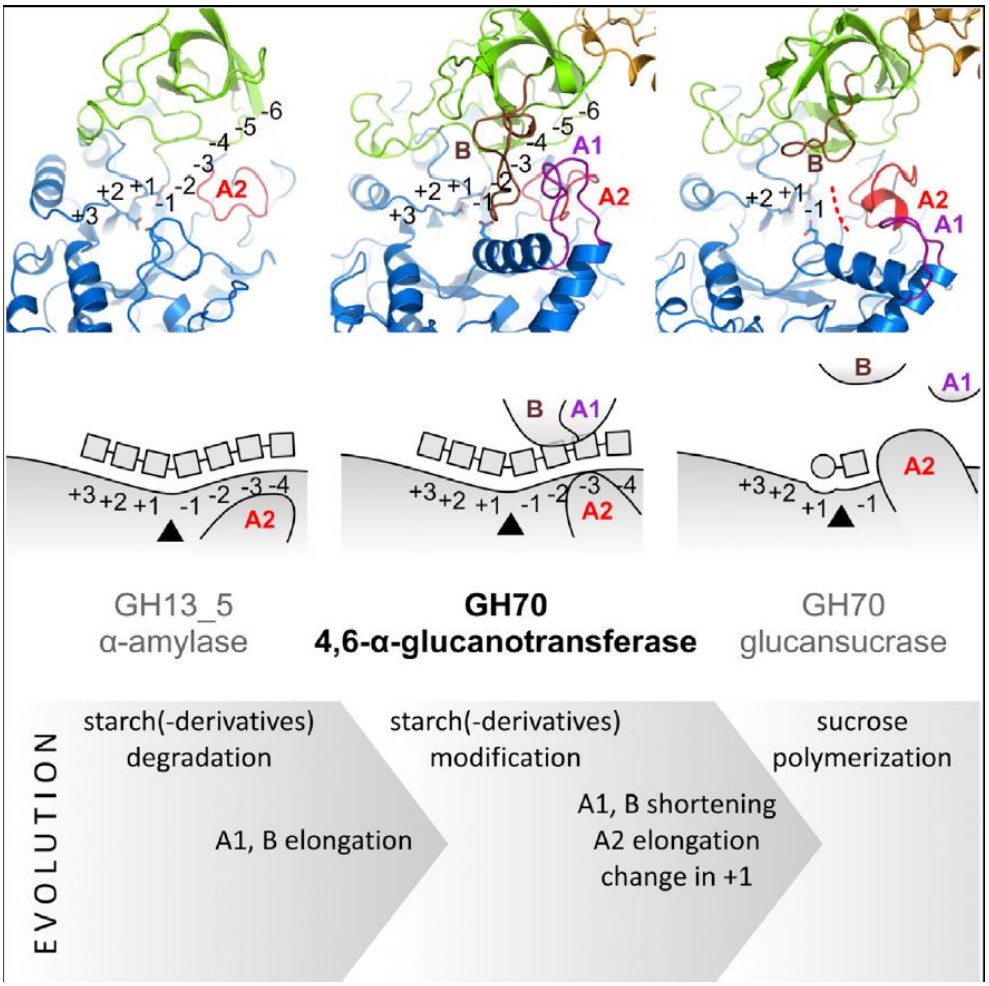


Figure 11: **Structural changes accompanying the proposed evolution of GH70 glucansucrases from starch-acting GH13 amylases in oral bacteria.** Arrangement of loops A1, A2, and B around the active site of GH13 α-amylases (left), α-4,6-glucanotransferases (middle), and GH70 glucansucrases (right), at the interface of the catalytic domain A (blue) and domain B (green). Looking down the $(\beta/\alpha)_8$ barrel, acceptor subsites are on the left and donor subsites on the right; catalytic residues are shown as sticks. Structures used are *Bacillus licheniformis* α-amylase (PDB: 1BLI), GtfB-ΔNΔV (PDB: 5JBD), and *Lactobacillus reuteri* 180 GTF180-ΔN (PDB: 3KLK). From (Bai et al., 2017).

Phylogenetic analyses of the GH70 enzymes (although it might be biased by the low number of available sequences) seem to corroborate this hypothesis. Indeed the α-4,6-glucanotransferases cluster in a distinct branch (Figure 12). This is also the case for the branching sucrases for which subtree is divided into two further branches, one corresponding to enzymes displaying an α-(1→2) linkage specificity and the other to enzymes with α-(1→3) linkage specificity. Interestingly, although that it remains to be explained, these two subgroups (branching-sucrase and α-4,6-glucanotransferases) are phylogenetically closest to each other compared with the rest of the enzymes from GH70 family.

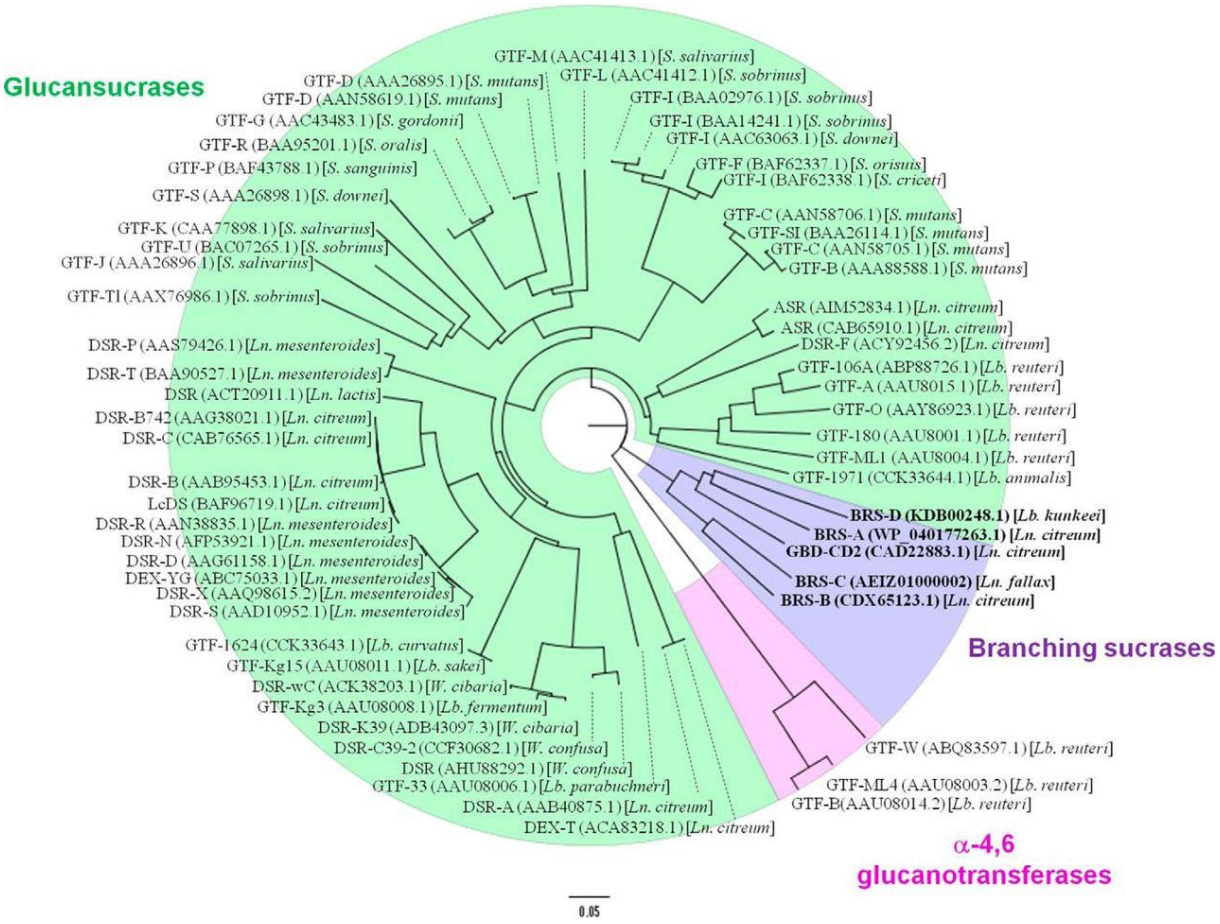


Figure 12: **Phylogenetic tree of GH70 enzymes. Each enzyme is labeled with its GenBank™ accession number and its origin.** The species of microorganism are indicated with *Ln.* for *Leuconostoc*, *Lb.* for *Lactobacillus*, *S.* for *Streptococcus*, and *W.* for *Weissella*. Glucansucrase enzymes are highlighted in *green*, α-4,6 glucanotransferases are in *pink*, and branching sucrases are in *bold* and *purple*. The *scale bar* corresponds to a genetic distance of 0.05 substitutions per position. From (Vuillemin et al., 2016).

II.4.2.2. Catalytic mechanism

The tridimensional characterization of GH70 members also allowed validating the catalytic mechanism of these enzymes. Indeed, subsites -1, +1 and +2 implicated in substrate and acceptor recognition – as defined by Davies in 1997 (Davies et al., 1997) –were identified in the structure of GTF-180-ΔN in complex with either sucrose or maltose (PDB: 3HZ3 and 3KLL) (Vujicic-Zagar et al., 2010a) (Figure 13).

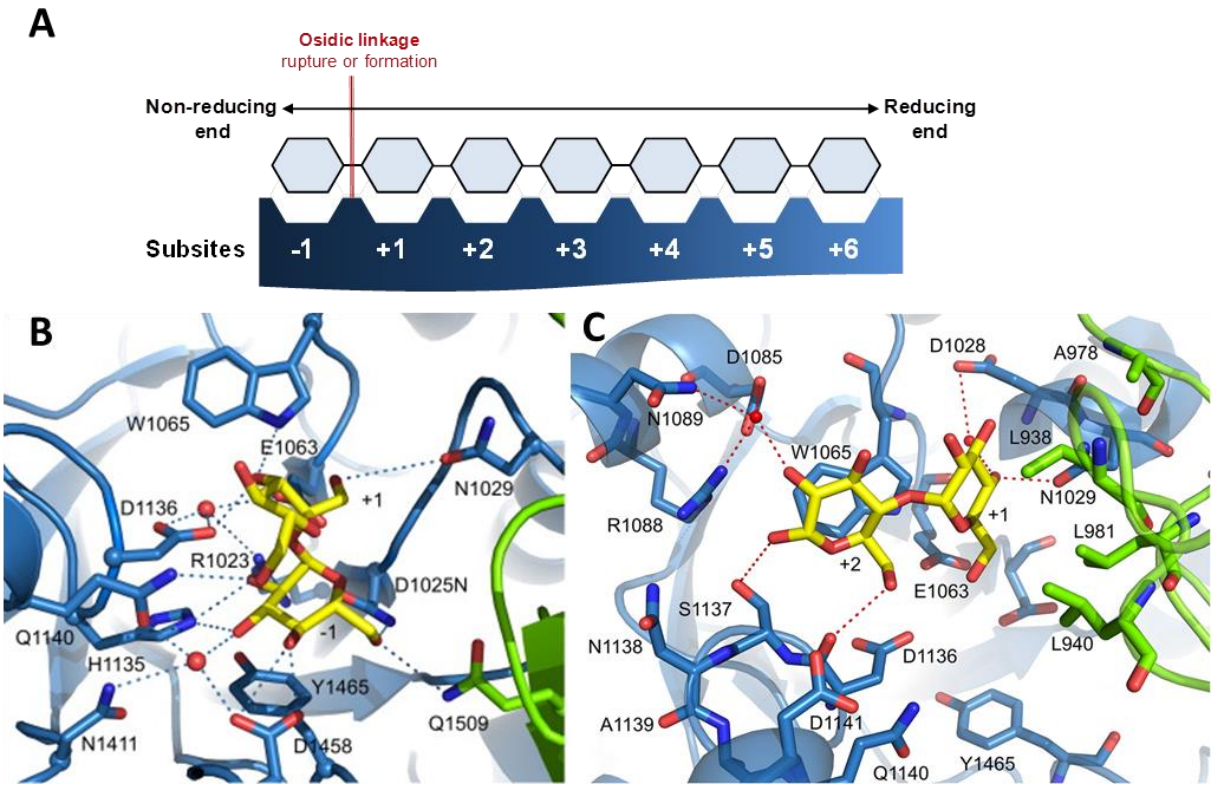


Figure 13: (A) **Schematic representation of the acceptor binding subsites, according to Davies proposed nomenclature**. From (Davies et al., 1997). (B and C) **Views of sucrose and maltose binding in the active site of GTF180 (D1025N)-ΔN.** (B) Sucrose (white stick representation) bound at subsites -1 and +1 in the D1025N mutant. Residues of the 980–984 loop are depicted as a transparent cartoon for clarity. Strictly conserved residues in the α-amylase superfamily are labeled in bold, except for D1504 which is not shown; (C) Only Maltose M1 (yellow stick presentation) bound in subsites +1 and +2 is shown. The glucosyl moiety of the sucrose from the GTF180-ΔN-sucrose complex (white stick representation) is superimposed. Leu938 in front of the +1 glucosyl residue was omitted for clarity. Residues from domains A and B are colored blue and green, respectively. From (Vujicic-Zagar et al., 2010a) and (Meng et al., 2016a).

Thus, the -1 subsite is defined by seven strictly conserved residues interacting with the glucosyl unit of sucrose, namely E1063 (A/B), D1025 (NU), D1136 (TSS), H1135 (motif IV), R 1023 (motif II), Y1465 (domain A) and Q1509 (motif I), (GTF180 numbering) (Figure 13). Four residues E1063 (A/B), W1065 (motif III), N1029 (motif II) and Q1140 (motif III) interact with the sucrose fructosyl moiety and are less conserved among GH70 members. Indeed, F2214 of GBD-CD2 replaces N1029, in addition, A2249 and G2250 of GBD-CD2 are found in the corresponding space occupied by W1065 in GTF180. Concerning α-4,6-glucanotransferases, a tyrosine and a lysine are found instead of W1065 and Q1140, respectively. Additionally to fructose accommodation, subsite +1 is also involved in the recognition of a variety of acceptors with different structures demonstrating a wide plasticity of this subsite. Its architecture as well as that of subsite +2 (Figure 13) is a determinant of the enzyme linkage specificity (Leemhuis et al., 2012).

The GH70 enzymes all adopt an α-retaining mechanism as proposed by Koshland (Koshland, 1953) and shown on Figure 14 and Figure 15, respectively. Concerning the sucrose-active enzymes, the reaction starts with the sucrose binding. The first step (glucosylation) leads to the formation of the β-D-glucosyl enzyme intermediate and the release of fructose through the formation of an oxocarbenium implicating the attack of the glucosyl unit anomeric carbon by the aspartate nucleophile residue and the protonation of the oxygen atom of the sucrose osidic linkage. The β-D-glucosyl enzyme covalent intermediate is further stabilized by a second aspartic residue (not shown on Figure 14). During the second step (deglucosylation), the deprotonated glutamic acid (acting as a base) activates an acceptor molecule which attacks the C1 of the glucosyl ring leading to either hydrolysis or transglucosylation when either water or a sugar act as acceptor, respectively. The linkage specificity is governed by the orientation of the incoming acceptor substrate driven by residues surrounding the positive subsites.

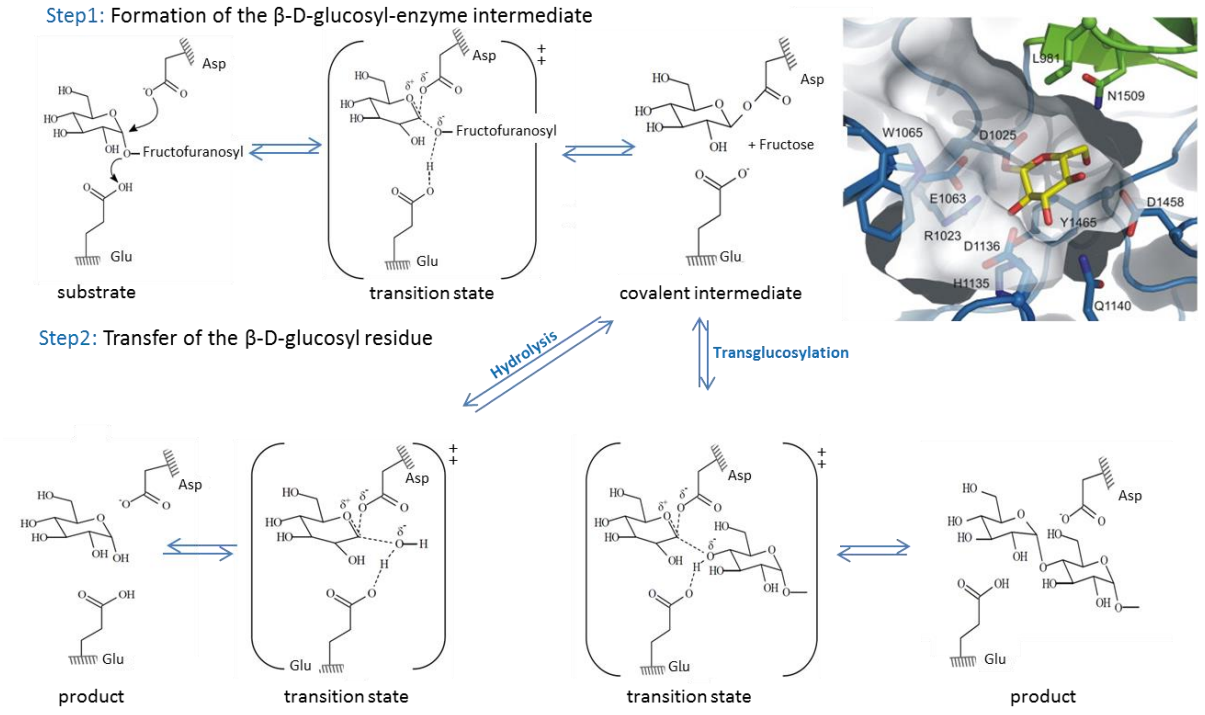


Figure 14: **The two-step α-retaining-type mechanism of glucansucrases** (adapted from Skov et al., 2001 and Withers, 2001). Nucleophile (Asp) and acid/base (Glu) catalysts are shown. (Up right) Model of the covalent glucosyl-enzyme intermediate in GTF180-ΔN (The glucosyl moiety is shown with yellow carbons; residues from domain A are shown with blue carbons and residues from domain B with green carbons) (Leemhuis et al., 2013a).

Regarding the α-4,6-glucanotransferases acting on starch and maltodextrines and proposed to act not only exo- but also endolytically (Figure 15), the donor substrates are proposed to occupy the subsites -3 (or -4) to +2. After binding, the donor substrates are cleaved by the acid/base residue that attacks on the glucosidic bond between subsites -1 and +1, accompanied by the nucleophile attack of the catalytic aspartate on the C1 of the glucosyl moiety in subsite -1, resulting in the formation of a

β-glycosyl-enzyme covalent intermediate. The deglucosylation step is then similar to that adopted by the GH70 sucrose-active enzymes.

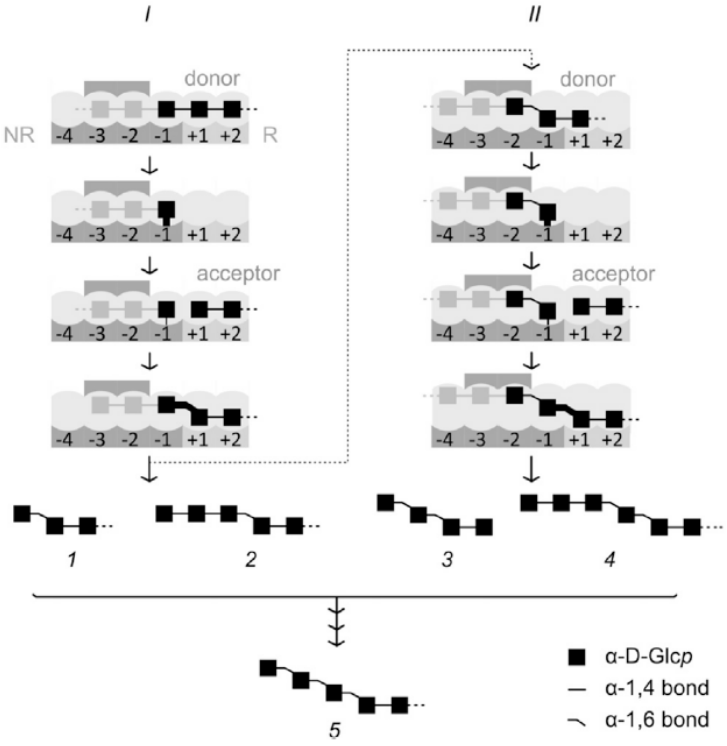


Figure 15: **Proposed Reaction Pathways of α-4,6-glucanotransferases.** R, reducing end; NR, non-reducing end; four of the six donor subsites are indicated, of which subsites -2 and -3 are in the tunnel. Newly formed bonds are indicated with thicker lines. Binding of maltooligosaccharide donor substrates (panel I) involves a single subsite, or multiple ones (gray units) in the tunnel, yielding mixed isomalto-/maltooligosaccharide transglycosylation products of types 1 and 2, respectively. Type 2 products can become donor substrates for a next cycle (panel II); whether they first diffuse out or stay bound to the enzyme and shift is currently unclear. The resulting transglycosylation products are of types 3 and 4. Eventually, all products are mixed isomalto-/maltooligosaccharides with α-1,6 linked glucosyl units at their non-reducing end.

II.4.3. Focus on domain V

II.4.3.1. Structural organization of domain V

Domain V is generally built up of both N- and C-terminal extremities of the polypeptide chain, except for domain V of the engineered DSR-E ΔN123-GBD-CD2, which only contains a N-terminal fragment (Figure 9). Since the nineties, domain V was shown to be involved in glucan chain fixation (Mooser and Wong, 1988; Abo et al., 1991; Monchois et al., 1998a, 1999; Suwannarangsee et al., 2007) and was accordingly called glucan binding domain (GBD). Often described to be located essentially at C-term extremity, GBD refers in fact to the whole domain V constituted of both extremal peptidic segments found at the N and/or C-term of the protein. The first evidence of the presence of non-catalytic domains showing glucan binding ability in GH70 enzymes was provided with polypeptide segments isolated by tryptic digestion of various streptococcal GSs (Mooser and Wong, 1988).

Further sequence and biochemical analyses revealed that these glucan binding domains were composed of repeated units, which were firstly classified into four groups (A, B, C and D) and then were proposed to have evolved from a common YG repeat. This motif is characterized by conserved aromatic amino acids and glycine ((N/x)x(D/x)x(G/x)ϕYϕϕxxxxGxxΨx(G/N/x)x(Ψ/x)Ψ(Ψ/x) where x represents a non-conserved residue, Ψ a hydrophobic residue, ϕ an aromatic residue). YG-repeats can be found at both N- and C-terminal part of domain V (Kralj, 2004a; Van Hijum et al., 2006) and share homology with the cell wall binding repeats of choline-binding proteins or toxins (Pfam CW_binding 1 (PF01473)) (Fernández-Tornero et al., 2001; Greco et al., 2006). From sequence similarities with these proteins, a model of the C-terminal segment of DsrP presenting a β-solenoid architecture was proposed (Olvera et al., 2007). Later on, from 3D-structure analyses, GH70 YG-repeats were shown to adopt a similar β-solenoid fold. Indeed, although the structures of GSs obtained so far only define part of the domain V (absence of around 700 residues), they revealed that a common structural module composed of YG-repeats as well as A- and C-type repeats is present in multiple copies (Figure 16). These structural modules consist of two or three consecutive β2/β3 units being approximately rotated by 60° one to another and comprising around 20 residues which form either a β-hairpin or a 3-stranded antiparallel β-sheet. Their regular arrangement along the domain V results in a global β-solenoid fold. Added to a common topology, these modules also share sequence similarities (Figure 16).

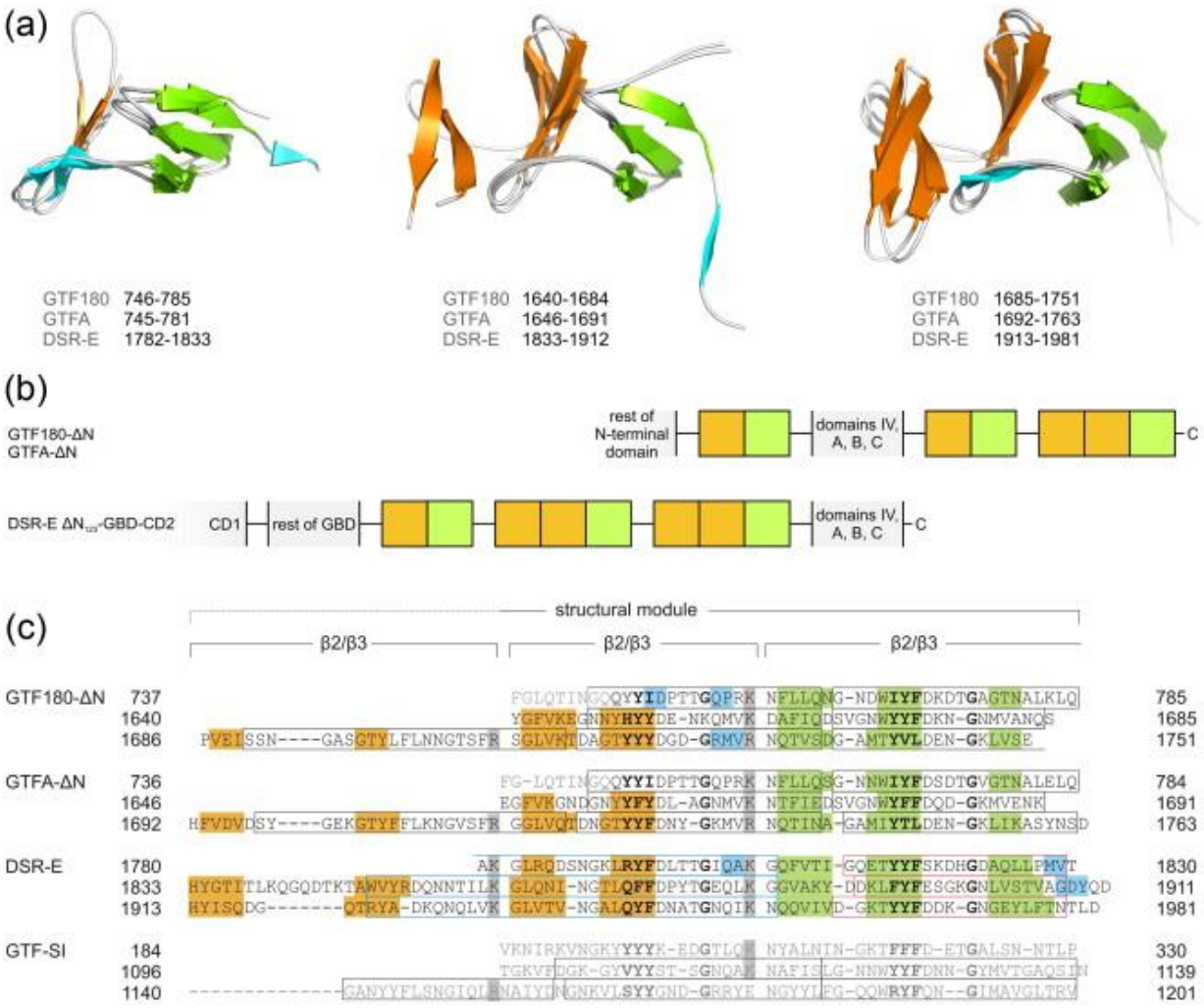


Figure 16: **Organization and combined structure/sequence alignment of domain V observed in the structures of GTF180-ΔN, GTFA-ΔN and DSR-E ΔN$_{123}$-GBD-CD2.** The β2/β3 motifs usually contain either a 2-stranded β-hairpin (orange) or a 3-stranded β-sheet (green); extra β-strands are shown in cyan. Two or three motifs form a common β-solenoid structural module. (a) Superposition of structural modules. (b) Schematic organization of domain V; boxes correspond to motifs. (c) Structural/sequence alignment of domain V with the aligned sequence of GTF-SI domain V below. Open boxes correspond to the YG-repeats (black lines), A-repeats (blue lines) and C-repeats (red lines); YG motifs are shown in bold. Grey boxes correspond to the basic residues between motifs. Residues in grey are not visible in the structure (Leemhuis et al., 2013a).

Among the different crystal structures solved so far it appears that the domain V - which is the most distant from the active site - can adopt variable positions. For instance, domain V of GTFA-ΔN is shifted of about 20° compared to that of GTF180-ΔN (Pijning et al., 2012). ΔN123-GBD-CD2 domain V displays a completely different orientation, folded back to domain IV and giving a more compact global form to the enzyme (Brison et al., 2012) (Figure 17). Added to a higher average B-factor value compared to the other domains, these observations highlight a certain flexibility of this domain.

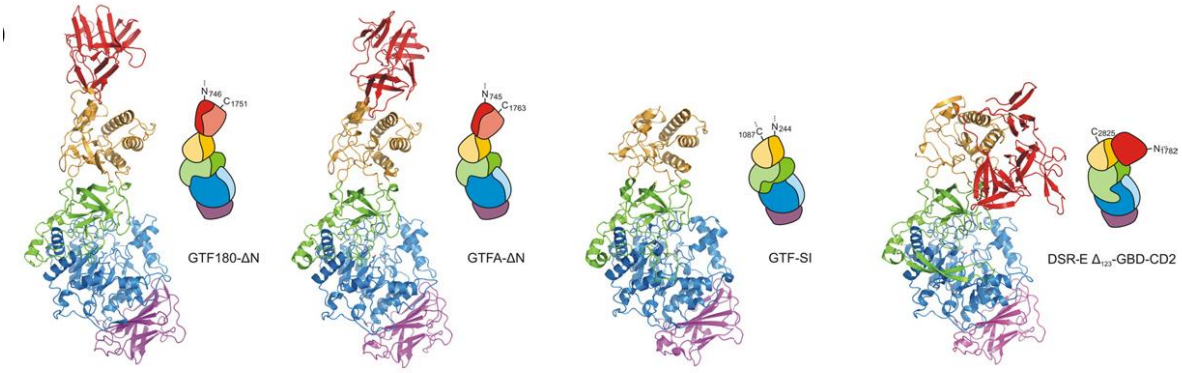


Figure 17: **Three-dimensional structures of glucansucrases.** Crystal structures of truncated glucansucrases in cartoon representation and schematic domain arrangement, colored by domain (A: blue, B: green, C: magenta, IV: yellow, V: red). From left to right: GTF180-ΔN (PDB: 3KLK), GTFA-ΔN (PDB: 4AMC, to be published), GTF-SI (PDB: 3AIE) and DSR-E ΔN_{123}-GBD-CD2 (PDB: 3TTQ). From (Leemhuis et al., 2013a).

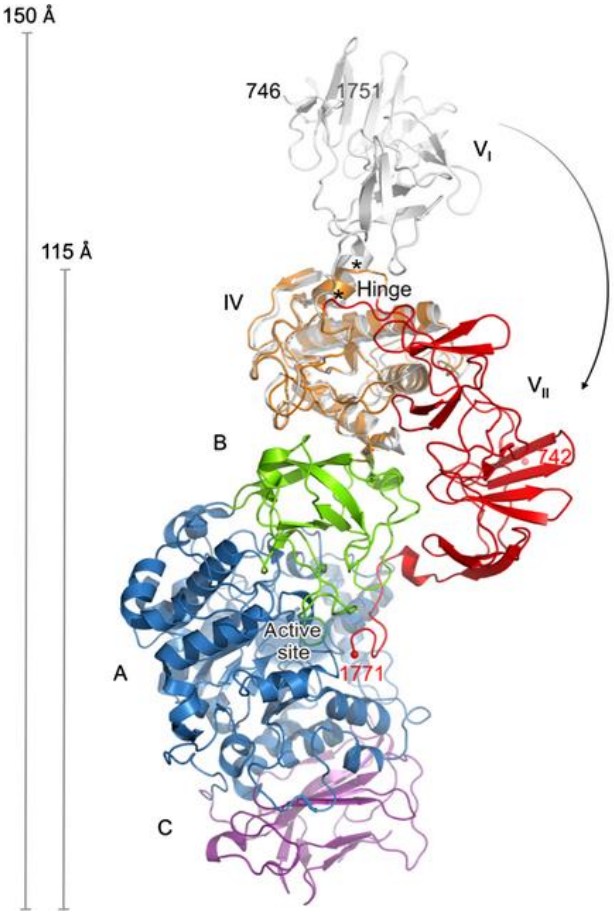


Figure 18: **Superposition of GTF180-ΔN I and GTF180-ΔN II crystal structures.** GTF180-ΔN I (grey, only domains IV and V are shown) and GTF180-ΔN II (molecule A; domains are colored the same as in Figure 17). Domain V makes an ~ 120° rotation about a hinge located between domains IV and V, near residues 794 and 1634 (indicated with an asterisk). Consequently, domain V is now near domains IV and B, whereas the C-terminal residues (visible up to residue 1771) bind at the protein surface near the active site cleft. The first and last visible residues of both crystal structures are indicated (746 and 1751 for GTF180-ΔN I; 742 and 1771 for GTF180-ΔN II). From (Pijning et al., 2014).

The ability of domain V to occupy different positions was further demonstrated by the elucidation of a new crystal form of GTF180-ΔN in which domain V is rotated of around 120° around a hinge between domains IV and V (Pijning et al., 2014) (Figure 18). SAXS analysis also revealed the positional variability of GTF180-ΔN domain V in solution between elongated and compact conformations (Pijning et al., 2014). Although no clear experimental evidence is available, the positional variability of domain V is certainly of functional relevance.

II.4.3.2. Functional role of domain V

Taken together, the remote position, the binding ability towards glucan and the flexibility of domain V suggest that it could serve as an anchoring zone for the growing carbohydrate chain bringing it towards and away the catalytic site. Several studies on domain V fragments or on full enzymes investigated this question first to characterize the binding abilities of this domain, second to understand its role in the catalytic mechanism.

For example, the glucan binding capacity of GtfI from *Streptococcus downei* was reduced by mutation of conserved aromatic residues of the previously described repeats (Shah et al., 2004). Transglucosylation, glucan binding properties and polymer size were affected by truncation of some repeats in other enzymes (Fabre et al., 2005; Joucla et al., 2006; Kralj, 2004b; Moulis et al., 2006b). Notably, stepwise truncations of the domain V of DSR-S *from L. mesenteroides* at both N- and C-terminal extremities induced progressive activity loss for the resulting variants as well as a progressive decrease of the produced polymer size (Moulis et al., 2006b) (Figure 19). In particular, the enzyme devoid of the whole domain V lost 99% of its catalytic activity and was only able to synthesize LMM dextran of around 10^4 g/mol (vs >10^8 for DSR-S vardel Δ4N). Deletion of the entire domain V of GTF180 heavily impaired polysaccharide synthesis (2% of the transferred available glucosyl units for GTF180-ΔNΔV vs 16% for GTF180-ΔN) in favor of oligosaccharide synthesis. However the polysaccharide molar mass produced by GFT180-ΔNΔV is equivalent to that of GTF180-ΔN (21x10^6 g/mol for GTF180-ΔNΔV and 22.6x10^6 g/mol for GTF180-ΔN). To note, the wild type protein produces a dextran of about 38 x10^6 g/mol (Kralj, 2004a). Thus, the effects of domain V deletions on enzyme catalysis vary from an enzyme to another rendering the elucidation of its role quite challenging.

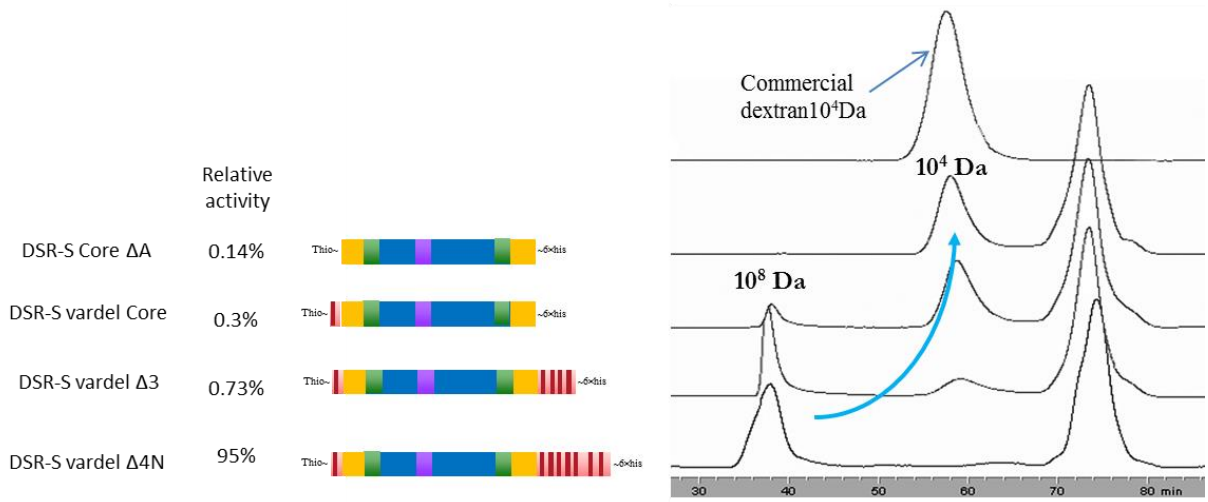


Figure 19: **Effect of stepwise deletion of the domain V of DSR-S from *L. mesenteroides* on produced dextran size.** Schematic representation of enzyme domains (pale red, domain V; yellow, domain IV; green, domain B; blue, domain A, purple, domain C; red bars, YG-repeats). Adapted from (Moulis et al., 2006b).

Other studies focused on GBDs separated from their catalytic cores reveal that the GBD fragments were able to fold autonomously and display high glucan affinity with K_d values in the nanomolar range (Kingston et al., 2002; Komatsu et al., 2007; Lis et al., 1995; Suwannarangsee et al., 2007)

Very recently, Brison *et al* solved the structure of the branching sucrase ΔN123-GBD-CD2 in complex with glucose, isomaltose and isomaltoriose molecules, providing the first structural evidences of interactions between the domain V and oligosaccharides. They further characterized glucan binding pockets sharing a common topology and composed of four YG repeats, 2 complete and 2 partial (Brison et al., 2016). They count approximately 80 residues and comprise three structural motifs with either β-hairpins or three-stranded antiparallel β-sheets. These pockets contain acidic or basic residues (D, E, K and R), aromatic residues (W or Y) and polar residues (T, S, Q and N). The glucosyl units of the sugar found in these pockets (namely V-K and V-L) are stabilized by a stacking interaction with an aromatic residue at the bottom of the pocket and a H-bond network with other residues of the pocket (Figure 20).

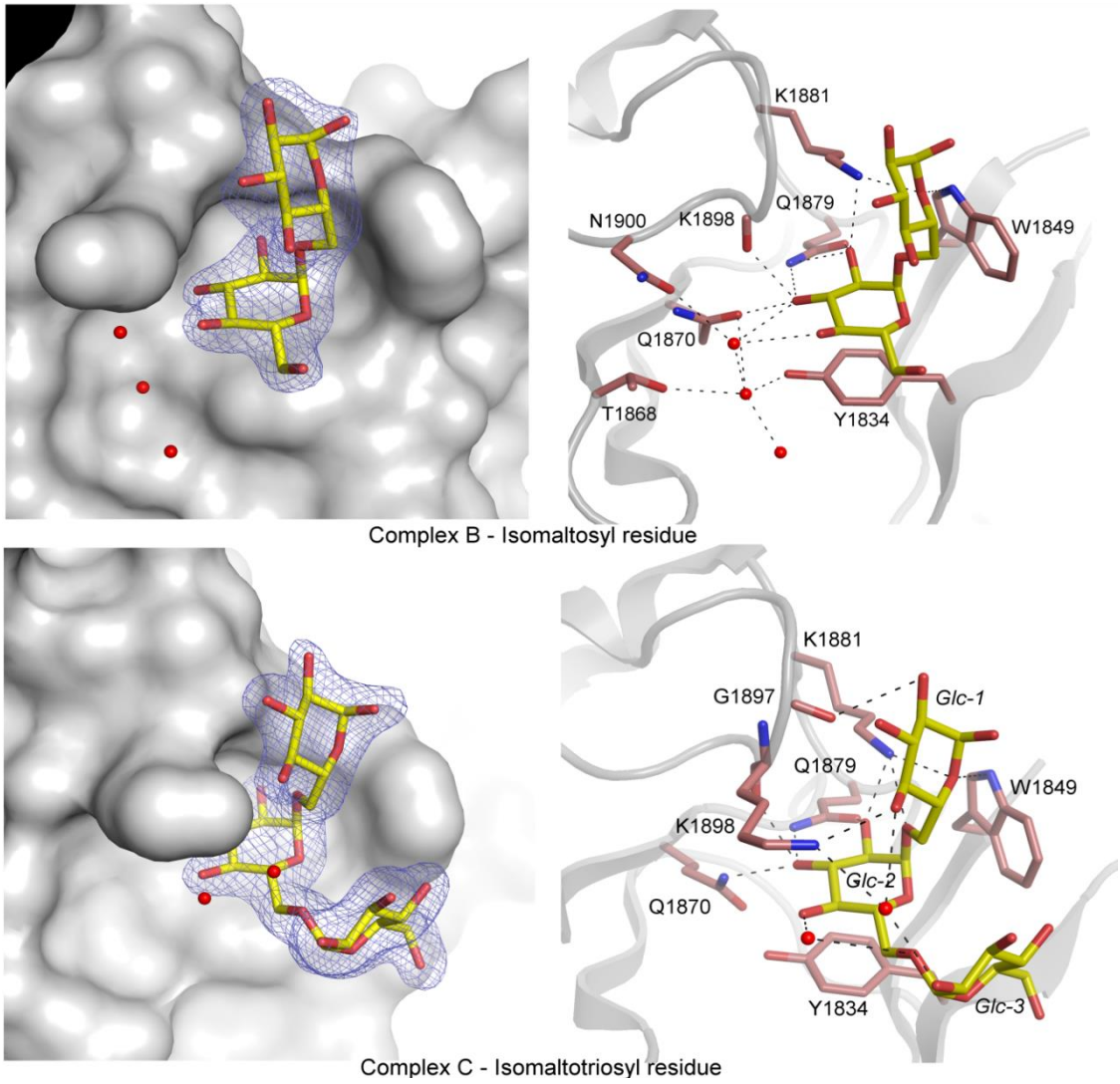


Figure 20: **Binding pocket V-K found in the GBD of ΔN_{123}-GBD-CD2 in complexes B (upper panels) and C (lower panels)** (Brison et al., 2016). For this pocket, in complexes B or C, an isomaltosyl or isomaltotriosyl residue was observed and shown as yellow sticks. The network of hydrogen bonding is shown on the right panels whereas the electron density map around carbohydrates is displayed on left panels. The residues and water molecules involved in binding are represented as pale red sticks and red spheres, respectively (right panels).

Further sequence analyses lead to the identification of similar pockets in the entire domain V of DSR-E (Figure 21) as well as in that of other GH70 enzymes proved to have dextran binding ability (Figure 22) (Brison et al., 2016). Also, in each structure of GTFA-ΔN and GTF180-ΔN, a pocket with a topology similar to the V-K and V-L binding pockets of ΔN_{123}-GBD-CD2 GBD could be identified. Residues of pocket V-K and V-L of ΔN123-GBD-CD2 which interact with the isomaltooligosaccharides are conserved in the sequence of the putative binding pockets found in the other enzymes indicating that these pockets are likely to be functional. In particular, one aromatic residue (Tyr or Phe resifdues) acting as a stacking platform, one tryptophan, one glutamine and one lysine are well conserved and may have a key role in glucan binding. However, experimental proofs remain to be found.

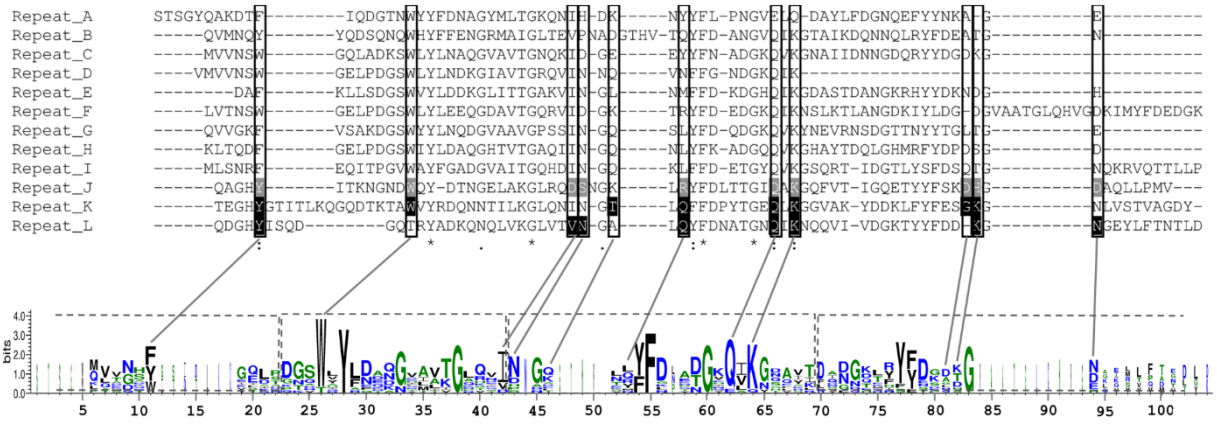


Figure 21: **Sequence alignment of the twelve repeats identified in the GBD of DSR-E** (Brison et al., 2016). Black highlighted residues are involved in sugar binding in repeats K and L while grey highlighted residues are proposed to play the same role in repeat J. Framed residues would participate in sugar binding for repeats A to I. A LOGO sequence based on this alignment is shown. The YG repeats are framed in dashed lines on the LOGO sequence.

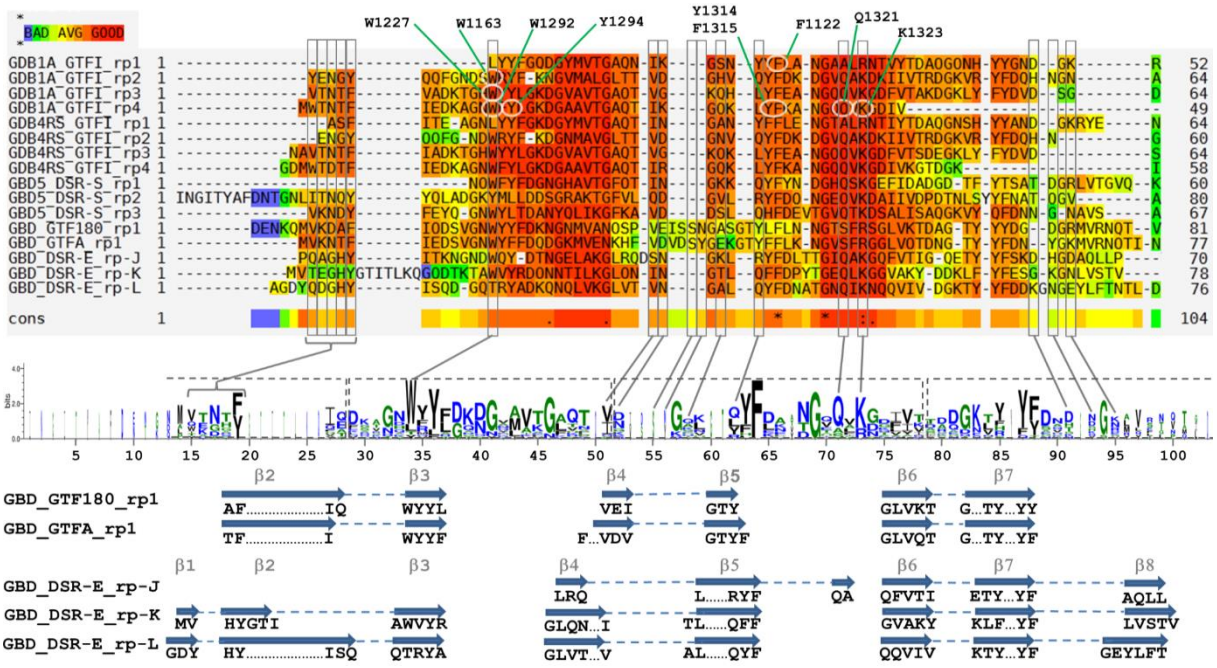


Figure 22: **Sequence alignment and LOGO sequence of the GBD repeats proved to bind α-glucans of GH70 glucansucrases** (Brison et al., 2016). GBD1A, GBD4RS and GBD5 bind dextrans with K_d values of 0.6 µM, 2.5 µM and 11.1 nM, respectively. Sequences are GBD1A_GTFI_rp# (repeats 1 to 4 of the GBD1A of GTFI from *S. downei*), GBD-4RS_GTFI_rp# (repeats 1 to 4 of the GBD-4RS of GFTI from *S. sobrinus*), GBD5_DSR-S_rp# (repeats 1 to 3 of the GBD5 of DSR-S from *L. mesenteroides* B-512F), GBD_GTF180_rp1 (repeat 1 of the GBD of GTF180-ΔN from *L. reuteri*), GBD_GTFA_rp1 (repeat 1 of the GBD of GTFA-ΔN from *L. reuteri* 121) and GBD_DSR-E_rp# (repeats J to L of the GBD of ΔN_{123}-GBD-CD2 from *L. citreum* NRRL-1299). Residues highlighted in red are well conserved while those tending toward blue are not. Residues framed in grey delineate the sugar-binding pockets and are pointed out in the LOGO sequence. Residues framed in dashed line on the LOGO sequence correspond to YG repeats. In the repeats of GBD1A, residues circled in white, were mutated by (Shah et al., 2004). Secondary structure elements, as determined by DSSP from PDB entries 3TTQ, 3KLK and 4AMC, are displayed below the LOGO sequence. β-strands, numbered according to Figure 8, linked by dashed lines represent either β-hairpins or three-stranded β-sheets.

Overall, those results let suggest that the domain V of GSs is involved in binding of their polymerization products as they constitute functional modules that can bind glucans by themselves. However, they do not seem to be directly involved in catalysis but likely facilitate glucan transfer to or away from the active site. In addition, the presence of multiple binding sites may enhance the affinity of the enzyme for carbohydrate polymers and increase the local concentration of carbohydrate chains in the vicinity of the active site and thereby influence the polymerization efficiency.

II.4.4. Polymerization mode: processive / non-processive and product specificity determinants

II.4.4.1. Mode of polymer elongation by glucansucrases

The mode of dextran synthesis has been a subject of debate for years. Robyt and co-workers (Robyt et al., 1974) who investigated the synthesis of dextrans by *L. mesenteroides* NRRL B-512F dextransucrase first proposed that polymer growth did not need any primer, involved two active sites and proceeded from the reducing end. One part of the scientific community agreed with this mechanism of elongation but some opponents, and among them the group of Greg Mooser (Mooser et al., 1991) were skeptical particularly because they did not believe in the existence of two catalytic sites. It is finally only since 2006 that the mode of dextran elongation involving only one active site was demonstrated for the DSR-S from *L. mesenteroides* NRRL B-512F, proceeding from glucosyl transfer onto the non-reducing end of a growing polymer chain (Moulis et al., 2006b). This mechanism is admitted today by the scientific community.

What was noticeable with DSR-S from *L. mesenteroides* NRRL B-512 is that the enzyme produces high molar mass polymer at the early stage of reaction, which is a typical feature of processive polymerases. However, a detailed analysis of the kinetic of polymer formation using sensitive analytical method (HPAEC-PAD: high-performance anion-exchange chromatography equipped with an ED40 pulsed amperometric detector) revealed that glucose was the first molecule to be produced and the glucose molecule was in turn glucosylated leading to a series of isomaltooligosaccharides after successive transfer onto the acceptor reaction products. The water molecule was the preferred acceptor but it was also shown that sucrose itself could also be used as a primer of dextran formation (Moulis et al., 2006b). Regarding other glucansucrases, the nature of the first primer often remains unclear. Structural analyses of the oligosaccharides produced at the beginning of the reaction by the reuteransucrase GTF-A revealed that sucrose (which is in high concentration at this time) is a good acceptor. Indeed, GTF-A starts to synthesize polysaccharides with the transfer of glucosyl units to the non-reducing glucose end of sucrose (Dobruchowska et al., 2013). LMM linear dextran produced by

S. sobrinus GTF-S3 also comprised a sucrose moiety at its extremity (Cheetham et al., 1991). Thus, depending on the enzyme, the nature of the preferred initial acceptor can vary.

The analysis of polymer formation with DSR-S further showed that there was no accumulation of the initially formed oligosaccharides in the medium as from a certain size corresponding to a degree of polymerization around 8, these compounds were rapidly elongated to yield HMM polymers (Moulis et al., 2006b). It was thus proposed that GSs adopt a multi-chain mechanism in the initial phase of the reaction and then switch to a single-chain elongation mode once the oligosaccharides initially formed have reached a sufficient size to be used as "primers" for the processive synthesis of polysaccharides. This mode of elongation was thus qualified as "semi-processive" (Moulis et al., 2006b). The switch from the distributive mode to the processive one was suggested to be mediated by the domain V of DSR-S as the enzyme totally or partially deleted of this domain adopts a distributive mode during all the elongation process (DSR-S core ΔA (Moulis et al., 2006b). Similar observations were then made with GTF180-ΔNΔV (Meng et al., 2015a)) further supporting the hypothesis that the anchoring zones of domain V would be required for the enzyme to act processively. However, these assumptions have not been supported by structural analysis to date, as only one branching sucrase in complex with isomaltooligosaccharides bound in domain V was solved. Such types of complex remain to be obtained with glucansucrases.

Finally, the termination of elongation has not been extensively studied. Is the polymer release governed by the polymer solubility, the strength of glucansucrase/polymer interaction or any other reasons? These are important questions, which have not been investigated in details yet.

II.4.4.2. Linkage specificity and polymer size

A large number of studies focusing on the linkage specificity are reported in the literature (André et al., 2010; Leemhuis et al., 2012, 2013a; Meng et al., 2016a). Before the availability of 3D-structures, mutagenesis studies were mostly directed towards the conserved motifs of the catalytic domain (motifs II, III and IV in particular, Table 3)(Côté and Skory, 2014; Hellmuth et al., 2008; Kralj et al., 2005b, 2006; van Leeuwen et al., 2008b; Moulis et al., 2006b; Van Leeuwen et al., 2009). From the obtained results, it appears that residues located downstream of the transition state stabilizer are involved in the linkage specificity of GH70 GSs. They participate in acceptor binding site shaping and orientate the way in which the acceptor substrates is accommodated and grafted on one specific position. From structural analysis, some residues of domain B were also shown to be implicated in the acceptor subsites shaping (Meng et al., 2014, 2015b). In particular, residues Leu938 and Leu940 of GTF180-ΔN were shown to be critical for product specificity. Mutations at these positions induced an increase of the α-(1→6) linkage proportion (Meng et al., 2014, 2015b). The L940W mutant even

lost its ability to synthesize α-(1→3) linkages resulting in the production of a totally linear dextran. To note, the molar mass of the products was also impacted by the mutation of this residues. These results also highlight the critical role of the loop (referred as B1, (Bai et al., 2017)) harboring the Leu940 in product specificities in terms of size or osidic linkage type (Meng et al., 2014). A residue located at the base of another loop in domain B (that will be referred as loop B2, see Chapter II) of the same enzyme, was shown to be involved in branch point formation (Meng et al., 2015b). Further studies focusing on structurally characterized subsites +1 and +2 also corroborate their implication in linkage specificity (Irague et al., 2013; Meng et al., 2015b, 2016c). Overall, these mutagenesis studies showed that the linkage specificity of GH70 GSs is determined by the interplay of different residues near the acceptor binding sites from both domains A and B. However, even if this knowledge is helpful for GH70 enzyme engineering, the synthesis of tailor-made α-glucans with given amounts of specific osidic linkage composition remains challenging.

Less is known on the determinants controlling the size of the produced polymers. As seen previously, it has been shown that domain V does have a role in the polymerization mechanism and hence on the molar mass of the products (see paragraph II.4.3.2. Functional role of domain V, p62). However, precise identification of each residues of this domain V directly involved is still needed. In addition, some residues close to the catalytic site were also shown to affect product size. Indeed, the mutations L940E and L940F of GTF180-ΔNΔV partially restored polysaccharide synthesis, which had been highly affected by the deletions of domain V (percentage of sucrose used for polysaccharide synthesis: 16.3, 2, 9 and 12.4 for GTF180-ΔN, GTF180-ΔNΔV, GTF180-ΔNΔV L940E and GTF180-ΔNΔV L940F, respectively) (Meng et al., 2015a). As mentioned before, the mutant GTF180-ΔN L940W produced a polymer with a molar mass of $6.3x10^6$ g/mol smaller than that of the wild type polymer. The hydrolytic activity was also reduced as well as the polysacharides/oligosaccharides ratio (Meng et al., 2014). Mutation of W1065 (subsite +2) of GTF180-ΔN into non-aromatic residues impaired the enzyme activity and abolished polysaccharide synthesis (Meng et al., 2017). Similarly, no detectable activity was observed for the mutants W491A or G (equivalent to W1065 of GTF180) of GTFI from *Streptococcus mutans* demonstrating its essential role in the catalytic activity. The GTF-R S628D (upstream of the transition state stabilizer) kept its catalytic efficiency but lost its ability to polymerize glucosyl units and only produce isomaltose, palatinose and leucrose, while the wild type enzyme displayed a dextran production yield higher than 80% (Hellmuth et al., 2008).

To conclude, the determinants involved in the polymer size and linkage specificity are diverse and located in α-glucan binding areas close to, or remote from the active site. However, if it is rather accurate to determine the linkage specificity from sequence and structure analysis, it remains much more difficult to predict the size of the polymer that will be synthesized by a glucansucrase,

indicating that additional studies are necessary to identify with precision the determinants participating in the control of the molar mass.

III. DSR-M and DSR-OK, the two templates of this study

Recently, data mining of different LAB strains allowed the identification of various new GH70 enzymes. Among them, DSR-OK from *Oenococcus kitaharae* DSM17330 present the particularity of synthesizing a dextran polymer with the highest molar mass (10^9 g/mol) reported to date whereas DSR-M from *Leuconostoc citreum* NRRL B-1299 produces a dextran of low molar mass (28 kg/mol) without any HMM dextran contamination, directly from sucrose as sole substrate. On another hand, those two enzymes both produce highly linear dextrans, and were thus considered as good models for investigating the structural features involved in the control of the polymer molar mass.

III.1. DSR-M

Leuconostoc citreum NRRL B-1299 was first described in 1954 by Jeanes et al. for its ability to produce dextrans comprising up to 35% of α-(1→2) glucosidic branches (Jeanes et al., 1954; Kobayashi and Matsuda, 1977). For this reason, enzymatic extract of this strain was industrially used for the production of prebiotic compounds (Paul et al., 1992) (see paragraph, I.2.2.2. Dextrans and isomaltooligosaccharides as prebiotics, p37). In the past, three GS-encoding genes (*dsrA*, *dsrB* and *dsrE*) were identified in this strain, the corresponding recombinant proteins were then biochemically characterized. From sucrose, DSR-A and DSR-B catalyze the synthesis of dextrans mainly composed of α-(1→6) linkages (85% for DSR-A, with 15 % α-(1→3) linkages; 100% for DSR-B (Monchois et al., 1996, 1998b)). As previously described, DSR-E glucansucrase is more atypical as it is the sole GH70 enzyme comprising two catalytic cores (Bozonnet et al., 2002; Fabre et al., 2005) (cf, II.3.1. Brief overview of the branching-sucrases, p46). The polymer produced by this enzyme was shown to contain only 5% α-(1→2) branching linkages, indicating that this enzyme cannot be solely responsible for the synthesis of the highly α-(1→2) branched dextran produced by the strain. In order to identify the full inventory of GH70-encoding genes in this strain, and particularly find the native enzyme responsible of the highly branched dextran synthesis, the genome of *L. citreum* NRRL B-1299 was sequenced. It resulted in the identification of three additional GH70 encoding genes, namely *brsA*, *dsrM* and *dsrDP* (Passerini et al., 2015).

Compared to other *L. citreum* strains, the *dsrDP* gene was identified in a sequence area specific to *L. citreum* NRRL B-1299 and which probably results from a phage-mediated sequence horizontal transfer. The corresponding recombinant protein was shown to synthesize a highly polydisperse polymer comprising five populations with a weight average molar mass estimated at i) 580 g/mol, ii)

2000 g/mol, iii) 4240 g/mol, iv) 9300 g/mol and v) >2×10^6 g/mol and mainly composed of α-(1→6) glucosidic linkages (Passerini et al., 2015).

Recombinant BRS-A is a branching sucrase able to graft α-(1→6) linear dextrans with α-(1→2)-linked glucosyl residues (cf. II.3.1. Brief overview of the branching-sucrases, p46).

The recombinant DSR-M was shown to be the first natural GH70 enzyme producing a dextran polymer composed exclusively of α-(1→6) linkages with a very low weight average molar mass (estimated at 28 kg/mol). In addition, DSR-M shows a broad pH and temperature tolerance with optima at pH 5.75 and 33 °C. Sequence analysis revealed that DSR-M is very similar to the GSE16-5 glucansucrase (94% identity) that was recently identified in *L. citreum* LBAE- E16 (Amari et al., 2015). The conserved motifs I to IV of DSR-M are very similar to the motifs found in dextransucrases specific for α-(1→6) linkage synthesis. Indeed, the 791-SEVQT (DSR-M numbering) of motif IV (Table 4) is shared by a majority of dextransucrase enzymes. Residues of subsites -1 and +1 -interacting respectively with the glucosyl and the fructosyl ring of sucrose in the inactive mutant of GTF-180ΔN in complex with sucrose (PDB : 3HZ3)- are conserved in DSR-M (Passerini et al., 2015; Vujicic-Zagar et al., 2010a).

The C-terminal domain V of DSR-M is much longer than domain V of GTF-180-ΔN. Eleven YG repeats (Giffard and Jacques, 1994), reported to play a key role in glucan binding (see II.4.3.1. Structural organization of domain V, p58), can be identified in DSR-M domain V. Interestingly, others repeats with yet unknown functions are found in the C-terminal DSR-M domain V (APY repeats), exhibiting more than 98% identity with the C-terminal region of *L. citreum* NRRL B-1355 alternansucrase and *L. citreum* CW28 inulosucrase (Argüello Morales et al., 2001; Joucla et al., 2006; Olivares-Illana et al., 2003). Deletion of these APY motifs in those enzyme did not alter their specificity (nature and size of polymers produced) Thus, DSR-M exhibits a dextransucrase-like catalytic core and an alternansucrase-like C-terminal region as already observed for GSE16-5 glucansucrase (Amari et al., 2015).

The main interesting feature of this novel GH70 enzyme is its ability to synthesize directly from sucrose low molar mass and very linear dextrans. As we have seen before, dextran applications largely depend on their size hence, the possibility to control this parameter in dextran production process is biotechnologically challenging (cf. I.3. Dextran production processes, p41). DSR-M is a candidate of choice for such purposes. Indeed, by playing on the sucrose initial concentration and the reaction temperature, a tight control of the polymer molar mass is allowed by the use of this new catalyst. These properties have been recently patented WO2016016544 (A1) (Vuillemin et al., 2015a).

Table 4: **Sequence alignment of functional conserved motifs I-IV of GH70 enzymes.** The numbering order of the motifs (II, III, IV and I) refers to the motifs originally defined in the family GH13. The two catalytic residues are indicated in red for the nucleophile residue (motif II) and the acid/base catalyst (motif III), and the transition-state stabilizer residue is shown in blue. The residues in green and orange correspond to residues of subsites -1 and +1, respectively. Grey shading represents enzymes encoded by the genome of *L. citreum* NRRL B-1299 (Passerini et al., 2015).

Protein		Motif II		Motif III		Motif IV		Motif I*	Specificity
GTF-I [Sd]	449	SIRVDAVDNVD	486	HVSIVEAWSDN	559	FARAHDSEVQDLIRD	931	ADWVPDQ	α-(1→3)
GTF-I [Ss]	443	SIRVDAVDNVD	480	HVSIVEAWSDN	553	FARAHDSEVQDIIRD	925	ADWVPDQ	
GTF-B [Sm]	1011	SIRVDAVDNVD	1048	HLSILEAWSDN	1120	FIRAHDSEVQDLIAD	1488	ADWVPDQ	
GTF-A [Lr]	1020	SVRVDAPDNID	1056	HINILEDWNHA	1128	FVRAHDNNSQDQIQN	1508	ADWVPDQ	α-(1→4) / α-(1→6)
GTF-O [Lr]	1020	SVRVDAPDNID	1056	HINILEDWNSS	1128	FIRAHDNNSQDQIQN	1508	ADWVPDQ	
ASR [Lm]	631	GIRVDAVDNVD	668	HLSILEDWNGK	762	FVRAHDYDAQDPIRK	1168	ADWVPDQ	α-(1→6) / α-(1→3)
GTF-W [Lr]	748	GFRVDAADNID	785	HLVYNEGYHSG	568	FVTNHDQR-KNVINQ	1216	EDLVMNQ	α-(4→6)
GTF-ML4 [Lr]	1012	GFRVDAADNID	1049	HLSYNEGYHSG	1121	FVTNHDQR-KNLINR	1479	EDIVMNQ	
DSR-CB4 [Lc]	526	GIRVDAVDNVD	563	HLSILEDWSHN	636	FVRAHDSEVQTVIAQ	1001	ADWVPDQ	α-(1→6)
DSR-C [Lm]	498	GIRVDAVDNVD	535	HLSILEDWSHN	608	FVRAHDSEVQTVIAQ	973	ADWVPDQ	
DSR-S [Lm]	547	GIRVDAVDNVD	584	HLSILEDWSHN	657	FVRAHDSEVQTVIAQ	1023	ADWVPDQ	
GTF-180 [Lr]	1021	GIRVDAVDNVD	1058	HINILEDWGWD	1131	FVRAHDSNAQDQIRQ	1503	ADWVPDQ	
DSR-A	361	GYRVDAVDNVD	399	HLSILEDWGDE	472	FIRAHDSEVQTIIAD	844	NDWVPDQ	α-(1→6)
DSR-B	401	GIRVDAVDNVD	449	HLSILEDWSHN	512	FVRAHDSEVQTVIAQ	877	ADWVPDQ	α-(1→6)
DSR-E (CD1)	520	GYRVDAVDNVD	557	HISILEDWDNN	630	FIRAHDSEVQTVIAQ	1010	NDWVPDQ	α-(1→6)
DSR-E (CD2)	2206	SIRIDAVDFIH	2243	HISLVEAGLDA	2317	IIHAHDKGVQEKVGA	2688	ADVVDNQ	α-(1→2)
DSR-DP	455	GIRVDAVDNVD	492	KISILEDWAWG	565	FIRAHDAESQDIFSN	940	ADYVPDQ	α-(1→6)
BRS-A	668	SIRIDAVDFVS	705	HLSLVEAGLDA	779	IIHAHDKDIQDKVGA	1151	ADVVANQ	α-(1→2)
DSR-M	673	SIRIDAVDNVD	710	HIHILEDWSPN	785	FIRAHDSEVQTIIAK	1177	ADFVPDQ	α-(1→6)

The variation of the molar mass with either the initial sucrose concentration or the reaction temperature was shown to be quasi-linear (Figure 23 and Figure 24). Thus, by controlling reaction conditions, DSR-M can synthesize dextran from 7 to 25 kg/mol in good yields (more than 81% of the available glucosyl units are recovered in dextran synthesis for each reaction) (Vuillemin et al., 2015a). These results are very promising from a biotechnological point of view as the use of DSR-M presents several advantages. Indeed: i) dextran synthesis is performed directly from sucrose without addition of acceptor molecules or dextranase activity, ii) when synthesized from sucrose concentrations below 400 g/L, DSR-M dextran polydispersity is low without detectable presence of high molar mass polymer, iii) the polymer size is easily controllable by adjusting sucrose concentration, iv) the production yields are high (>80%) and v) DSR-M activity is rather stable on a large range of pH and temperature (80% of the maximal activity from 28 to 45 °C and pH from 4.2 to 6) and vi) this enzyme is very efficient since it is activated by its own product, reducing the synthesis duration (cf. also chapter II).

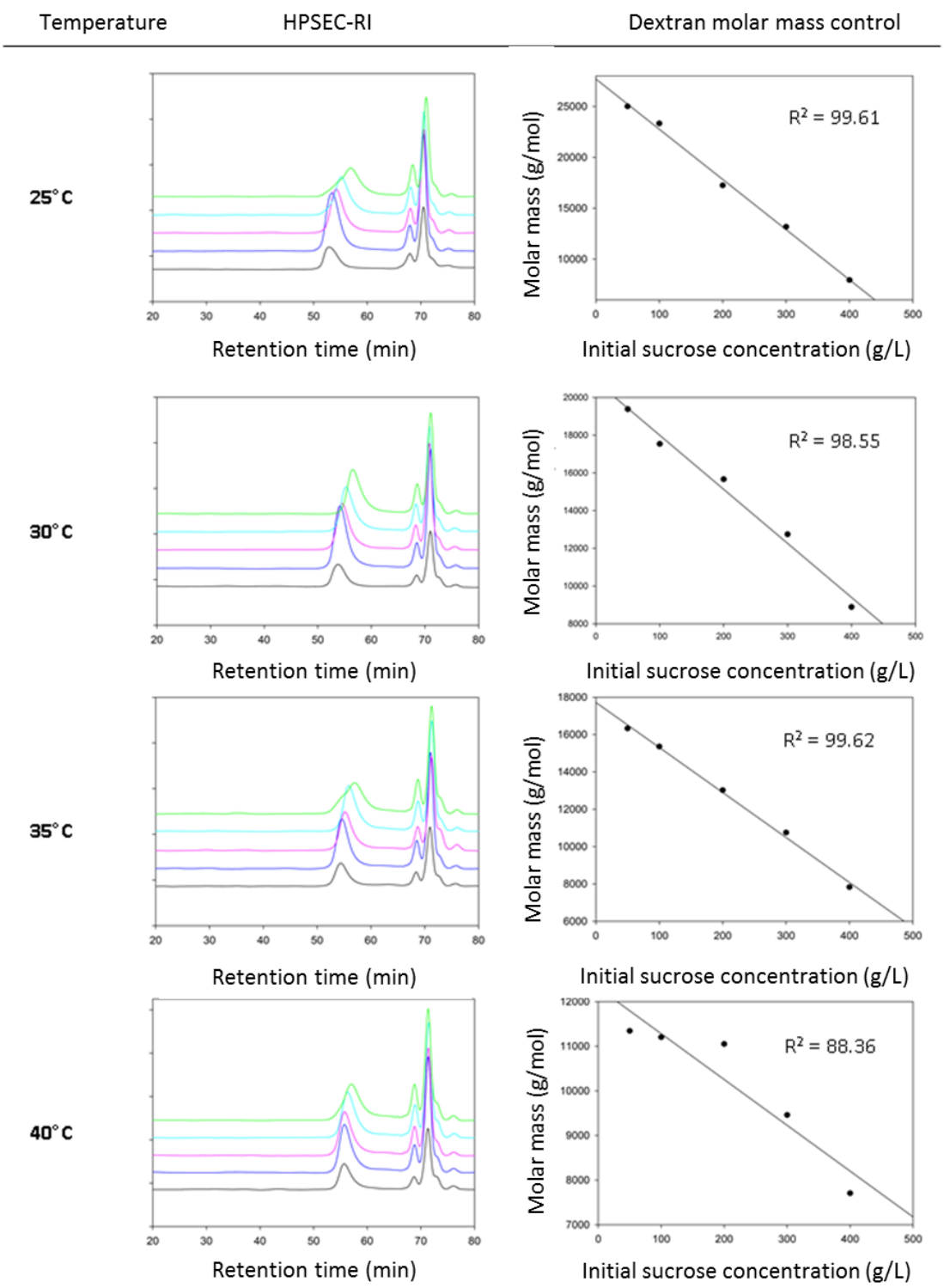


Figure 23: **Molar mass control of dextran produced by DSR-M as a function of sucrose initial concentration at a certain temperature.** From (Vuillemin et al., 2015a).

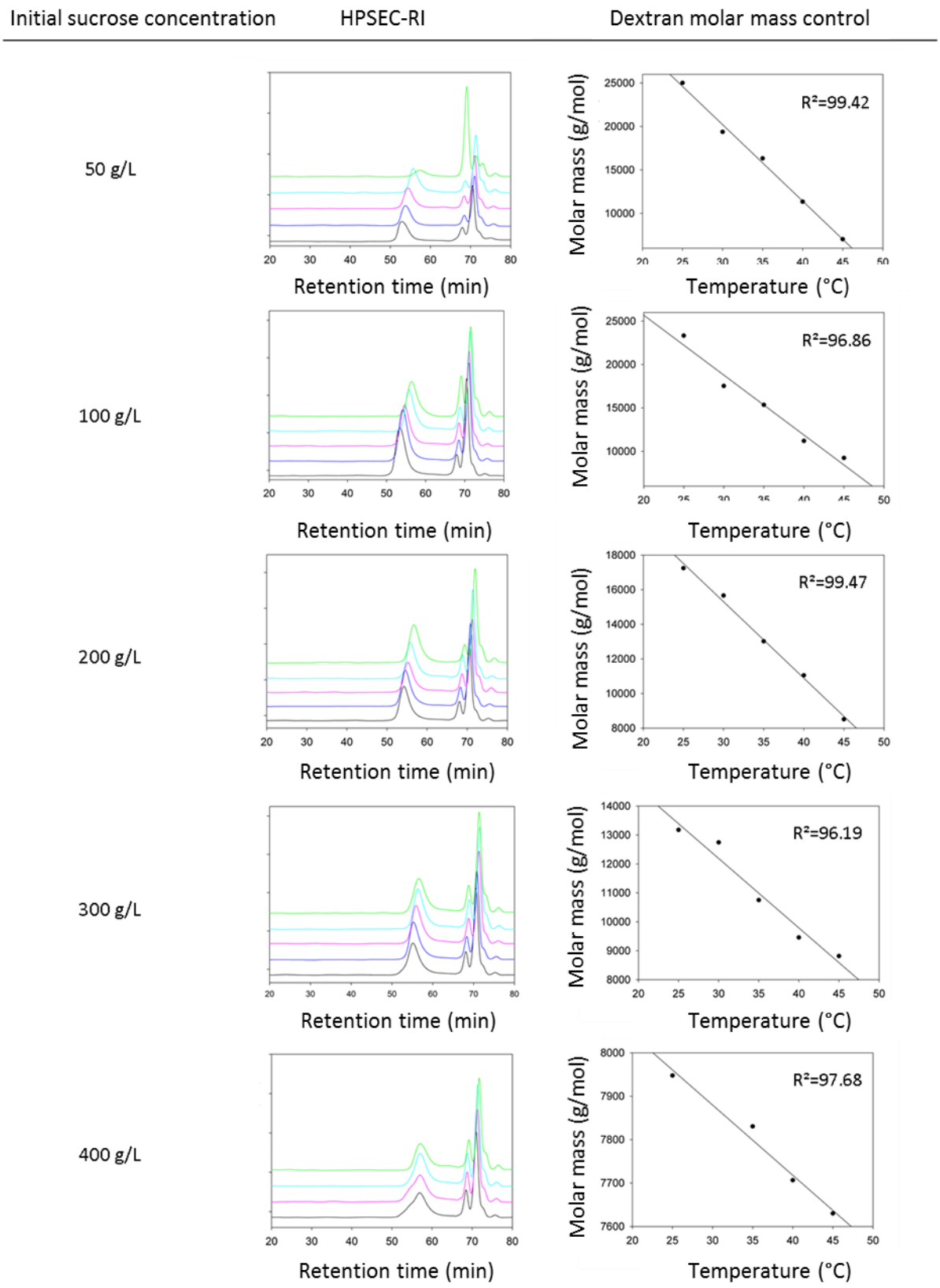


Figure 24: **Molar mass control of dextran produced by DSR-M as a function of temperature and sucrose initial concentration.** From (Vuillemin et al., 2015a).

III.2. DSR-OK

The lactic acid bacteria of *Oenococcus kitaharae* DSM17330 was isolated in Japan from composted distilled Shochu residue. The genome of this strain was sequenced in 2012 by Borneman et al. (Borneman et al., 2012), and further genome mining allowed the identification of putative dextransucrase gene named *dsr-ok*. The resulting recombinant protein is the first characterized GH70 enzyme in the genus *Oenococcus*. This enzyme produces a very HMM dextran with physical properties never observed before (Vuillemin et al., 2017).

Sequence analyses revealed that DSR-OK shares the highest amino acid identity of 58% with the known dextransucrase DSR-D from *L. mesenteroides* LCC4 (58% id. and 71% similarity, on 96% coverage) and with other putative dextransucrase sequences sharing 99% identity on 100% coverage with DSR-D sequence, such as the DSR-S from the industrial strain *L. mesenteroides* NRRL B-512F. As for DSR-M, the conserved motifs I to IV of the catalytic domain are similar to those found in dextransucrases specific for α-(1→6) linkages (Table 5). In particular, DSR-OK motifs are very similar to those of DSR-S from B-512F, only one amino acid diverging between the two enzymes, Ile 461 of DSR-OK being replaced by Val 550 of DSR-S. Essentially, the sequence 574-SEVQT of motif IV is characteristic of the α-(1→6) linkage specificity (Kralj et al., 2005b).

Table 5: **Protein sequence alignment of DSR-OK conserved motifs (I-IV)**. The two catalytic residues are indicated in bold red (NU) for nucleophile residue and (A/B) for the acid/base catalyst, and the transition state stabilizer (TSS) is represented in bold blue. The residues in bold green and in bold orange correspond to residues of subsite -1 and +1, respectively (Vuillemin et al., 2017).

		Motif II		Motif III		Motif IV		Motif I	Specificity
GTF-I [Sd] [1, 2]	449	SIRVDAVDNVD	486	HVSIVEAWSDN	559	FARAHDSEVQDLIRD	931	ADWVPDQ	
GTF-B [Sm]	1011	SIRVDAVDNVD	1048	HLSILEAWSDN	1120	FIRAHDSEVQDLIAD	1488	ADWVPDQ	α-(1→3)
GTF-SI [Sm]	473	SIRVDAVDNVD	510	HLSILEAWSDN	583	FIRAHDSEVQDLIRD	954	ADWVPDQ	
GTF-A [Lr]	1020	SVRVDAPDNID	1056	HINILEDWNHA	1128	FVRAHDNNSQDQIQN	1508	ADWVPDQ	α-(1→4) /
GTF-O [Lr]	1020	SVRVDAPDNID	1056	HINILEDWNSS	1128	FIRAHDNNSQDQIQN	1508	ADWVPDQ	α-(1→6)
ASR [Lm]	631	GIRVDAVDNVD	668	HLSILEDWNGK	762	FVRAHDYDAQDPIRK	1168	ADWVPDQ	α-(1→6) / α-(1→3)
DSR-CB4 [Lc]	526	GIRVDAVDNVD	563	HLSILEDWSHN	636	FVRAHDSEVQTVIAQ	1001	ADWVPDQ	
DSR-C [Lm]	498	GIRVDAVDNVD	535	HLSILEDWSHN	608	FVRAHDSEVQTVIAQ	973	ADWVPDQ	
GTF-180 [Lr]	1021	GIRVDAVDNVD	1058	HINILEDWGWD	1131	FVRAHDSNAQDQIRQ	1503	ADWVPDQ	α-(1→6)
DSR-S [Lm]	547	GIRVDAVDNVD	584	HLSILEDWSHN	657	FVRAHDSEVQTVIAQ	1023	ADWVPDQ	
DSR-D [Lm]	547	GIRVDAVDNVD	584	HLSILEDWSHN	657	FVRAHDSEVQTVIAQ	1023	ADWVPDQ	
DSR-OK [Ok]	458	GIRIDAVDNVD	495	HLSILEDWSHN	568	FVRAHDSEVQTVIAQ	936	ADWVPDQ	
		NU		**A/B**		**TSS**			

The optimal pH and temperature of DSR-OK activity were determined at 5.0 and 30 °C, respectively. With a half-life time of 111 h at 30°C for the non-purified fraction, and 20 h for the purified form, respectively, this enzyme is quite stable. As a comparison, DSR-S vardel Δ4N – a truncated variant of the dextransucrase DSR-S from *L. mesenteroides* NRRL B-512F that is a highly efficient polymerase for the synthesis of high molar mass dextran – displays a half-life time of only 24 hours (crude extract) and 50 min (purified form) at the same temperature (Moulis et al., 2006a). The kinetic parameters of

DSR-OK were also determined and fit with an uncompetitive substrate inhibition model with a Km of 11.7 mM ± 1.2 mM of sucrose, a Vmax of 4055 ± 52 µmoles of fructose released.s^{-1}.g^{-1}, corresponding to a kcat of 691 s^{-1} with a sucrose inhibition constant Ks estimated at 1.78 ± 0.13 M. DSR-OK efficiency is closed to that of DSR-S vardel ΔN which displays an equivalent kcat of 584 s^{-1}, but is more inhibited with a Ks of 326 mM sucrose (Moulis et al., 2006a).

Further biochemical characterization of this new catalyst revealed its ability to produce a very HMM dextran with very good yield (89% of the sucrose transferred glucosyl units). The synthesized dextran is composed of 97.6% of α-(1→6) linkages and 2.4% of α-(1→3) bonds and its size was determined by AF4-MALLS to be superior to 10^9 g/mol (Vuillemin et al., 2017). This is the longest dextran described to date and its rheological properties are very singular. While commercial dextrans generally adopt a Newtonian behavior, viscosity of 5% aqueous solution of DSR-OK dextran decreased with increasing shear rates, which is characteristic of a non-Newtonian polymer solution (this was also the case for the dextran produced by DSR-S vardel ΔN or some other dextrans previously reported). The viscosity of DSR-OK dextran solution is very important and much higher than that of other dextrans (Figure 25). DSR-OK product is also the first dextran for which a yield stress is described which makes it very interesting for biotechnological applications (Figure 25). Indeed, polysaccharides presenting yield stress (like xanthan or chemically modified gums) are often used for particles or foam suspensions in food, cosmetic or in oil and gas industries. Thus, DSR-OK dextran shows an atypical rheological behavior and a high viscosity compared to other dextrans. This polymer is hence of great interest for industrial applications in which it could be used as texturing agent for instance. For these reasons the enzyme and its product were recently patented WO2015193492 (A1) (Vuillemin et al., 2015b).

In the goal to find alternatives to guar gum uses which is submitted to important price fluctuations, industrial partners get interested in the potential of DSR-OK dextran. Hence, three patents related to the use of DSR-OK in particular conditions and the use of its product modified or not have been recently filed by Rhodia Operations (Solvay): FR3045608, FR3045668, FR3045667 (CASTAING et al., 2017; LABEAU et al., 2017a, 2017b). Noteworthy, DSR-OK was not significantly inhibited by salts or surfactants in contrast with the majority of its counterparts. The first patent reports on dextran synthesis in the presence of salts, and describes a process which could be directly used to enhance oil recovery (LABEAU et al., 2017b). The other one is related to dextran synthesis in the presence of surfactant, which would allow the *in situ* synthesis of polymers increasing the viscosity of cosmetic or detergent preparations (CASTAING et al., 2017). The last one concerns the carboxylation of DSR-OK dextran to further enhance its rheological properties (LABEAU et al., 2017a).

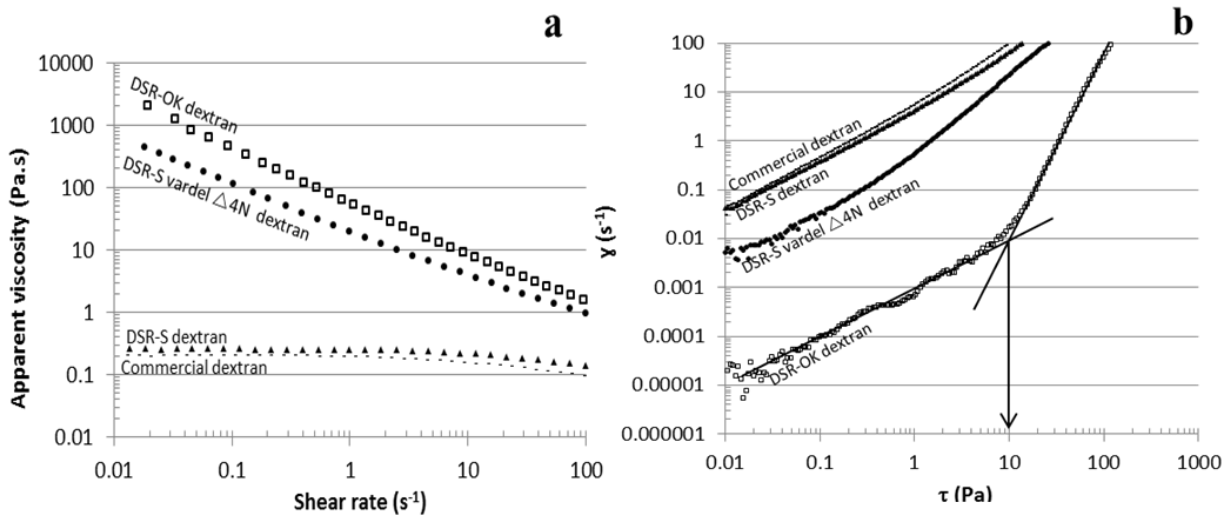


Figure 25: **Rheological analyses of 5% (w/v) aqueous dextran solutions**. DSR-OK dextran (□), DSR-S vardel Δ4N dextran (•) native DSR-S dextran (▲) and commercial dextran (-) a) Apparent viscosity as a function of shear rate for; b) Yield stress measurements.

DSR-M and DSR-OK -synthesizing the smallest and the highest dextran reported to date, respectively- are thus two novel GH70 enzymes with important biotechnological potential but also represent candidates of choice to deepen our knowledge on the structural determinants responsible of GH70 product size specificity.

Thesis objectives

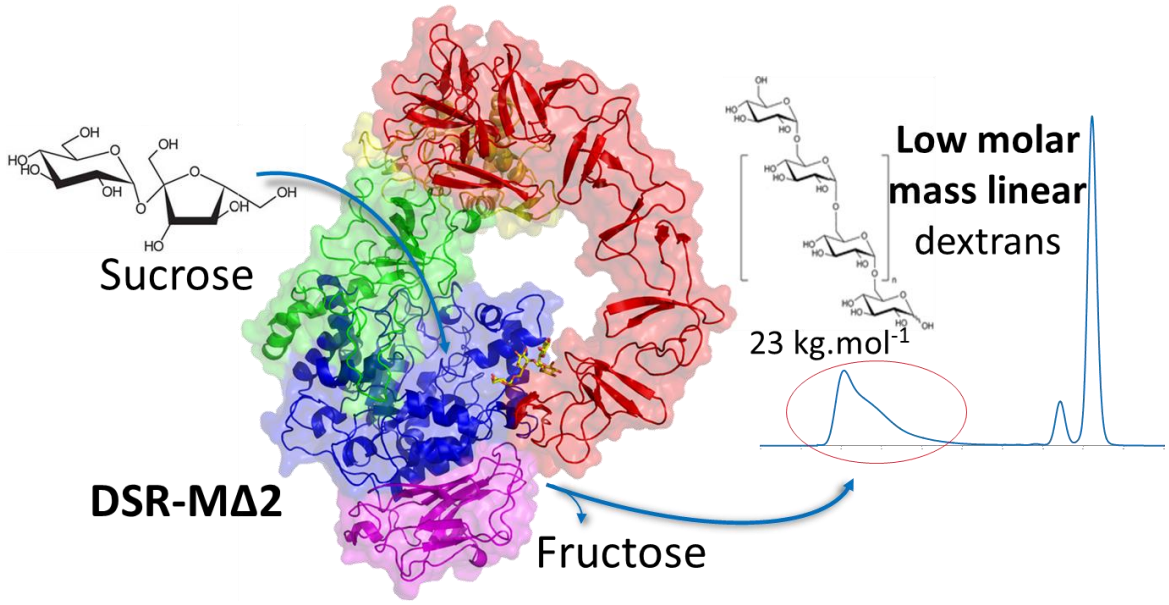
As previously shown, DSR-M from *Leuconotoc citreum* NRRL B1299 and DSR-OK from *Oenococcus kitaharae* DSM17330 - two novel GH70 enzymes recently characterized in the team – produce very different dextrans. Indeed, DSR-OK synthesizes the biggest dextran described to date while DSR-M produces only low molar mass very linear dextrans. These two templates were selected as the model enzymes for our study.

Both enzymes share a very similar architecture with a predicted organization in five domains, the catalytic triad and the conserved motifs I, II, III and IV. Notably, the "SEVQT" sequence in the motif IV, which is reported as a signature of α-(1→6) specificity is present in both sequences. Also, these enzymes share a large domain V with common YG repeated units. It was already shown that these repeats are involved in interactions with the growing glucan chains. Indeed, isolated domain V fragments were proved to bind dextrans with high affinity constants and truncation of these repeats in an efficient polymerase impacted the size of the produced dextrans. In addition and in parallel with the present project, two 3D-structures of a GH70 branching sucrase in complex with isomaltosyl residues enabled the identification of sugar binding pockets (composed of YG-repeats) in the domain V of the enzyme. From these findings, we speculated that the determinants controlling the size of the polymers produced by our enzyme candidates could be located in their domain V. To verify our hypothesis and acquire a deeper insight on the mode of polymer elongation as catalyzed by GH70 glucansucrases, a multi-disciplinary approach, including structural and biochemical characterization, as well as mutagenesis studies was undertaken with our two candidates to identify differences and bring answers to the following questions:

- Why do certain GH70 polymerases synthesize high molar mass polymers while others produce low molar mass polymers?
- Is the elongation mechanism the same?
- What determinants are responsible of their specificity?
- What is the influence of domain V with regard to the role of the catalytic domain?

Chapter II

Investigations on the determinants responsible for low molar mass dextran formation by DSR-M dextransucrase



Marion Claverie, Gianluca Cioci, Marlène Vuillemin, Nelly Monties, Pierre Roblin, Guy Lippens, Magali Remaud-Simeon, Claire Moulis

ACS Catalysis (2017)

Abstract

Certain enzymes of the GH70 family dextransucrases synthesize very high molar mass dextran polymers, whereas others produce a mixed population of very high and low molar mass products directly from sucrose substrate. Identifying the determinants dictating polymer elongation would allow the tight control of dextran size. To explore this central question, we focus on the recently discovered DSR-M enzyme from *Leuconostoc citreum* NRRL B-1299, which is the sole one that naturally, exclusively and very efficiently produces only low molar mass dextrans from sucrose. Extensive biochemical and structural characterization of a truncated form of DSR-M (DSR-MΔ2, displaying the same biochemical behavior as the parental enzyme), X-ray structural analysis of complexes with sucrose and isomaltotetraose molecules, together with accurate monitoring of the resulting polymer formation reveal that DSR-MΔ2 adopts a non-processive mechanism attributed to i) a high propensity to recognize sucrose as a preferred acceptor at the initial stage of catalysis, ii) an ability to elongate oligodextrans irrespective of their size and iii) the presence of a domain V showing a weak ability to bind to the growing dextran chains. In this study, we present the 3D-structure with the largest defined domain V reported to date in the GH70 family, and map sugar binding pockets on the basis of the structure of the complex obtained with isomaltotetraose. Altogether, these findings give insights on the interplay between the domain V and the catalytic site during polymerization. They open promising strategies for GH70 enzyme engineering aiming at modulating glucan size.

Key-words

Glucansucrases, dextransucrases, dextran, crystal structure, GH70 family, non-processive polymerization, glucan-binding domain, carbohydrate-binding protein.

I. Introduction

Glucansucrases (GS) from the glycoside hydrolase family 70 (GH70) are α-transglucosylases found in lactic acid bacteria (LAB) such as *Leuconostoc*, *Lactobacillus*, *Streptococcus*, and *Weissella* (Cantarel et al., 2009; André et al., 2010; Meng et al., 2016a). From sucrose, an economical and abundant agro-resource, they catalyze the polymerization of the glucosyl units to form homopolymers of variable size and structure, notably with different types of α-osidic bonds and degree of branching. Among the resulting polymers, dextrans contain more than 50% of α-(1→6) linkages and usually display a very high molar mass (HMM) (10^6 to 10^8 g/mol). The linear dextrans (more than 95% of α-(1→6) osidic bonds) were the first microbial polysaccharides to find industrial applications as infusion fluids, volume expanders, iron-carriers, anti-coagulants or vaccine adjuvants after derivatization (Naessens et al., 2005; Vettori et al., 2012). These applications mainly concern dextran fractions of average molar mass of 1, 10, 40 or 70 x10^3 g/mol, obtained by partial acid hydrolysis and solvent-fractionation of the native polymer produced by fermentation (Basedow and Ebert, 1977). Dextran70 (70x10^3 g/mol) is referenced in the List of Essential Medicines by the World Health Organization (19th WHO model List of Essential Medicines, 2015). The chemical processes behind the production of these clinical fractions require several steps, and are, consequently, quite costly, time-consuming and non-eco-friendly (Wolff et al., 1954; Vettori et al., 2012), stimulating the search for alternative and greener processes. Dextransucrase and dextranase activities have for example been combined to develop a one-step synthetic process of low molar mass dextrans (Gan et al., 2014; Goulas et al., 2004b; Khalikova et al., 2005; Ölçer and Tanriseven, 2010; Zhang et al., 2013). Reactions in the presence of sucrose and sugar acceptors such as glucose or maltose were also investigated to favor the synthesis of low molar mass dextrans, but these conditions never totally suppressed the formation of high molar mass polymers (Dols et al., 1997; Kothari et al., 2012; Paul et al., 1986).

Control of the size and structure of the polymers synthesized by glucansucrases would benefit from a better comprehension of the polymerization mechanism and, in particular, of the protein structural determinants involved in polymer size control or linkage specificity. Glucansucrases as flexible proteins of high molecular weight (~150-200 kDa) are difficult to crystallize, and for years, structure/function relationships were essentially based on primary structure analyses. However, quite recently, three 3D-structures of truncated GH70 glucansucrases differing by their linkage specificity were solved: GTF180-ΔN from *Lactobacillus reuteri* 180 (Vujicic-Zagar et al., 2010b), GTFA-ΔN from *L. reuteri* 121 (Pijning et al., 2012) and GTF-SI from *Streptococcus mutans* MT8148 (Ito et al., 2011). Two additional crystal structures of GH70 enzymes complete the picture, that of ΔN123-GBD-CD2, a representative of the branching sucrase subgroup specialized in dextran branching (Brison et al., 2012; Moulis et al., 2016), and that of Gtf-B-ΔNΔV from *L. reuteri* 121, a representative of 4,6-α-

glucanotransferase subgroup, which uses maltooligosaccharides as donor instead of sucrose (Bai et al., 2017). All structures have in common a 5 domain organization with a global U-shape folding. Domains A, B and C are also found in related GH13 enzymes, which form with GH70 and GH77 enzymes the clan GH-H, whereas domains IV and V are unique to the GH70 family. Except for domain C, these domains are formed by two discontinuous segments from both N- and C-terminal polypeptides. The catalytic domain A adopts a $(\beta/\alpha)_8$ barrel fold and contains the three catalytic amino acids involved in the α-retaining mechanism: the aspartate nucleophile and the glutamate acid-base catalyst involved in the β-D-glucosyl enzyme formation, along with the third aspartate acting as a transition state stabilizer (Leemhuis et al., 2013a; Meng et al., 2016a; Moulis et al., 2016). Sequence comparison combined with structural analyses also suggested the presence of residues involved in linkage specificity in this domain.

Our understanding at the molecular level of how glucansucrases control the size of their polymers is still limited. Indeed, some GSs such as the *L. mesenteroides* NRRL B-512F DSR-S and its variant DSR-S vardel Δ4N are efficient polymerases, producing almost uniquely very HMM glucans (10^8 g/mol) (Moulis et al., 2006b; Irague et al., 2012). Others, such as the dextransucrase GTF-180 from *Lb. reuteri* 180 or the alternansucrase ASR from *L. mesenteroides* NRRL B-1355 synthesize two glucan populations, a first one of very high molar mass (around 10^7 gmol) and the other comprising oligosaccharides or polymers of medium size (10^3-10^4 g/mol) (Côté and Robyt, 1982; Joucla et al., 2006; Meng et al., 2015a). Several studies suggest that domain V, located at the N and C-terminal extremity of the enzymes, plays a role in the polymer size through its capacity to bind glucans.(Kingston et al., 2002; Monchois et al., 1998a; Moulis et al., 2006b) Indeed, domain V fragments are able to fold independently and strongly bind α-glucans with dissociation constants in the nanomolar range(Kingston et al., 2002; Komatsu et al., 2007; Lis et al., 1995; Mori et al., 2011; Suwannarangsee et al., 2007). Called accordingly glucan binding domain (GBD), this domain V contains repeated motifs derived from a common YG motif (Giffard and Jacques, 1994). The available crystal structures of GH70 enzymes unfortunately mostly concern truncated constructs, in which only a small part of the native domain V is visible (Brison et al., 2012; Leemhuis et al., 2013a; Vujicic-Zagar et al., 2010b). Only very recently, several complexes of the ΔN123-GBD-CD2 branching sucrase with glucose, isomaltose (I2) or isomaltotriose (I3) provided the first structural evidence of sugar binding pockets with a common topology and directly interacting with carbohydrates in domain V (Brison et al., 2016). In *L. mesenteroides* NRRL B-512F DSR-S, successive suppressions of these repeat units induced a progressive loss of activity and decrease of the polymer size (Moulis et al., 2006b). The variant deleted of the entire domain V kept only 0.14% activity relatively to the native enzyme, totally lost its ability to produce HMM dextran (10^8 g/mol for the native form) and only synthesized low

molar mass (LMM) dextran of about 10^3 g/mol. In contrast, truncation of the domain V of GTF180-ΔN yielded a catalytically quasi-fully active enzyme (only 25% reduction) but induced a change of the distribution between oligosaccharides and HMM dextran populations, a lower amount of glucosyl units being incorporated in HMM polymers (2% for GTF180-ΔN-ΔV versus 16.3% for GTF180-ΔN) to the profit of oligosaccharide synthesis. However, the size of the HMM dextran produced by GTF180-ΔN-ΔV remains unchanged (2.10^7 g/mol). In addition, the mutation of Leu940 located in the acceptor substrate binding site of GTF180-ΔN-ΔV partially restored its capacity to synthesize HMM polysaccharide showing that both the active site and the domain V are important to control polymer size (Meng et al., 2015a). Overall, these findings suggest that the role of domain V may vary from one enzyme to another, but structural data to fully understand the interplay between the various domains during polymerization and their implication in the control of the polymer size are still lacking.

To address these questions, the enzyme DSR-M, recently isolated from *L. citreum* NRRL B-1299, is of particular interest (Passerini et al., 2015). Indeed, compared to all GS characterized to date, this enzyme is the first one reported to naturally synthesize linear LMM dextrans directly from sucrose. As sequence analysis did not reveal any obvious discriminating features that could explain this specificity, we report here the three dimensional X-ray structures of the protein alone, or in complex with either sucrose or isomaltotetraose (I4: [α-D-Glcp-(1→6)]$_3$-Glcp) which were solved at 3.2 Å, 3.6 Å and 3.7 Å, respectively. The latest displays the largest structure of a domain V reported so far, in which 4 putative binding pockets with a topology close to that described in ΔN123-GBD-CD2 branching sucrase could be identified. The complex with I4 further revealed a sugar molecule in one of these pockets, providing a rational basis to explore the role of the GBD in determining the length of the resulting dextrans. Specific loop structures around the catalytic site were further examined by site-directed mutagenesis, and their functional impact was evaluated by analysis of the resulting polymers. From our data, a distributive mode of polymer formation emerges for DSR-M, with an involvement of the domain V to the length of the resulting dextrans.

II. Results

II.1. Design and characterization of truncated variants DSR-MΔ1 and DSR-MΔ2

As previously reported, the 229 kDa DSR-M enzyme is among the largest of the GH70 family (Passerini et al., 2015). Based on sequence alignment, the enzyme is predicted to adopt a U-shape fold resulting in an organization in five domains (Figure 26). In domain V, YG repeats are found at both N- and C-terminal extremities. One particularity is the C-terminal extremity of the protein, where YG repeats are followed by another series of APY repeats nearly identical to those found in the C-terminal region of *L. citreum* NRRL B-1355 alternansucrase (ASR) and *L. citreum* CW28 inulosucrase (Passerini et al., 2015).

Initial attempts to overexpress the *dsr-M* encoding gene in *E. coli* all led to the production of proteins degraded at their N- and C-terminal extremities. To circumvent this problem, shorter genes were designed and cloned into *E. coli*. DSR-MΔ1 encodes the DSR-M enzyme devoid of its signal peptide and of the last 632 amino acids at its C-terminal extremity (Figure 26). A similar C-terminal truncation of the ASR alternansucrase did not change its enzymatic activity and/or specificity (Joucla et al., 2006). Production levels of DSR-MΔ1 in *E. coli* BL21 star DE3 reached 10,000 U per liter of culture, with a specific activity of 60 U/mg of protein. The enzyme was purified to homogeneity for crystallization, but SDS-PAGE analysis of the first crystals revealed protein degradation at its N-terminus, also confirmed by Edman sequencing. Thus, an additional removal of 148 amino acids from the DSR-MΔ1 was attempted to construct the DSR-MΔ2 form (Figure 26). For DSR-MΔ2, we obtained 10,000 U/L of culture and determined a specific activity of 67 U/mg of protein.

With 292 mM sucrose as a starting substrate, HPSEC and HPAEC-PAD reveal that DSR-MΔ1 and DSR-MΔ2 both produce low molar mass dextrans without any trace of very high molar mass polymer (>10^6 g/mol). Comparing the polymer distribution of full-length DSR-M and its truncated versions, polymers slightly shorter than those produced by the full enzyme are produced in higher amounts with DSR-MΔ1 and DSR-MΔ2, indicating that the introduced truncations only slightly affect the global product profile (Figure 27, Table S3). The sucrose hydrolysis ratio is very low (inferior to 4%) for all three enzymes. Moreover, ^{1}H NMR spectra further confirm that these LMM dextrans all contain more than 99% of α-(1→6) linkages, so the truncation does not change the linkage specificity (Passerini et al., 2015). On this basis, DSR-MΔ2 was considered as a good model to pursue the biochemical and structural characterization.

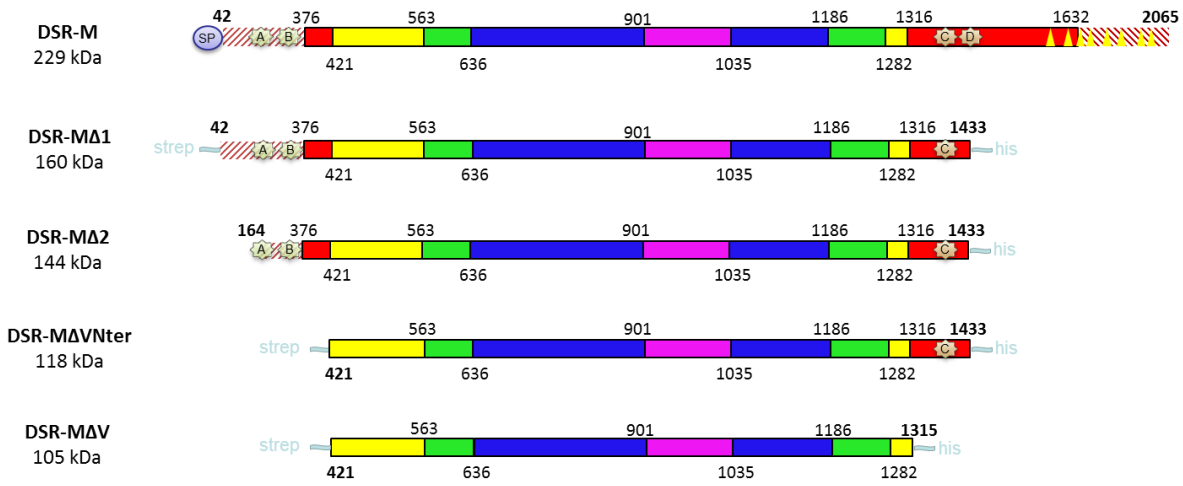


Figure 26: **Schematic structural organization of DSR-M and the truncated variants,** based on amino acid alignment with GTF-180-ΔN and structural analysis. Hatched red lines correspond to non-aligned zones. DSR-M structural domains: i) domain V in red, ii) domain IV in yellow, iii) domain A in blue, iv) domain B in green and v) domain C in magenta. Blue circle represents the signal peptide, APY motifs are indicated with triangles and pale yellow stars marked as A, B, C and D represent putative glucan binding pockets.

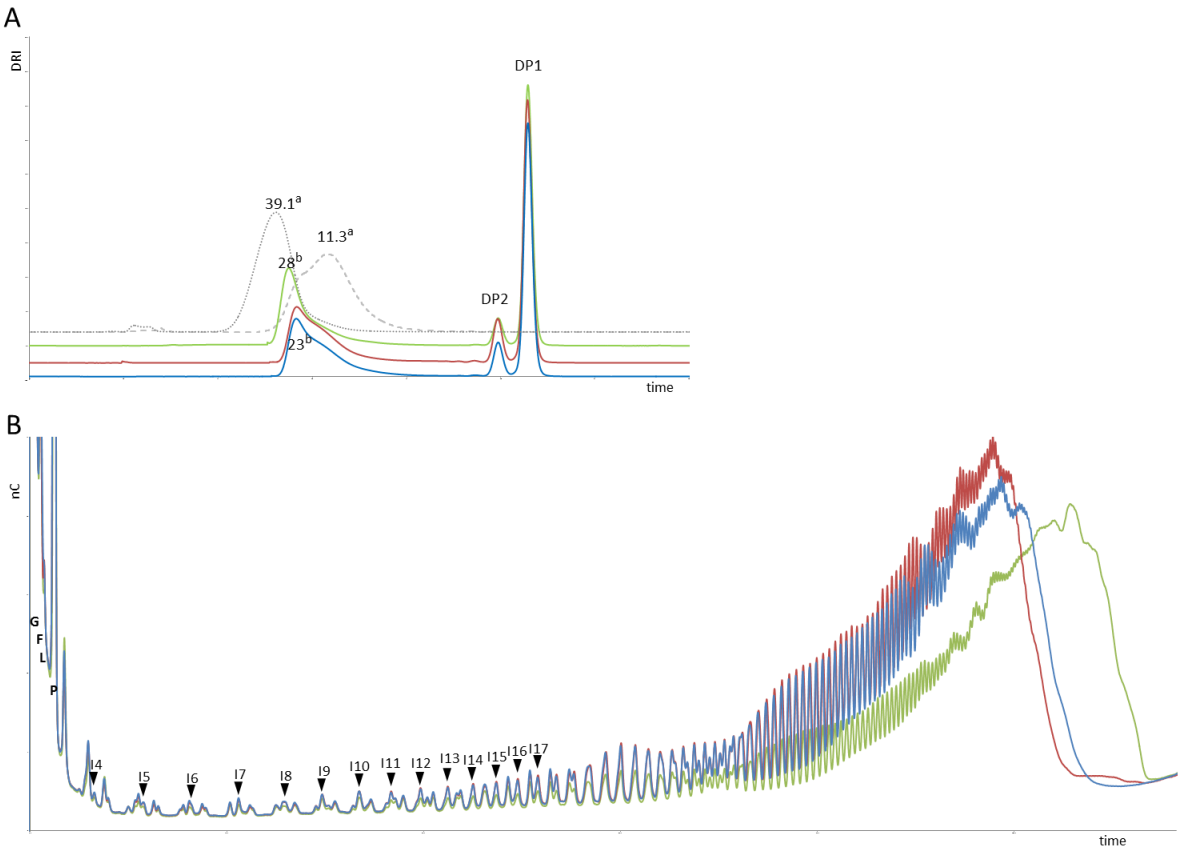


Figure 27: **Analysis of the products synthesized by DSR-M variants from 292 mM sucrose.** A, HPSEC chromatogram after total sucrose consumption, (green, DSR-M; red, DSR-MΔ1; blue, DSR-MΔ2; grey dash, commercial dextran 11300; grey points, commercial dextran 39100). DP1, monosaccharides; DP2, disaccharides. B, HPAEC-PAD profile of DSR-MΔ2; G, glucose; F, fructose; L, leucrose; P, palatinose; I4 to I20, isomaltooligosaccharides of DP4 to DP20. (a) molar mass of commercial standards in kg/mol (b) determined molar mass at peak apex in kg/mol.

II.2. Kinetics of polymer formation reveal that DSR-M Δ2 can accept different chain initiators

HPAEC monitoring of sucrose consumption and polymer synthesis catalyzed by DSR-MΔ2 from 292 mM sucrose (Figure 28) shows a constant sucrose consumption rate of ≈0.76 mM/min during the first 10 min of reaction, corresponding to 2% of sucrose consumption. In a second phase, from 15 min to 90 min, it substantially increases to 3.3 mM/min before gradually diminishing until the end of the reaction that is reached within less than two hours (Figure 3). Oligosaccharides synthesized during the first reaction step (until 10-15 min) are not isomaltooligosaccharides (IMOS), as judged from their retention times. After incubation with an invertase, they were digested into IMOS (Figure S2). Sucrose hence is present at the extremity of each oligosaccharide, acting as the chain initiator being glucosylated through the formation of an α-(1→6) linkage onto its glucosyl unit. After the initial reaction phase, the concentration of oligosaccharides of higher DP significantly increases, and IMOS can be detected, showing that glucose (obtained after release onto water) can also be used as chain initiator. All these produced polymers were efficiently digested by dextranase (Figure S3), confirming the very high content of α-(1→6) linkages determined by NMR. Near the final phase, after 75 min of reaction, two thirds of sucrose are consumed leading to a significantly increased fructose concentration. Two additional series of oligosaccharides result from the initial glycosylation of this fructose (leading to sucrose isomers such as leucrose for example) and their subsequent elongation (Figure S4). At the end of the reaction, 84% of the glucosyl units issued from sucrose are recovered in polymers, 3% in glucose (hydrolysis activity), 8% in leucrose, and 5% in other oligosaccharides of unknown structure and of DP inferior to 12.

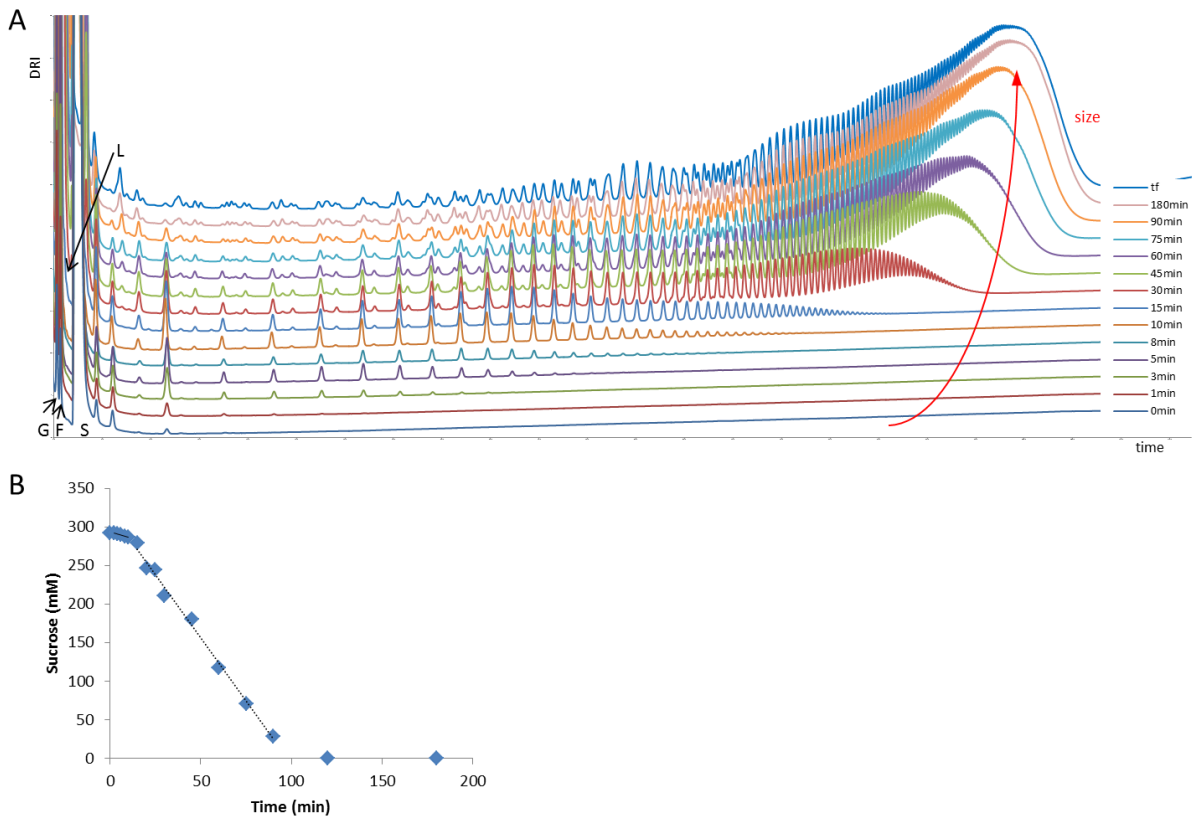


Figure 28: **Monitoring of oligosaccharide and polysaccharide production during polymerization catalyzed by DSR-MΔ2.** A, HPAEC-PAD chromatograms during the polymerization reaction. G, glucose; F, fructose; S, sucrose; L, leucrose. B: Sucrose consumption during the reaction.

II.3. 3D structures of DSR-MΔ2 and DSR-MΔ2 E715Q in complex with sucrose or isomaltotetraose

The structure of the free DSR-MΔ2 (PDB entry: 5LFC) was solved by molecular replacement using GTF-SI as template (PDB entry: 3AIE (Ito et al., 2011)). In this structure, the 108 residues at the N-term extremity of the domain V were not clearly visible. Two other structures were then obtained DSR-MΔ2 E715Q in complex with either sucrose in the active site (PDB entry: 5O8L) or isomaltotetraose (I4) (PDB entry: 5NGY). In this latter structure and despite a low resolution, we could manually build the entire N-term part of the domain V in which an I4 molecule could be identified (Figure 29 and Figure S1). As for the other GH70 enzymes of known structures, the polypeptide chain adopts a "U-shape" folding and is organized in five domains: domains A comprising a $(\beta/\alpha)_8$ barrel, domain B (from β-strand 3 to α-helix 3), and domains C, IV and V (Vujicic-Zagar et al., 2010b; Ito et al., 2011; Pijning et al., 2012; Bai et al., 2017; Brison et al., 2012). Among them, only domain C is formed by a continuous fragment (Figure 29). The final models of DSR-MΔ2 and DSR-MΔ2 E715Q in complex with sucrose comprise residues 288 to 1436 and 334 to 1433, respectively. In the structure of DSR-MΔ2

E715Q I4-complex, residues 171 to 1435 are visible, making it the GH70 structure with the most complete domain V solved to date.

The active site of DSR-MΔ2 in domain A displays a pocket shape at the bottom of a large cavity located at the interface of domain A and B, near which a calcium binding site can be identified at the same position as in the other GH70 glucansucrases (Vujicic-Zagar et al., 2010b; Brison et al., 2012). Superposition of the structures of the DSR-MΔ2-sucrose complex with that of the GTF180-ΔN-sucrose complex (PDB: 3HZ3 (Vujicic-Zagar et al., 2010b)) shows that conserved residues of the subsite -1 (Arg675, Asp677, Glu715 His789, Asp790, Asn1068, Asp1118, Tyr1127, Asp1178 and Gln1183) and +1 (Gln794, Trp717 and Asn681) align well with a 0.23 Å root mean square deviation of the Cα. Asp677, Glu715 and Asp790 hence act as the putative nucleophile, general acid/base catalyst and transition state stabilizer, respectively (Figure 30).

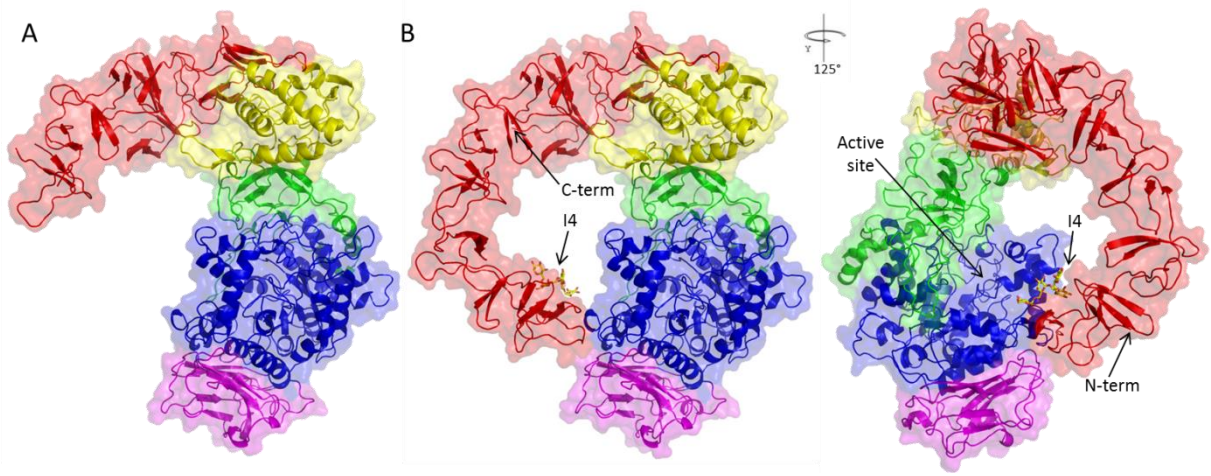


Figure 29: **DSRM-Δ2 structures**. A, structure of DSRM-Δ2 solved at 3.2 Å resolution. B, structure of DSRM-Δ2 E715Q in complex with I4 (inactive mutant) solved at 3.7 Å resolution. Magenta, domain C; blue, domain A, which includes the (β/α)8 barrel; green, domain B; yellow, domain IV; red, domain V.

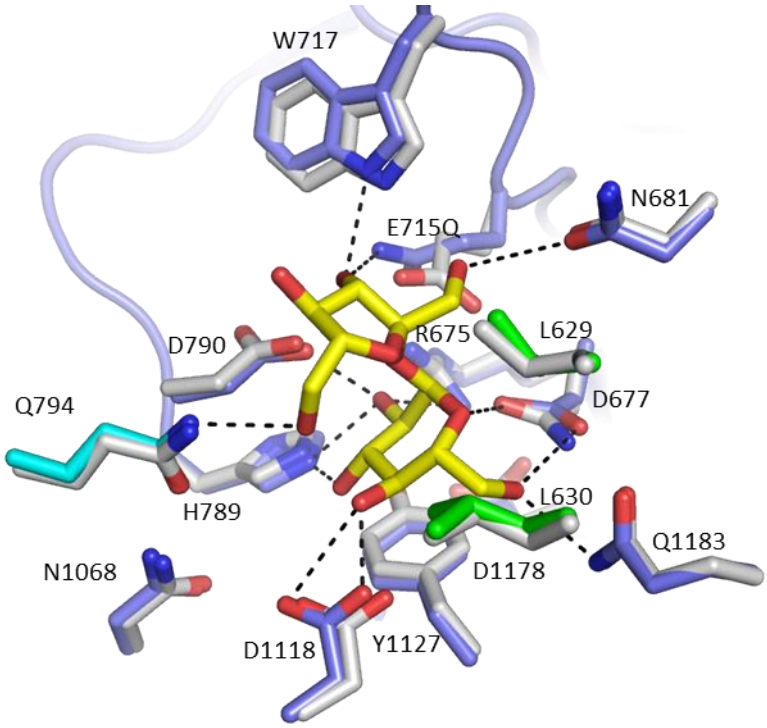


Figure 30: **Superposition of subsites -1 and +1 of DSR-MΔ2 E715Q: sucrose-complex with subsites -1 and +1 of GTF180-ΔN: sucrose-complex**. The DSR-M catalytic residues are Asp677 (nucleophile), Glu715Gln (acid/base), and Asp790 (transition state stabilizer). Sucrose is shown with yellow carbons. Residues of the inactive GTF180-ΔN mutant (D1025N) that interact with sucrose are represented in gray. The carbon atoms of their structural equivalents in DSR-MΔ2 are shown in blue (domain A), cyan (subdomain H1-H2), and green (domain B).

A distinct feature of DSR-MΔ2 is the presence of four loops surrounding its catalytic cleft, namely loop A1 (located at the N-term part between strand β7 and helix α7), loop A2 connecting strand β2 and helix α2 (Figure 31), and loops B1 / B2 (found in the N-terminal part of domain B between strand β3 and helix α3).

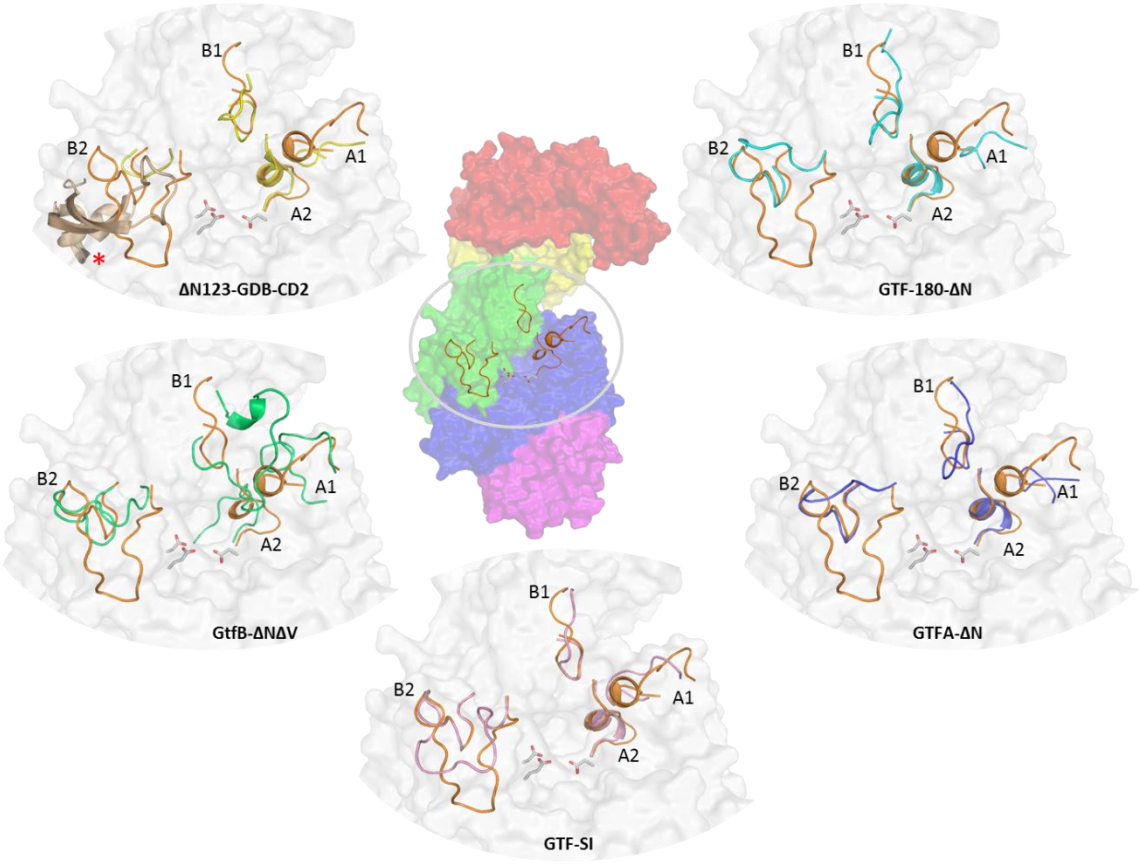


Figure 31: **Secondary structure differences between DSR-MΔ2 and other GH70 enzyme structures**. Center: DSR-MΔ2 (same colors than Figure 29) loops A1 (805-820), A2 (1115-1125), B1 (568-578) and B2 (601-626) in orange. Sides: in yellow, green, pink, blue and cyan, loops A1, A2, B1 and B2, of ΔN123-GDB-CD2 (3TTQ), GtfB-ΔNΔV (5JBD), GTF-SI (3AIE), GTFA-ΔN (4AMC) and GTF-180-ΔN (3KLK), respectively; surface of DSR-MΔ2 in grey; catalytic residues are represented as grey sticks. *Red star: special loop of ΔN123-GDB-CD2 (2731-2796) covering the area of loop B2.*

Loop A1 of DSR-MΔ2 contains a small H1' helix (residues Val812-Leu817) inserted between the two helices of the subdomain H1 – H2 common to other GH70 enzymes. In DSR-MΔ2, this subdomain comprises residues from Gln793 to Gln836 and is located between strand β7 and helix α7 of the $(\beta/\alpha)_8$ barrel. To assess the functional relevance of this insertion, Asp813 and Leu816 - two residues of H1' pointing towards the active site - were mutated into alanine, and the reaction products of the mutant enzymes were analyzed. Although the specific activity of both mutants was two-fold lower, the resulting product characteristics were not significantly impacted by the mutations (Figure S5, Table S3). When comparing the structures of DSR-MΔ2 sucrose-complex (or of the glycerol complex) with that of the I4-complex, we notice a shift (around 2 Å) of the H1 – H2 subdomain with a concomitant reorientation of Gln794 side chain. This latter interacts with the glycerol molecule or fructosyl ring of sucrose in DSR-MΔ2 sucrose-complex, respectively, but it is pointing away from the active site in the I4-complex (Figure 32 and Figure S1). Interestingly, a rearrangement of this H1 – H2 subdomain with a conformational change of the equivalent Gln2326 side chain (equivalent to Gln794

of DSR-MΔ2) was also observed in the structure of ΔN123-GBD-CD2: isomaltotriose-complex (4TTU) compared with the apo-enzyme (Brison et al., 2016).

Loop B1 is slightly shorter in DSR-MΔ2 and GTF-SI than in GTF180-ΔN and ΔN123-GBD-CD2 and is very different from its equivalent in GTF-B. One particularity of B1 in DSR-MΔ2 is its proximity with loop A1 due to the presence of the extra helix H1'. As the loop B1 residue Leu940 of GTF180-ΔN, oriented towards the +1 subsite, was shown (Meng et al., 2014, 2015a) to be involved in almost every aspect of the polymerization - linkage specificity, hydrolysis activity, size determination and polysaccharide versus oligosaccharide synthesis ability - the equivalent residue Leu575 in DSR-MΔ2 was replaced by a tryptophan. The mutation resulted in a 50% reduction of dextran molar mass (Table S3) but did not change the linkage specificity. Finally, loop B2 (582–632) of DSR-MΔ2 fills a larger space than in the other GH70 enzymes, rendering the binding cleft around the active site less open. Notably, the empty space left by the short (25 amino acids) loop B2 of ΔN123-GBD-CD2 is filled by another long loop emerging from the C-term part of domain B from 2731 to 2796 (Figure 31).

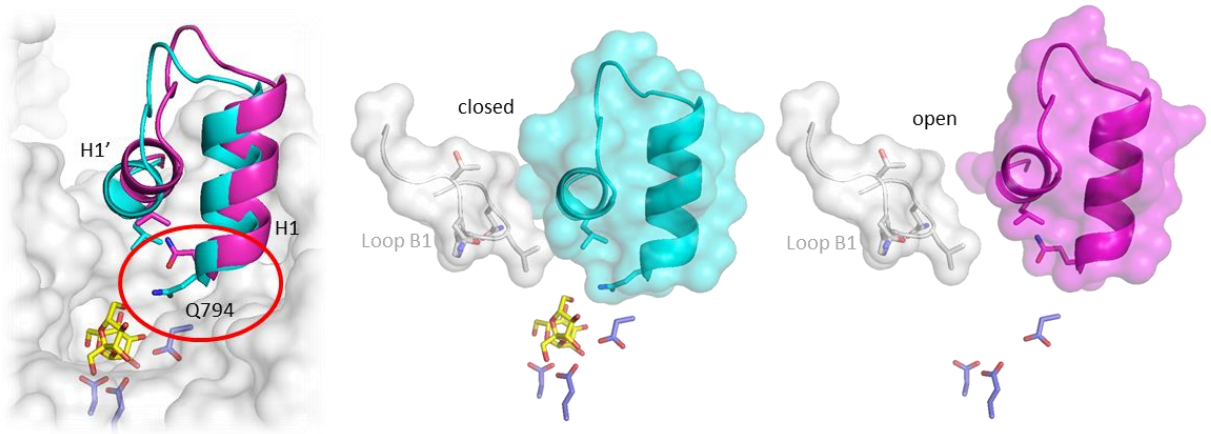


Figure 32: **Alternative conformation of Gln794 and subdomain H1 – H1' motion upon sucrose binding in the active site**. Left: Superposition of inactive mutant structures in complex with sucrose (cyan) and in complex with I4 (magenta), illustrating the movement of the subdomain H1-H2 and the alternative conformation of Gln794 when sucrose is bound in the active site. "Closed" (middle) and "open" (right) conformation of the catalytic cleft are represented. DSR-MΔ2 surface is shown in grey, catalytic residues and sucrose are represented as blue and yellow sticks, respectively.

Domain IV of DSR-MΔ2 superimposes well with the corresponding domain of GTF180-ΔN (rmsd, 1.2 Å on Cα), and secondary structure elements are well conserved except for a β-hairpin (residues 547-560) at the N-terminal border between domain IV and B, unique to DSR-MΔ2.

Domain V of DSR-MΔ2 (5LFC), comprising the first 249 and the last 117 residues located respectively at the N- and C-term ends of the protein, adopts the same orientation as in ΔN123-GBD-CD2 and in the orthorhombic apo-form of GTF180-ΔN (PDB entry: 4AYG (Pijning et al., 2014)), with its

extremities pointing to the catalytic $(\beta/\alpha)_8$ barrel. Our structure of the DSR-MΔ2 E715Q:I4-complex, including residues 171 to 1435 and hence the largest domain V solved so far, shows a similar orientation for domain V. Residues 178-179 and 194-195 at its tip are located at a short distance of 5 Å to helix α6 and α7 of the central $(\beta/\alpha)_8$ barrel that harbors the catalytic site. Several Praseodymium ions that intercalate in the crystal packing potentially help to stabilize this domain V. The observation that the GBD is well structured in the E715Q:I4-complex crystal structure but not in the 2 other structures presented, prompted us to investigate the solution structure of DSR-M by small angle X-ray scattering (SAXS). SAXS curves qualify DSR-MΔ2 as a monomeric protein with a radius of gyration (Rg) of 40 Å and a maximum interparticle distance (Dmax) of ~125 Å, closely matching the theoretical values calculated from the crystal structure of the I4-complex (39.5 Å and 117 Å, respectively) (Figure 33). Indeed, rigid-body fitting of the crystal structure, with modelling of flexible N- and C- terminus, resulted in a good fit to the experimental data (X=1.8), suggesting that the observed crystal conformation is highly probable in solution. On the other hand, observing the ab-initio SAXS envelop, it seems that the protein in solution adopts a "horse shoe" shape with the bend point located between domains IV and V. Although we could imagine a certain flexibility of the domain V, the Kratky analysis suggests that the particle is globally quite compact.

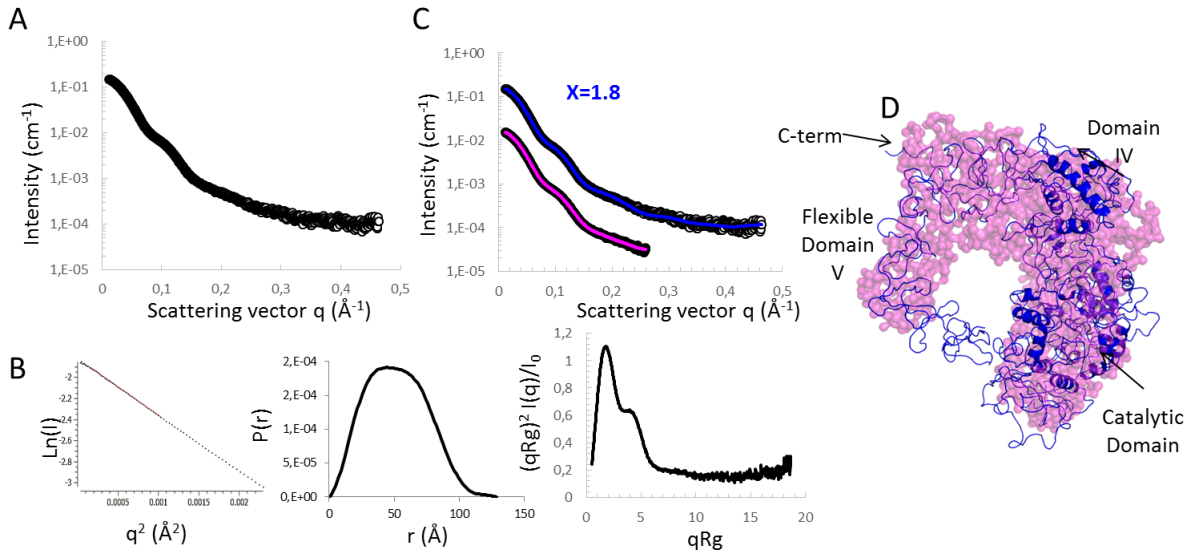


Figure 33: **SAXS analysis.** A) Scattering curve of DSR-MΔ2. B) Guinier fit, P(r) function and dimensionless Kratky plot. C) Fits of the ab-initio envelop (purple dots) and of the solution model (blue dots) against the experimental data (black circles). D) Superposition of the ab-initio envelop (purple) and of the solution model (blue).

II.4. Functional implications of domain V

We observed a large part of domain V in the structure of the DSR-MΔ2 E715Q:I4-complex, and we could attribute residual electron density to an isomaltotetraosyl residue (Figure 34) in a pocket

named pocket V-A comprising residues 172 to 245 of domain V. This pocket adopts a similar topology to that of the sugar binding pockets recently characterized (Brison et al., 2016) in the branching enzyme ΔN123-GBD-CD2. With all the precautions that should be taken considering the low resolution of our structure, our data enabled the identification of an aromatic residue (Tyr180) in stacking interaction with the second glucosyl unit of the oligosaccharide (Glc-2, numbering starting from the reducing end). In addition, the O2 atom of Glc-2 also interacts with Gln217 and Lys219, its O3 atom with Gln208, Gln217 and Leu236 (via the oxygen atom of its main chain carbonyl group) and its O4 atom with Tyr238. The side chain of Lys219 also interacts with the oxygen atom of the osidic bond between Glc-1 and Glc-2, and with the O5 of Glc-1. Glc-3 seems to be weakly stabilized by van der Waals interactions with Tyr180 and Tyr187. Finally, the O2 atom of Glc-4 interacts with Tyr180 (via the oxygen atom of its main chain carbonyl group). Sequence alignments of DSR-M domain V with the repeats corresponding to the binding pockets of DSR-MΔ2 and ΔN123-GBD-CD2 (Figure 35) allowed the identification of three other repeats that we designate as repeats V-B, -C and -D (Figure 26 and Figure 35). Despite the low resolution, our data indicate that pocket V-A and putative pockets V-B and V-C share a similar topology (Figure S6). However, in putative pocket V-C, the residues equivalent to the conserved Tyr180, Gln217 and Lys219 of pocket V-A are replaced by Phe1342, Glu1387 and Arg1389. Dextran binding ability was further investigated by affinity gel electrophoresis in the presence of 68.4 kg/mol dextran (Figure S7). Similar to ΔN123-GBD-CD2, a clear delay of migration was visualized for DSR-MΔ2, while no significant delay was observed for a construct devoid of the repeats (DSR-MΔV, see below). Finally, DSR-MΔ2 affinity for I6 (used in crystallization assays) was corroborated by ITC analysis, which gave an approximate Kd_{I6} value of 10 mM (Figure S7).

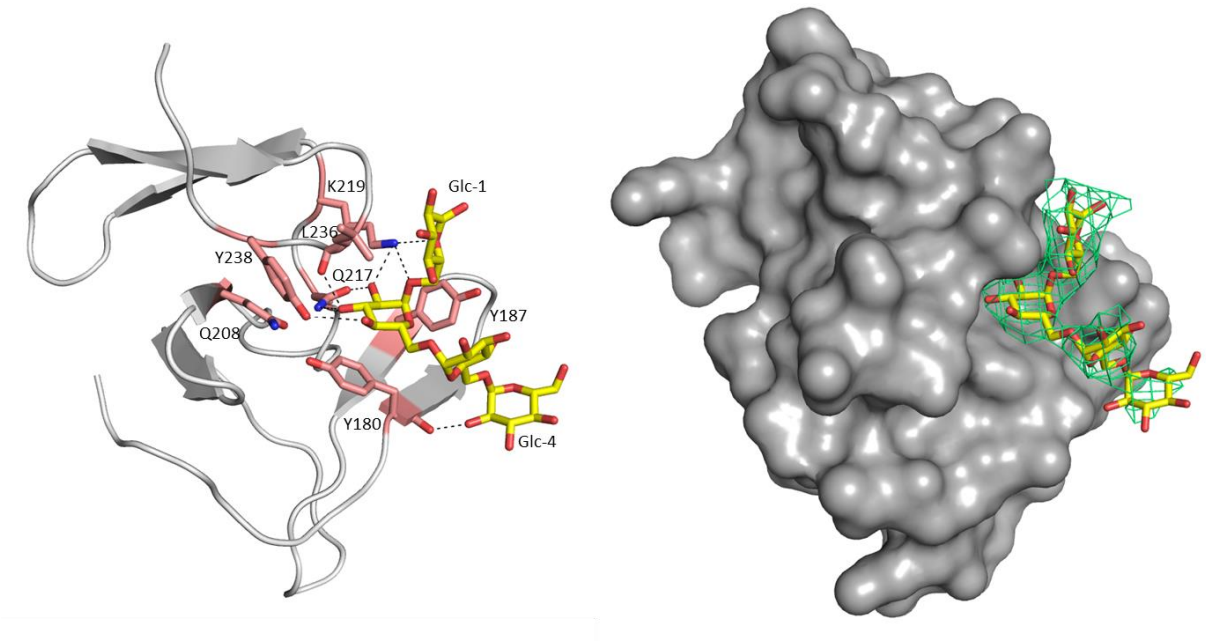


Figure 34: **Binding pocket V-A found in the GBD of DSR-MΔ2.** An isomaltotetraosyl residue is shown as *yellow sticks*. The network of hydrogen bonding is shown on the *left panel*, whereas the electron density map around carbohydrates is displayed on the *right panel*. The residues involved in binding are represented as *pale red sticks* (*left panel*). The difference electron density map (*Fo - Fc*) around carbohydrates was contoured at 3σ (in *green*).

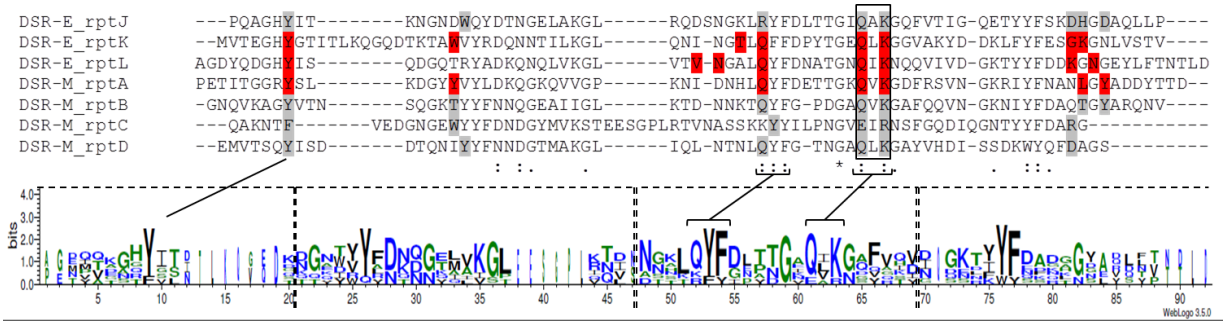


Figure 35: **Sequence alignment of the 3 repeats identified in the GBD of ΔN123-GBD-CD2 (from DSR-E) with those identified in DSR-MΔ2 representing putative glucan binding pockets.** *red* highlighted residues are involved in sugar binding in repeats K and L of ΔN123-GBD-CD2 (DSR-E) and A of DSR-M, whereas *gray* highlighted residues are proposed to play the same role in other repeats. The "QxK" motif is framed. A LOGO sequence based on this alignment is shown. The YG repeats are framed in *dashed lines* on the LOGO sequence.

To assess the contribution of domain V to the polymerization process, we further truncated DSR-MΔ2 and deleted the putative sugar binding pockets V-A and V-B located at the N-terminal part of domain V (DSR-MΔV-Nter, residues 421 to 1433, Figure 26). With comparable specific activity (70 U/mg of protein) to that of DSR-MΔ2, the molar mass of the dextran polymers produced by DSR-MΔV-Nter from 292 mM sucrose (16 kg/mol) proved significantly lower than that of DSR-MΔ2 dextran (23 kg/mol) (Figure 36, Table S3) without affecting the protein polymer synthesis ability (still 85% of glucosyl units transferred into dextran). In a second construct, domain V including the putative sugar

binding pocket V-C was removed (DSR-MΔV, positions from 421 to 1315, Figure 26). Again, DSR-MΔV retained a specific activity of 70 U/mg of protein, and 85% of the glucosyl residues from sucrose were incorporated into dextran polymers of 16 kg/mol (Figure 36, Table S3). The additional removal of the putative C-terminal pocket V-C hence did not impact the size of the dextran produced. Finally, HPSEC analysis of the acceptor reaction products obtained from 39 kg/mol commercial dextran and 292 mM sucrose shows that both DSR-MΔ2 and DSR-MΔV enzymes, in addition to the ab initio synthesis produced directly from sucrose, are able to further glucosylate exogenous dextran molecules to a high polymerization degree that is slightly lower for DSR-MΔV than for DSR-MΔ2 (Figure S8).

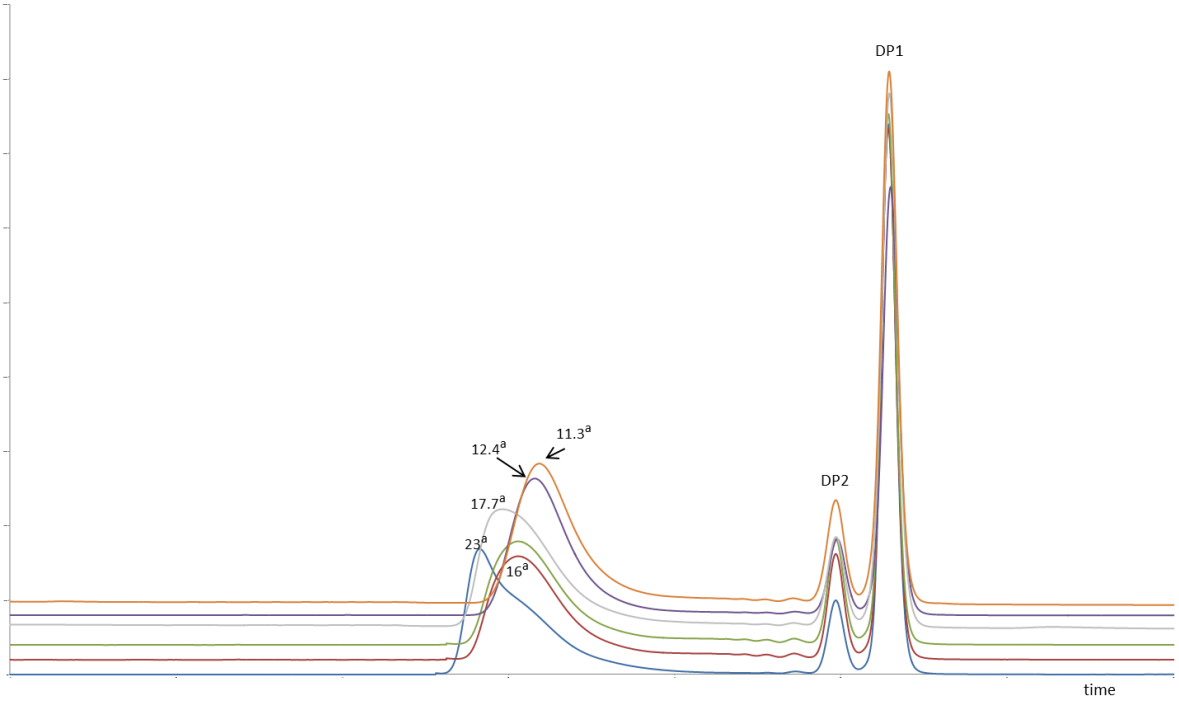


Figure 36: **HPSEC analysis of the products synthesized by DSR-M domain V variants from 292 mM sucrose.** Blue, DSR-MΔ2; red, DSR-MΔV; green, DSR-MΔV-Nter; grey, DSR-MΔ2-Y264A; purple, DSR-MΔ2-Y180A; orange, DSR-MΔ2-YY180-264AA. [a] *determined molar mass at peak apex in kg/mol*

Because truncation of the putative binding repeats might have a long range effect by destabilizing the horse shoe structure, we mutated individual amino acids implicated directly in the sugar binding. Tyr180 and Tyr264 of pocket V-A and pocket V-B were obvious candidates for mutation, and we constructed two single (DSR-MΔ2-Y180A, DSR-MΔ2-Y264A) and one double mutant (DSR-MΔ2-YY180-264AA). The single mutation in the pocket V-A (DSR-MΔ2-Y180A) had a pronounced effect on the final dextran size, even more important than the deletion of the whole domain V, with dextrans of only 12 kg/mol versus 16 kg/mol for DSR-MΔV (Figure 36, Table S3). Single mutant DSR-MΔ2-Y264A (mutation in pocket V-B) produced dextrans of 18 kg/mol molar mass, also contrasting with the 23 kg/mol of DSR-MΔ2 dextran. The double mutant (DSR-MΔ2-YY180-264AA) produced dextran

chains comparable to those of the DSR-MΔ2-Y180A single mutant, with only a marginal additional decrease for the mean molar mass of the resulting dextrans.

III. Discussion

In comparison with the other characterized glucansucrases from GH70 family, DSR-M from *L. citreum* NRRL B-1299 is presently the sole enzyme able to produce, from sucrose, only short dextrans of low molar mass. It furthermore presents an excellent efficiency, with 85% of the transferred glucosyl units being incorporated into polysaccharides. DSR-M hence appears as a good enzyme for the production of LMM dextrans directly from sucrose in a single-step eco-friendly process. Why DSR-M limits the polymerization process to the synthesis of dextran chains of around 30 kg/mol molar mass whereas other dextransucrases such as DSR-S vardel Δ4N or GTF-180ΔN produce polymer chains of more than several millions of g/mol was the central question addressed in this study.

For semi-processive enzymes such as DSR-S vardel Δ4N (a glucansucrase deleted of a part of the variable region located at N-term and of 4 N repeats at C-term) (Moulis et al., 2006b), HMM polymers (10^8 g/mol) can already be observed at the early stage of synthesis. For DSR-MΔ2, the polymer size gradually increases with time, indicating a chain elongation mode typical of distributive enzymes. We further show that, during the first 10 min of the reaction, sucrose serves as a primer and velocity is low and constant. After this initial stage, the sucrose-consumption rate increases fourfold and remains constant till substrate depletion (Figure 28). This reveals that glucosylation of long chain acceptors is faster than that of sucrose or glucose and further suggests that enzyme deglucosylation may be more efficient in the presence of acceptors of a superior length. However, above a certain size, the length of the acceptor chain does not further impact the glucosyl transfer velocity, explaining why the sucrose consumption remains constant and reflecting the extreme efficiency of DSR-MΔ2 to elongate oligodextrans. Overall, from this biochemical analyses, the ability of DSR-M to produce exclusively LMM dextran could be attributed to i) a high propensity to use sucrose as both donor and acceptor at the early stage of the reaction, thereby generating a high number of growing dextran chains and ii) an equivalent ability to elongate oligodextrans irrespective of their size. The X-ray structures of the enzyme and the inactive mutant DSR-MΔ2 E715Q were solved. Globally, DSR-MΔ2 is structurally organized in five domains as the other GH-70 enzyme structures solved so far. However, in domain A, several loops surrounding the catalytic cleft are different from those exhibited by the other GH70 enzymes. Inserted between the two helices of the H1 - H2 subdomain, loop A1 of DSR-MΔ2 is atypical (Figure 31 and Figure 32). Interestingly, sucrose accommodation in the active site induces a re-orientation of Gln794 side chain, and leads to a motion of the complete H1 – H2 subdomain. Such a motion was already observed for ΔN123-GBD-CD2 in complex with isomaltosyl

residues (4TTU)(Brison et al., 2016) and the equivalent glutamine in GTF-180ΔN (Gln1140) was shown to be implicated in linkage specificity, molar mass and branching of the produced polymer (Van Leeuwen et al., 2009). In DSR-MΔ2, it could assist sucrose accommodation, fructose release and acceptor positioning, and reflects the functional importance of loop dynamics during catalysis. Loop B1 is another important loop, located at the upper part of the catalytic cleft and near the acceptor subsite +1. In GTF180-ΔN, loop B1 residues Leu938 and Leu940 were shown to influence the linkage specificity, hydrolysis activity and polymer size (Meng et al., 2014). The mutation of Leu575 in DSR-MΔ2 (equivalent to Leu940 of GTF-180ΔN) reduced by 50% the polymer size, indicating a possible role of this residue and by extension of loop B1 in dextran acceptor accommodation. Finally, loop B2 contributes to the width of the binding cleft: either open, in GTF180-ΔN and GTFA-ΔN, or more compact like in GTF-SI and DSR-MΔ2. In DSR-MΔ2, it is exceptionally long, potentially contributing to the protection of the glucosyl-enzyme intermediate from water attack and thereby favoring transglucosylation. Indeed, sucrose hydrolysis by DSR-MΔ2 is really minor (3%) compared to 24% and 10% for the reactions catalyzed by GTF180-ΔN and GTFA-ΔN, respectively (Meng et al., 2014, 2015c). Clearly, the loops shaping the catalytic cavity are determinant for the efficacy and specificity of DSR-MΔ2 and could represent interesting targets for further enzyme engineering.

The X-ray-structure of the entire domain V of DSR-MΔ2 inactive mutant was determined, in complex with I4 (Figure 29). In the crystal, the enzyme looks like a horse shoe, and the SAXS data in solution indicate that this is not a crystallization artefact but rather reflects the flexibility of domain V, as previously demonstrated for that of GTF180-ΔN (Pijning et al., 2014). The proximity of domain V with the active site indicates a possible interplay of these two domains, which is further supported by the identification of an IMOS binding pocket at the N-term of DSR-MΔ2. This is the first structural evidence of carbohydrate-protein interactions in domain V of a polymerase, and the binding pocket resembles that recently identified in the branching sucrase ΔN123-GBD-CD2 in complex with isomaltosyl residues. Residues Tyr180, Gln208 and Lys219 of this V-A pocket interact with different atoms of the oligosaccharide (Figure 34), and play the same role as Tyr1834, Gln1879 and Lys1881 in the pocket V-K of the branching sucrase ΔN123-GBD-CD2 (Brison et al., 2016). Further sequence and structure analysis of DSR-M domain V enabled the identification of three supplemental candidates as sugar binding pockets (Figure 35). Isothermal titration calorimetry, as well as affinity gel electrophoreses, confirmed that DSR-MΔ2 affinity for dextran chains is mainly mediated by domain V, as no binding to dextran could be detected for DSR-MΔV (Figure S7). However, compared to the nanomolar affinity estimated for dextran binding of GTF-I and DSR-S domains V fragments(Komatsu et al., 2007; Suwannarangsee et al., 2007), DSR-MΔ2 Kd values were estimated at 10 mM for IM6 and

300 µM for 68.4 kg/mol dextran, respectively, that are extremely high but in agreement with the non-processive character of the enzyme.

Finally, to assess the functional role of this domain V, we compared the enzymatic activity of various truncated or point mutants with that of DSR-MΔ2. Deletion of one or more of the pockets does not affect enzyme specific activity, as well as its polymer synthesis ability. These results contrast with those reported (Meng et al., 2015a; Moulis et al., 2006b) for polymerases synthesizing HMM polymers such as DSR-S vardel Δ4N and GTF180-ΔN, for which the deletion of the entire domain V resulted in a quasi-total or 25% loss of activity, respectively. However, elimination of the entire domain V of DSRM-Δ2 causes a reduction of the LMM dextran size distribution from 23 kg/mol (DSR-MΔ2) to 16 kg/mol (DSR-MΔV). Binding to domain V thus enhances the ability of the enzyme to synthesize longer chains, and suggests that the proximity of domain V to the active site allows a possible interplay of these two domains in enzyme catalysis. As elimination of pockets V-A and V-B in DSR-MΔV-Nter, or pockets V-A, V-B and V-C in variant DSR-MΔV has the same impact on the size of the produced dextrans, pocket V-C at the C-terminal extremity of domain V may not be functional. Closer inspection of its structure and sequence revealed that the QxK motif is replaced by the EIR sequence in this putative pocket. Then, presence of this motif (QxK) together with the spatial proximity of a Tyr might be a defining signature for a functional binding pocket. The single mutations Y180A of pocket V-A and Y264A of pocket V-B result in mutants producing dextrans of lower molar mass than that of DSR-MΔ2, highlighting the crucial role of the tyrosine in oligosaccharide binding and elongation.

IV. Conclusion

Overall, our extensive biochemical and structural analyses reveal that DSR-MΔ2 adopts a clear non-processive mechanism - with a preference for sucrose as initial acceptor of D-glucosyl units, then an ability to elongate dextran growing chains irrespective of their size - that could be due to singularities in the loops delineating the catalytic cleft. We also report here the most complete 3D-structure of a GH70 family member, with a large part of its domain V solved in complex with an isomaltotetraose molecule. The original horseshoe shape of the global structure reveals a high flexibility of the domain V and the identification of several sugar binding pocket proposed to locally increase dextran concentration around the active site for chain elongation. However, their weak binding affinity and low number suggest that anchoring of the nascent dextran chains to DSR-MΔ2 domain V is less important than in HMM dextran synthesizing enzymes such as DSR-S vardel Δ4N, explaining why DSR-MΔ2 is distributive and incapable of producing HMM dextran. Altogether, these findings provide new

mutation targets for GH70 enzyme engineering aiming at developing new sustainable processes for tailor-made dextran production.

V. Material and methods

Residue numbering refers to the whole size DSR-M enzyme (GenBank accession number: BN964_01347).

V.1. Construction of DSR-MΔ1, DSR-MΔ2 and DSR-MΔV deletion mutants

dsrm Δ1, dsrm ΔV-Nter and *dsrm ΔV* genes were amplified by PCR from pET-55/DsrM plasmid DNA template (Passerini et al., 2015) using the primers described in (Table S1). The addition of CACC sequence (underlined) to the 5'-forward primers allowed the correct insertion of genes into the pENTR/D-TOPO® vector (Life Technologies, Carlsbad, CA, USA). From a positive entry clone, LR recombination (Gateway® LR Clonase® II enzyme mix, Life technologies) was performed with pET-55-DEST destination vector (Novagen). Expression clones were selected on LB agar plates supplemented with 100 µg/mL of ampicillin. Plasmids were extracted with Sigma-Aldrich GenElute HP Plasmid Miniprep kit, verified by restriction analyses, and the genes of interest were sequenced (GATC Biotech, Constance, Germany). *E. coli* TOP10 competent cells (Life Technologies) were used for all cloning experiments. Concerning the construction of DSR-MΔ2 (form corresponding to the crystalized protein), the deletion of the first 151 amino acids (including the strep tag) from DSR-MΔ1 was performed following the method described by Wang and Malcolm, 1999 (primers in Table S1). Briefly, this method consists of the use of a two-stage procedure, based on the QuikChange Site-Directed Mutagenesis protocol (Stratagene, La Jolla, CA, USA), in which a pre-PCR, single-primer extension stage before the standard protocol, allows the deletion of a sequence of interest.

V.2. Protein expression and purification

Transformed *E. coli* BL21 star DE3 cells were grown in modified ZYM5052 (Studier, 2005) (with the following changes: 0.1% lactose, 0% glucose and 1% glycerol), supplemented with ampicillin (100 µg/mL) at 21 °C under agitation (150 rpm). After 26-hours incubation, cells were harvested by centrifugation, resuspended in lysis buffer (20mM phosphate sodium buffer, pH 7.4, 500 mM NaCl, 20 mM imidazole) supplemented with EDTA free anti-protease tablets (Roche, Basel, Switzerland) and disrupted by sonication. After centrifugation, recombinant enzymes were recovered in the soluble fraction of the crude cell extract, ready for purification.

Protein purification was performed using the ÄKTAxpress (GE Healthcare, Little Chalfont, UK) at 8 °C, with a first step consisting in 6xHis tag affinity chromatography (for details, see (Vuillemin et al.,

2016)) followed by a size exclusion step on superose 12 column 16x60 (GE Healthcare) from which the protein preparation is eluted with crystallization buffer (30mM MES pH 6.5, 100mM NaCl, 0.05g/L $CaCl_2$), or with 50 mM sodium acetate buffer, pH 5.75, 100mM NaCl, 0.05g/L $CaCl_2$ for biochemical characterization. Protein purity was verified by SDS-Page gel electrophoresis. Protein concentration was estimated by spectroscopy at 280nm, using a NanoDrop instrument (Wilmington, DE, USA). Theoretical molar extinction coefficients and molecular weights were calculated using the ExPASy ProtParam tool (http://web.expasy.org/protparam).

V.3. Activity assays

Activity was assayed using the 3,5-dinitrosalicylic method (Miller, 1959). One unit of DSR-M variant is defined as the amount of enzyme that catalyzes at 30°C the production of 1 µmol of fructose per minute, from 292 mM sucrose, in 50 mM sodium acetate buffer pH 5.75.

V.4. Enzymatic reaction and product characterization

Enzymatic reactions were performed at 30 °C on 292 mM sucrose in 50 mM sodium acetate buffer at pH 5.75, using 1 U/mL enzyme (corresponding to 0.014 mg/mL for DSR-MΔ2), stopped by 5 min incubation at 95 °C, and samples were stored at -20 °C until further analyses. For kinetic studies, samples were taken at regular intervals until total sucrose depletion. The reaction mixtures were then analyzed by HPAEC-PAD and HPSEC.

HPAEC-PAD (high performance anion exchange chromatography with pulsed ampero-metric detection) analyses were performed using a CarboPac TM PA100 analytical column (2 mm x 250 mm) coupled with a CarboPacTM PA100 guard (2 mm x 50 mm). Glucose, fructose, leucrose (5-0-α-D-glucopyranosyl-D-fructose) and sucrose were separated using a sodium acetate gradient (6–500 mM) in 150 mM NaOH over 36 min (0.250 mL/min), and quantification was performed using standards of these sugars at 5, 10, 15 and 20 mg/kg. Calculation of the percentages of glucosyl moieties derived from sucrose and incorporated into free glucose ($\%G_{glucose}$) or leucrose ($\%G_{leucrose}$) was done using molar concentrations at initial and final time of the reaction (t0 and tf): $\%G_{glucose} = [glucose_{tf}] / [sucrose_{t0}]$ and $\%G_{leucrose} = [leucrose_{tf}] / [sucrose_{t0}]$. Longer sodium acetate gradients varying from 50 to 136 min were used for oligosaccharide content analyses.

HPSEC analyses were performed using Shodex OH-Pak SB-804 and SB-802.5 columns (Showa Denko, Minato-ki, Tokyo, Japan) in series, coupled with a Shodex OH-Pak SB-G guard column and placed in a 30 °C oven. Samples were diluted in water to a maximum of 10 g/kg of total sugars. Elution with water was performed at a flow rate of 0.3 mL/min. After total sucrose consumption, percentage of glucosyl residues from sucrose incorporated into polymer ($\%G_{dextran} = (342/162) \times Area_{dextran\ tf}$ /

$Area_{sucrose\ t0}$) was calculated. The weight-average molar masses at peak apex of synthesized dextrans were determined using a calibration curve with standards of fructose, sucrose, and dextrans of 1.5, 11.3, 39.1 and 68.4 kg/mol (Sigma-Aldrich), at 10 g/kg each.

V.5. Crystallization and Data collection

Freshly purified enzyme was concentrated using centrifugal filter device (Amicon Ultra, 4 Ultracel, 50 kDa; Millipore). All crystallization experiments were carried out at 12 °C by the sitting-drop vapour-diffusion method, using MRC 96-well microplates (Molecular Dimensions, Newmarket, England) and a Nanodrop ExtY crystallization instrument (Innovadyne Technologies, Santa Rosa, USA) to prepare 400 nl droplets. Initial screening has been performed using JSCG+, PACT and PEGS I&II screens (Qiagen Hilden, Germany) from which we identified one condition in the presence of potassium thiocyanate. After condition optimization, the best DSR-MΔ2 crystals were obtained within a week with a 1:1 (v:v) ratio of protein to precipitant solution (PEG 3350 20%, 0.1 M 1,3-bis[tris(hydroxymethyl)methylamino]propane pH 6.5, 0.15 M KSCN, protein concentration: 20 – 22 mg/mL, cryoprotection solution: mother liquor + 10% (w/v) glycerol, for the apo crystals; PEG 3350 18%, 0.1 M 1,3-bis[tris(hydroxymethyl)methylamino]propane pH 6.5, 0.225 M KSCN 0.01 M Praseodyne Acetate (additive), protein preparation: 4 mg/mL + 10 mM I6, cryoprotection solution: mother liquor + 20% (w/v) ethylene glycol + 50 mM I6 (10 min), for the inactive IMOS-complex; PEG 3350 20%, 0.2 M sodium citrate tribasic dihydrate 0.1 M Bis-Tris propane 6.5 protein preparation: 6,00 mg/mL + 10 mM sucrose + 5 mM $SrCl_2$, cryoprotection solution: Cryo 15% ethylene glycol + 10mM Sacc for the inactive sucrose-complex).

Diffraction data were collected at beamlines ID23-2 and ID30-b at the European Synchrotron Radiation Facility (Grenoble, France). The diffraction data were collected at 100K. The data sets were integrated using XDS (Kabsch, 2010) and scaled using SCALA from the CCP4 suite (Winn et al., 2011). The same set of Rfree reflections has been generated on the first 3.2 Å dataset and transferred to the other datasets.

V.6. Structure determination

The DSR-MΔ2 structure was initially solved at 3.2 Å resolution by molecular replacement with PHASER (McCoy et al., 2007) using the structure of GTF-SI (Protein Data Bank entry 3AIE) as the search model. Manual rebuilding cycles using COOT (Emsley et al., 2010) alternated with restrained refinement using REFMAC5 (Murshudov et al., 1997) resulted in an incomplete model, lacking an important part (108 residues) of the N-terminal GBD. Residual electron density at the bottom of the catalytic pocket could be attributed to a glycerol molecule coming from the cryobuffer. This model has been deposited as 5LFC at the PDB and was used to solve the structure of the sucrose complex

(inactive mutant, at 3.6 Å) and of the later obtained I4-complex at 3.7 Å resolution. For this last dataset, as the N-terminal GBD was entirely visible in the chain A (see Figure S1), its structure has been reconstructed by placing small GBD fragments (obtained by homology modelling) followed by manual rebuilding cycles and refinement. Although the enzyme was co-crystallized and soaked with I6, only four residues were clearly visible in the electron density. The resulting structure will thus be referred as I4-complex. The tetrasaccharide was reconstructed by first superimposing the structure of ΔN123-GBD-CD2 in complex with isomaltotriose (4TTU). Praseodymium ions have been modelled as single atoms as the electron density does not allow the modelling of the metal coordination sphere. Data collection and refinement statistics are presented in Table S2.

The final models have been validated using WHATIF and MOLPROBITY webservers (http://swift.cmbi.ru.nl/whatif; http://molprobity.biochem.duke.edu/). Coordinates and structure factors have been deposited at the PDB (entries 5LFC, 5NGY and 5O8L for DSR-MΔ2, DSR-MΔ2-E715Q: I4-complex and DSR-MΔ2-E715Q: sucrose-complex structures, respectively).

V.7. SAXS measurements and processing

Small-angle X-ray scattering (SAXS) experiments were performed on the SWING beamline(Bizien et al., 2016) at the SOLEIL synchrotron, Gif-sur-Yvette, France. The wavelength was set to 1.033 Å. A 17x17 cm^2 Aviex CCD detector was positioned 1800 mm from the sample, with the direct beam off-centered. The resulting exploitable q-range was 0.006-0.55 $Å^{-1}$, where $q = 4\pi\sin\theta/\lambda$, considering 2θ as the scattering angle. The samples were circulated in a thermostated quartz capillary with a diameter of 1.5 mm and 10 µm wall thickness positioned inside a vacuum chamber. 50 µl volume of sample was injected onto a size-exclusion column (Bio SEC3 300, Agilent, Santa Clara, CA, USA) equilibrated in MES buffer (30 mM MES pH 6.5, 100 mM NaCl, 0.05 g/L $CaCl_2$) using an Agilent high-performance liquid-chromatography (HPLC) system and eluted directly into the SAXS capillary cell at a flow rate of 200 µL/min at a temperature of 10 °C. SAXS data were collected online throughout the elution time and a total of 149 frames, each lasting 2 s, were recorded separated by a dead time of 0.5 s between frames. The transmitted intensity was continuously measured with an accuracy of 0.1% using a diode embedded in the beam stop. For each sample, the stability of the associated radius of gyration and the global curve shape in the frames corresponding to the main elution peak were checked, and the resulting selection of curves were averaged as described previously (David and Pérez, 2009). The recorded curves were normalized to the transmitted intensity and subsequently averaged using FOXTROT. The same protocol was applied to buffer scattering. Rg values were determined by a Guinier fit of the one-dimensional curves using the ATSAS package (Petoukhov et al., 2012). The p(r) function was calculated using the GNOM program and the corresponding *ab initio* envelopes were

calculated using the GASBOR program. The solution model including the N-term and C-term residues has been calculated and fitted to the experimental data using the AllosMod-FOXS program (https://modbase.compbio.ucsf.edu/allosmod-foxs/).

V.8. Mutagenesis studies

Mutants E715Q (inactive mutant), Y180A, Y264A, L575W, D813A and L816A were constructed by inverse PCR (oligo-mediated introduction of site-specific mutations) using *dsrm-Δ2* plasmid as template and the primers described in Table S1. PCR amplification was carried out with Phusion DNA polymerase (0.5 U) for 16 cycles (95 °C, 15 s; 55 or 60°C, 20 s; 72 °C, 8 min). The parental plasmid template was digested with DpnI and PCR products were purified using a Qiaquick spin column, following the manufacturer recommendations. *E. coli* TOP10 cells were transformed with the plasmid. Resulting clones were selected on LB agar plates supplemented with 100 µg/ml of ampicillin. Plasmids were extracted with Sigma-Aldrich GenElute HP Plasmid Miniprep kit, verified by restriction analyses, and the genes of interest sequenced (GATC Biotech). DpnI restriction enzyme and Phusion® High-Fidelity DNA Polymerase were purchased from New England Biolabs (Beverly, MA, USA). Oligonucleotides were synthetized by Eurogenetec (Liège, Belgium). *E. coli* BL21 star DE3 cells (Life Technologies) were used as hosts for mutant productions. The reaction products from 292 mM sucrose were analyzed by HPSEC and HPAEC-PAD and compared with those obtained with the wild-type DSR-MΔ2 enzyme.

VI. Acknowledgements

This work was in part supported by French Ministry of Higher Education and Research, and the Agence Nationale de la Recherche (ANR). We gratefully thank Pauline Bondy for enzyme construction and recombinant expression help, Florent Grimaud and Etienne Severac for technical support and Thibault Craviari for crystallization trial assistance. We are also very grateful for ITC analyses realized by Annabelle Varrot and Emilie Gillon from structural and molecular glycobiology group, CERMAV-CNRS (Grenoble, France). We greatly acknowledge staff from the European Synchrotron Radiation Facility (ESRF, Grenoble, France), as well as the SWING team of synchrotron Soleil (Gif-Sur-Yvette, France) for SAXS acquisition. We thank the ICEO facility (PICT, Toulouse, France) for providing access to HPLC equipment and protein purification systems. Samuel Tranier and the structural biophysics group of the Integrated Screening Platform of Toulouse (PICT, Toulouse, France) are greatly acknowledged for access to crystallization facility and for their help on synchrotron data collection.

VII. Supplementary information

Table S1: **Primers used to generate DSR-M variants** (deletion or site-directed mutagenesis). Underlined bases were necessary for the directional insertion of the purified PCR products in pENTR/D-TOPO® vector (Life Technologies)

Primer Name	Nucleotide Sequence	Main features
DSR-MD1for	5'-CACCCAAACGCCGGTTGGTACAACACAG-3'	Amplification of dsrMΔ1 gene
DSR-MD1rev	5'-TTTTGCCATCGTACCATCGTTATT-3'	
DSR-MDVfor	5'-CACCCAAAACAGTGTTGCTAC-3'	Amplification of dsrMΔV gene
DSR-MDVrev	5'-TTGATTCATAAGCTGCTTGGGC-3'	
DSR-MDVfor	5'-CACCCAAAACAGTGTTGCTAC-3'	Amplification of dsrMΔV-Nter gene
DSR-MD1rev	5'-TTTTGCCATCGTACCATCGTTATT-3'	
DSR-MD2for	5'-GGAGATATACCATGGCAAGCTTTGAAAAAGCCCCTGATTC-3'	Amplification of dsrMΔ2 plasmid
DSR-MD2rev	5'-GAATCAGGGGCTTTTTCAAAGCTTGCCATGGTATATCTCC-3'	
L575Wfor	5'-CTAACCATTGGCAAAATGGGGCG-3'	Mutation of position 575
L575Wrev	5'-CATTTTGCCAATGGTTAGTAGAGCC-3'	
E715Qfor	5'-CACATTTTACAAGATTGGTCTCC-3'	Mutation of position 715 - inactive mutant
E715Qrev	5'-CCAATCTTGTAAAATGTGAATGTG-3'	
D813Afor	5'-CCCACTGTTGCTAAGGCTTTAC-3'	Mutation of position 813
D813Arev	5'-GCCAGTAAAGCCTTAGCAACAG-3'	
L816Afor	5'-GATAAGGCTGCACTGGCTAAG-3'	Mutation of position 816
L816Arev	5'-CTATCCTTAGCCAGTGCAGC-3'	
Y180Afor	5'-GGGCGTGCTAGTTTGAAAGATGG-3'	Mutation of position 180
Y180Arev	5'-CAAACTAGCACGCCCACCAG-3'	
Y264Afor	5'-GTAAAAGCAGGCGCTGTTACTAAC-3'	Mutation of position 264
Y264Arev	5'-CTGACTGTTAGTAACAGCGCCTG-3'	

Table S2: **Data collection and refinement statistics**

	DSR-MΔ2	DSR-MΔ2 E715Q I4 Complex	DSR-MΔ2 E715Q Sucrose Complex
Space group	$P\ 2_12_12_1$	$P\ 2_12_12_1$	$P3_121$
Unit-cell parameters (Å)	a=105.8 b=128.7 c=229.3 α=β=γ=90°	a=105.9 b=128.8 c=234.5 α=β=γ=90°	a=b=183.6 c= 146.8 γ=120°
No. of molecule/AU	2	2	1
Matthews coefficient ($Å^3$/Da)	2.70	2.75	2.38
Solvent content (%)	54.32	55.4	48.3
Wavelength (Å)	0.8726	0.8729	0.9762
Resolution range (Å)	49.23-3.20 (3.37-3.20)*	50.00-3.70 (3.90-3.70)*	50.00-3.60 (3.79-3.60)*
No. of observed reflections	270317 (31461)*	146438 (19782)*	127013 (19196)*

No. of unique reflections	51658 (7408)*	34388 (4910)*	32876 (4780)*
Completeness (%)	98.7 (98.4)*	98.3 (97.3)*	98.4 (99.0)*
Multiplicity	5.2 (4.2)*	4.3 (4.0)*	4.0 (4.0)*
I/σ(I)	4.9 (1.7)*	3.6 (1.4)*	4.9 (1.1)*
R_{merge}	0.150 (0.453)*	0.195 (0.528)*	0.152 (0.724)*
CC ½	0.99 (0.87)*	0.98 (0.82)*	0.99 (0.70)*
Wilson B ($Å^2$)	55.8	62.3	72.0
Refinement			
Rwork/Rfree	0.221/0.248	0.192/0.236	0.198/0.228
RMSD Bond lengths (Å)	0.009	0.007	0.008
RMSD Bond angles (°)	1.207	1.098	1.26
Rama Favored (%)	95.0	95.0	95.0
Rama Allowed (%)	5.0	5.0	5.0
Mean B; all atoms ($Å^2$)	54.0	63.0	90.2

* refer to statistics for the outer resolution shell

Table S3: **Overview of DSR-M variants constructed in this study** (deletion or site-directed mutagenesis) and impact on the mean molar mass of produced dextrans.

Enzyme form	Characteristic	MM*
DSR-M	full enzyme	28
DSR-MΔ1	deletion of signal peptide and C-term region including pocket V-D	23
DSR-MΔ2	deletion of disordered region at N-term of DSR-MD1	23
DSR-MΔ2 Y180A	Tyr mutation in pocket V-A	12.4
DSR-MΔ2 Y264A	Tyr mutation in pocket V-B	17.7
DSR-MΔ2 YY-180-264-AA	Tyr mutation in pockets V-A and V-B	11.3
DSR-MΔ2 L575W	Leu mutation in loop B1	12.5
DSR-MΔ2 E715Q	Inactive mutant	-
DSR-MΔ2 D813A	Asp mutation in helix H1' (loop A1)	22.2
DSR-MΔ2 L816A	Leu mutation in helix H1' (loop A1)	23.8
DSR-MΔV-Nter	Deletion of the N-terminal domain V including pockets V-A and V-B	16
DSR-MΔV	Deletion of the entire domain V including all the pockets	16

*Determined produced dextran molar mass at peak apex in $kg.mol^{-1}$

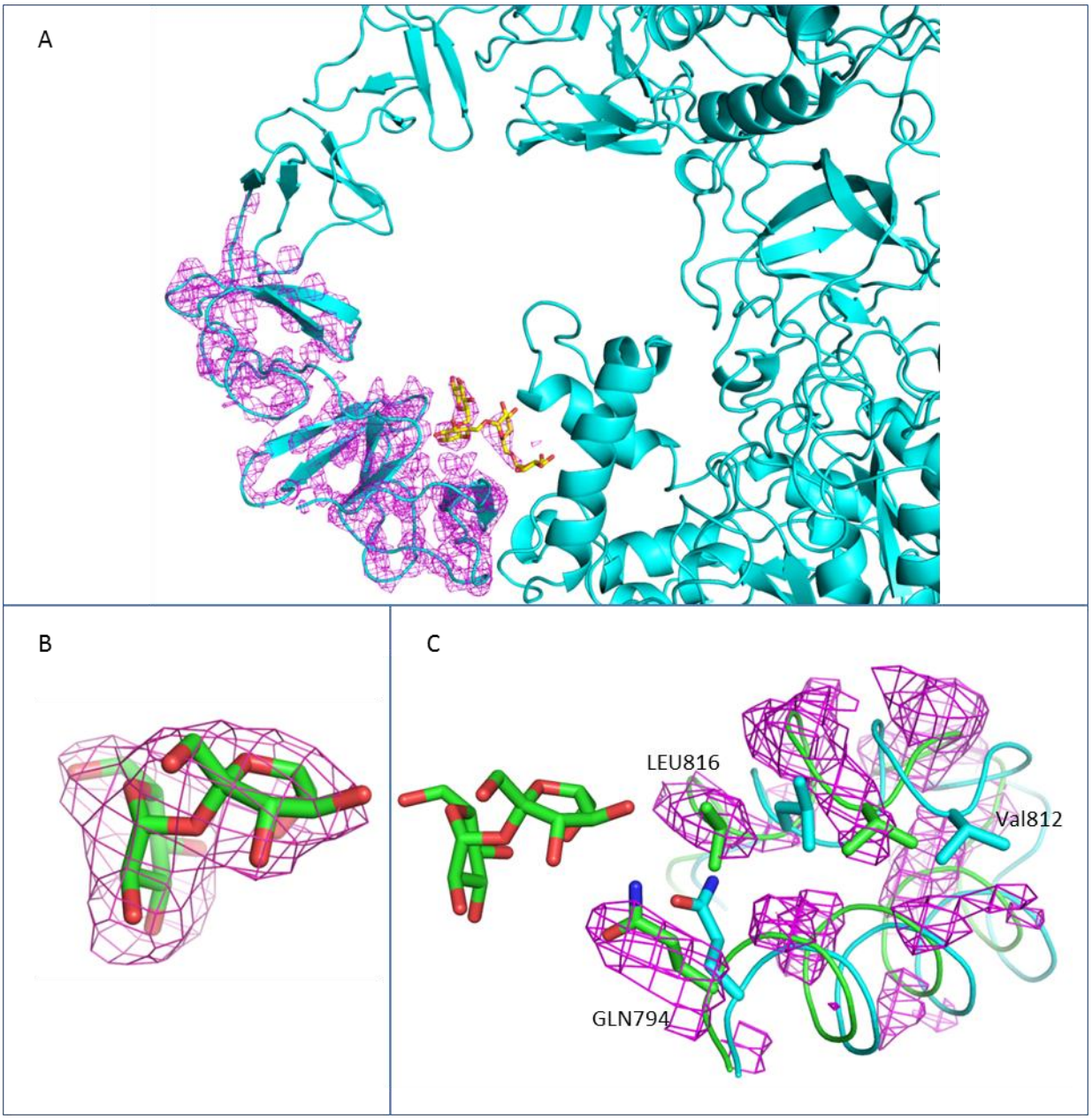


Figure S1: **Electron density maps**. **A**, Difference Fo-Fc map (level = 3σ), calculated after the molecular replacement, shows the presence of the N-terminal part of the GBD in the structure of the I4-complex (5NGY). **B**, Difference Fo-Fc map, contoured at 3σ, for the sucrose molecule bound in the active site (5O8L). **C**, Loop A1 mouvement is illustrated by fitting an open loop conformation (5NGY, cyan structure) against the sucrose-complex dataset that shows a closed loop conformation (5O8L, green structure is superimposed). Difference Fourier map Fo-Fc is contoured at 3σ.

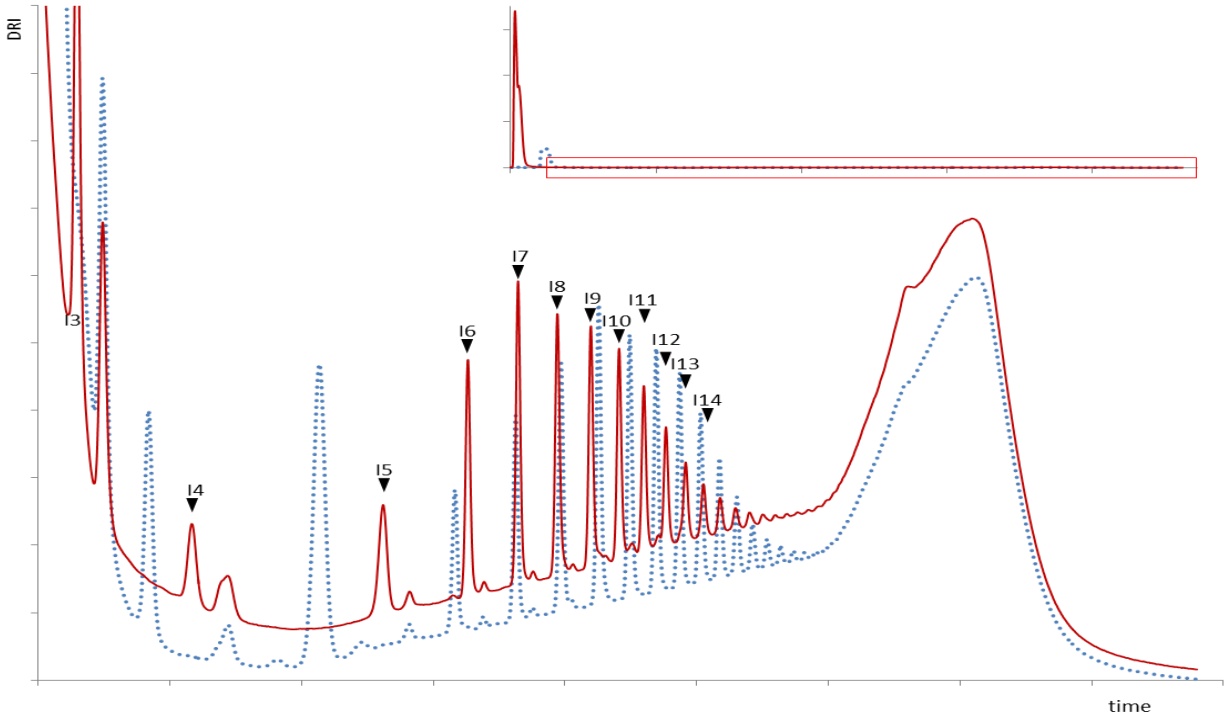


Figure S2: **Invertase digestion**. HPAEC-PAD chromatogram of products formed after 5 min of reaction on 292 mM sucrose by DSR-MΔ2, before (blue dash) and after (red) partial digestion with invertase.

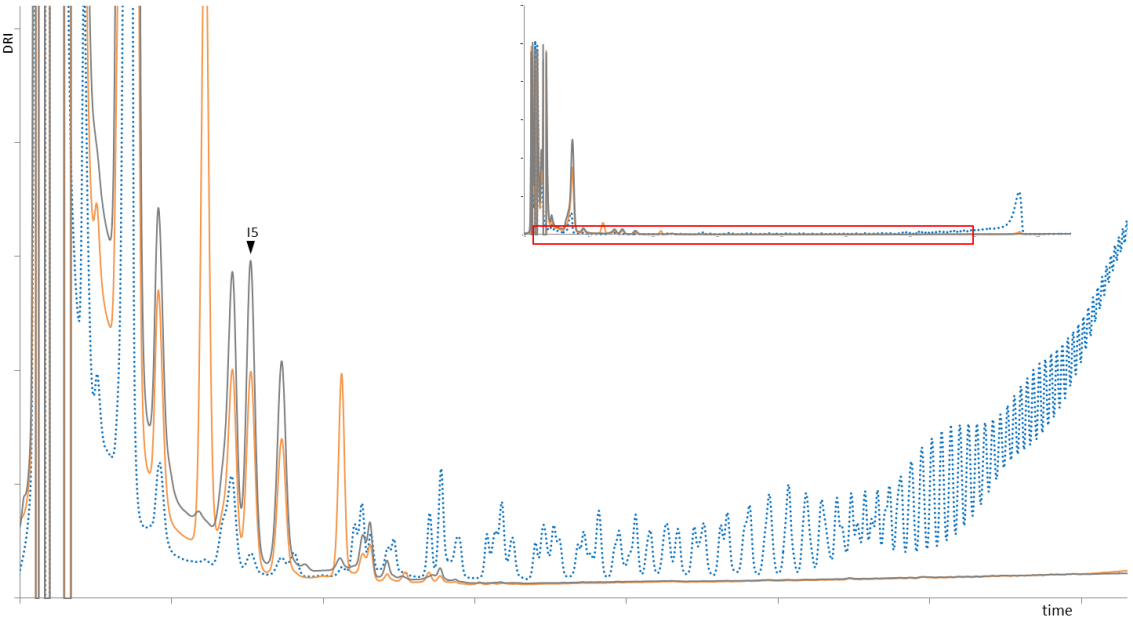


Figure S3: **Dextranase digestion**. HPAEC-PAD chromatogram of products formed after total sucrose depletion, before (blue dash) and after partial digestion with dextranase (orange) and with dextranase and invertase (grey).

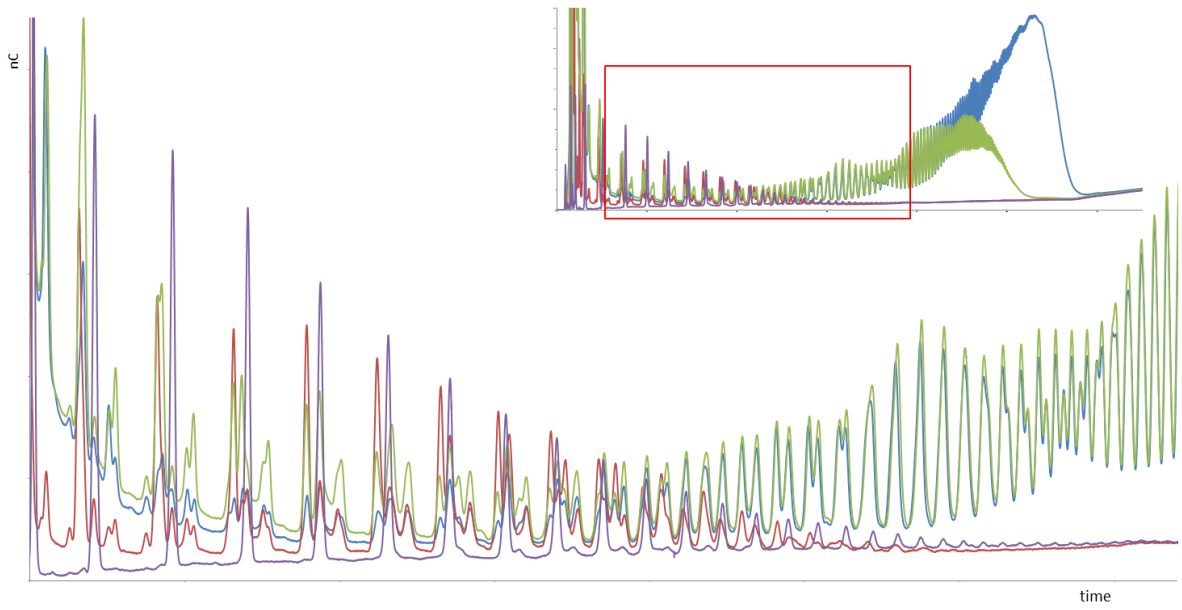


Figure S4: **Acceptor reaction of DSR-MΔ2 on leucrose and fructose**. HPAEC-PAD chromatogram of products formed after total 292 mM sucrose depletion in standard conditions, on sucrose alone (blue), on sucrose + 30 g/L leucrose (red), on sucrose + 30 g/L fructose (green); commercial dextran 1,5 kg/mol (purple).

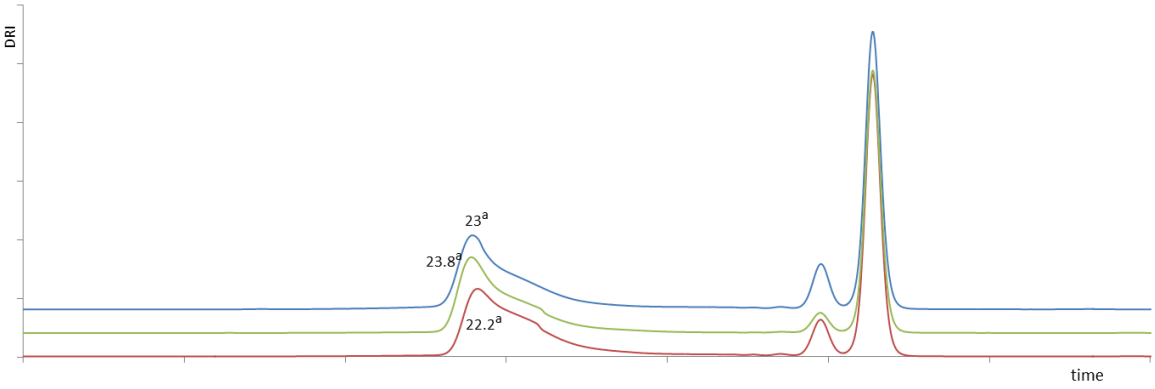


Figure S5: **HPSEC analysis of the products synthesized by DSR-MΔ2 variants from 292 mM sucrose.** Blue, DSR-MΔ2; green, DSR-MΔ2 L816A; red, DSR-MΔ2 D813A. [a] *determined molar mass at peak apex in kg.mol^{-1}*

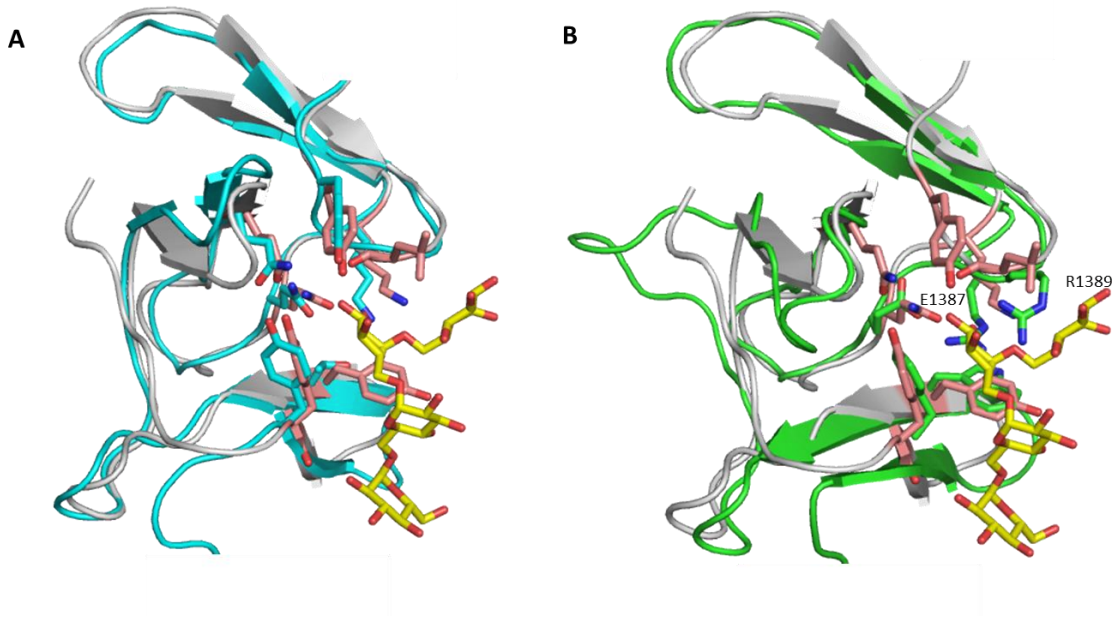


Figure S6: **Superposition of putative pockets V-B and V-C with DSR-MΔ2 pocket A showing topology similarities**. A, superposition of V-A and V-B in cyan (rmsd of 1.5 Å on Cα). B, superposition of V-A and V-C in green (rmsd of 2.63 Å on Cα).

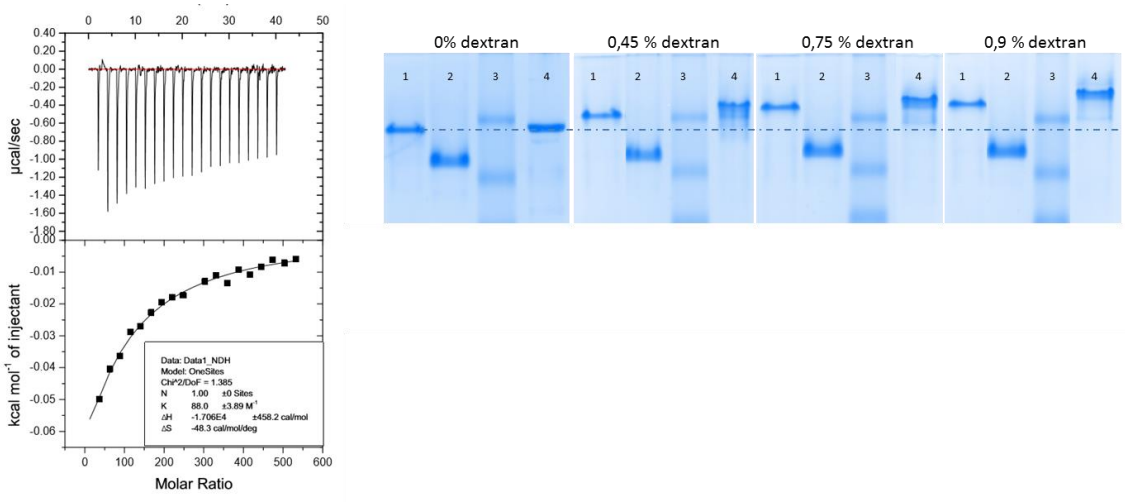


Figure S7: **Binding affinity of DSR-M variants towards glucans.** A, ITC measurements* with DSR-MΔ2 and IM6 (isomaltohexaose). B, Affinity gel electrophoresis of DSR-M variants in gels containing variable amounts of dextrans. Affinity gels for which the content in dextran is expressed in%(w/v). Lanes 1 correspond to DSR-MΔ2, lanes 2 correspond to DSR-MΔV, lanes 3 correspond to protein standards, lanes 4 correspond to ΔN-123-GBD-CD2 used as a positive control. (For details on the method see (Brison et al., 2016))

***ITC measurements.** Isothermal titration calorimetry (ITC) experiments were performed using an ITC200 calorimeter (Microcal-Malvern Instruments). The experiments were carried out at 25°C with the protein and ligands dissolved in 30 mM MES pH 6.5, 100mM NaCl and 0.05 gL^{-1} $CaCl_2$. The protein concentration was checked by measuring A_{280} by using a theoretical molar extinction coefficient

of 207 660 M^{-1} cm^{-1} and a molecular weight of 144 128 daltons. The protein solution was placed in a 200 µl sample cell and the titration was performed with 20 injections of 2 µl carbohydrate ligands whilst stirring at 1000 $rev.min^{-1}$. Due to the low affinity, experiments were done in large excess of ligand using 100 mM of I6 with protein at 39.3 µM. A blank titration using 100 mM I6 injected into buffer has been performed and subtracted from the raw titrations curves using the Origin 7 software. Two independent titrations were performed and the experimental data were fitted to a theoretical titration curve using supplied Origin 7 software, with DH (enthalpy change), Ka (association constant) as adjustable parameters and n (number of binding sites per protein monomer) fixed at 1. Free energy change (ΔG) and entropy contributions (TΔS) were derived from the equation $\Delta G = \Delta H - T\Delta S = -RT \ln Ka$ (with T the absolute temperature and R = 8.314 $Jmol^{-1}K^{-1}$).

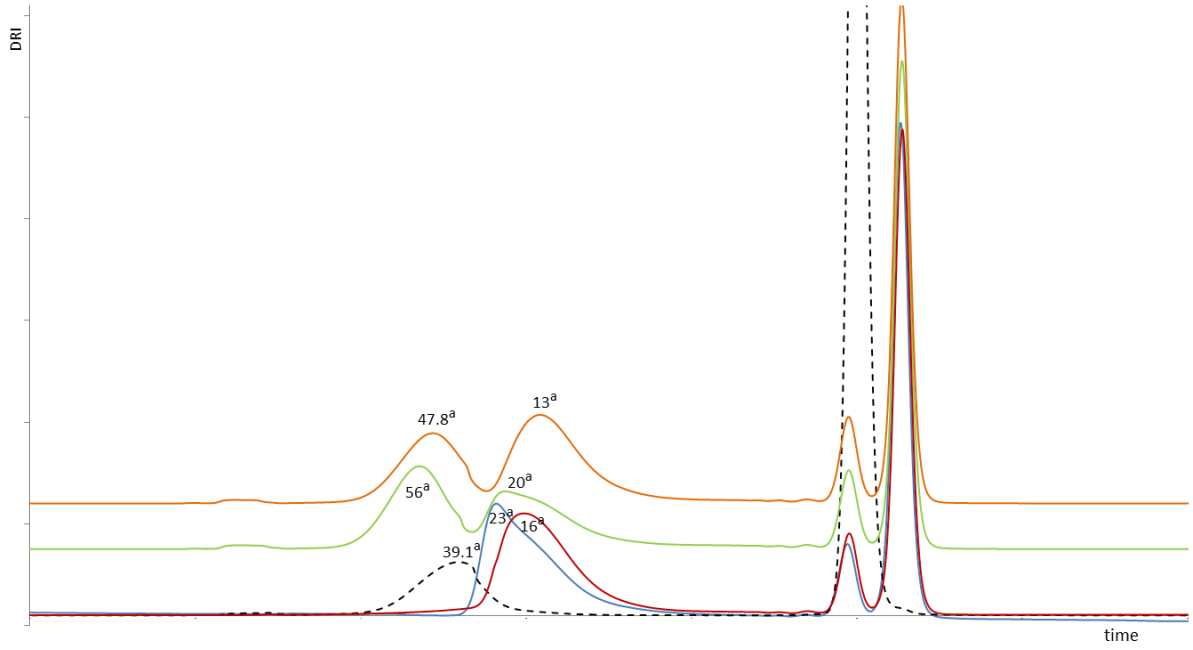


Figure S8: **HPSEC analysis of the products synthesized by DSR-M variants from 292 mM sucrose and 15g/L dextran 39,100.** Dash line, reaction on sucrose and dextran initial time; Blue, DSR-MΔ2 on sucrose only; red, DSR-MΔV on sucrose only; green, DSR-MΔ2 on sucrose and dextran; orange, DSR-MΔV on sucrose and dextran. [a] *determined molar mass at peak apex in kg.mol*$^{-1}$

Chapter III

Futile cycle engineering of the DSR-M dextransucrase modifies the resulting polymer length

Claverie Marion, Cioci Gianluca, Guionnet Matthieu, Schörghuber Julia, Lichtenecker Roman, Moulis Claire, Remaud-Simeon Magali and Lippens Guy

To be submitted for publication

I. Introduction

The glycoside hydrolase family 70 comprises α-transglucosylase enzymes, also called glucansucrases (GSs), which synthesize glucose polymer from sucrose, a cheap and renewable agro resource. Depending on the enzyme specificity, the α-glucans vary in terms of size, structure, types of glucosidic bonds (α-(1→2), α-(1→3), α-(1→4), or α-(1→6)) and degree of branching. This variety of structures confers to glucans different properties, making them useful to pharmaceutical, fine chemicals, material and food industries.

To date, structures of six truncated GH70 enzymes have been solved: GTF180-ΔN from *Lactobacillus reuteri* 180 (Vujicic-Zagar et al., 2010b), GTFA-ΔN from *L. reuteri* 121 (Pijning et al., 2012), GTF-SI from *Streptococcus mutans*(Ito et al., 2011), the ΔN123-GBD-CD2 of DSR-E from *L. mesenteroides* (Brison et al., 2012) and very recently Gtf-B-ΔNΔV from *L. reuteri* 121 (Vujicic-Zagar et al., 2010b; Pijning et al., 2012; Ito et al., 2011; Brison et al., 2012; Bai et al., 2017) and DSRM-Δ2 from *L. citreum* B-1299. All structures reveal a common 5 domain organization with a global U-shape folding. Domains A, B and C are common to GH13 α-amylases and domains IV and V are unique to GH70. Except for domain C, these domains are formed by two discontinuous polypeptides from both N- and C-terminal polypeptides. The active site comprises the catalytic triad of GH70 enzymes, which display a α-retaining mechanism in which sucrose cleavage occurs through the formation of a β-D-glucosyl covalent intermediate. An aspartic and glutamic acid act as the nucleophile and the acid-base catalyst, respectively, whereas a second aspartate functions as a transition state stabilizer. 3D structures have deepened our understanding of the GS mechanism, in particular their linkage specificities, thereby underscoring the potential of enzyme engineering to produce tailor-made α-glucans (Meng et al., 2016a). However, less is known concerning the factors that determine the final size of the products. Identifying such determinants could contribute to the synthesis of glucans with both controlled linkage composition and molar mass.

The inventory of the GH70 enzymes encoded by *Leuconostoc citreum* NRRL B-1299 led to the identification of three novel α-transglucosylases. One of them, DSR-M, presents the particularity to naturally synthesize a very linear and low molar mass dextran (of around 28 kg/mol) directly from 100 g/L sucrose (Passerini et al., 2015). Three structures of this enzyme (free enzyme, sucrose and IM4 complexes) were recently solved (Claverie et al., 2017). Structural and biophysical characterization showed that the low affinity of its glucan binding pockets in domain V toward the growing glucan chain is one factor explaining its inability to produce longer polymers. Nevertheless,

other structural features such as a long loop near the catalytic domain (loop B2) also distinguish DSR-M from other GSs. Here we present the structure of a new protein-ligand complex that displays a better define electron density in this area, allowing for the first time to characterize the binding of isomalto-oligosaccharides (IMOS), the natural product of the enzyme, near the catalytic domain. A tryptophan residue (Trp624) specific to DSR-M is found at the basis of loop B2 in direct interaction with a glucosyl unit of the bound glucan, and plays a major role in the polymerization capacity of DSR-M.

Mutation of this residue highlights the critical role of Trp624 on the polymer length. NMR spectroscopy shows that the Trp624 residue additionally serves to dynamically shape the anchoring region for the nascent dextran chains, and thereby equally governs water accessibility to the active site. The enhanced cleavage of the covalent β-D-glucosyl intermediate results in a larger pool of released glucose units that can efficiently be elongated, and thereby contributes to the small lengths of the final polymers. We integrate the structural and functional aspects into a theoretical model of the enzymatic process, whereby binding of a nascent chain does not necessarily lead to productive elongation (Bar-Even et al., 2015; Cantarel et al., 2009). Enhanced "futile cycling" increases the pool of effective seeds for the dextran chains, and thereby directly determines the length distribution of the final polymers. Our model with the length dependent probability (p) of a productive encounter between enzyme and substrate as its sole adjustable parameter confirms that changing the interaction energy for the nascent chains is an efficient strategy to modulate the resulting polymer length. The limited size of the products of DSR-M W624A makes it of great interest for food and pharmaceutical industrial applications.

II. Results and discussion

II.1. The crystal structure of DSR-MΔV in complex with an isomaltotetraose defines novel anchoring points

The active site of DSR-M comprises the catalytic triad of GH70 enzymes involved in the α-retaining mechanism. An aspartic and glutamic acid act as the nucleophile and the acid-base catalyst, respectively, whereas a second aspartate functions as a transition state stabilizer (Claverie et al., 2017; Leemhuis et al., 2013a).

We previously solved the crystal structure of an inactive mutant of DSR-MΔ2 (a truncated form of DSR-M comprising a large part of domain V), in which the acid-base catalyst was mutated to a glutamine (DSR-MΔ2-E715Q) (Claverie et al., 2017). Here, the same DSR-MΔ2-E715Q mutant was co-crystallized in the presence of pure isomaltohexaose (I6), and crystals were further soaked with high

amount of this compound. The crystal structure was solved at 3.9 Å (Data statistics are summarized in Table S4). Nevertheless, the electron density maps unambiguously showed the presence of a substrate in the enzyme active site that could be attributed to four of the six glucosyl units of the I6 molecule (Figure S9). As the electron density is contiguous, the I4 molecule can be seen as a reaction product immediately after elongation. This allowed mapping of the DSR-M subsites -1, +1, +2 and +3, and the glucosyl units found in these subsites are referred to accordingly. Glc+1 and Glc-1 are stabilized by a network of hydrogen bonds with selected side chain atoms: the O2 atom of Glc+1 interacts with Tyr626 and its O3 with Asn681, whereas for Glc-1, its O2 atom interacts with Asp790, Gln715 (replacing Glu715 in the active enzyme) and His789, its O3 with His789, its O4 with Asp1118, and finally its O5 and O6 with Asp677. Also, Tyr1127 is involved in a hydrophobic stacking interaction with Glc-1. Similarly to the +2 glucosyl unit of the maltose in the complex with GTF180-ΔN (pdb code: 3KLL), Glc+2 is in a stacking interaction with Trp717 (the equivalent of Trp1065 in GTF180-ΔN from Lactobacillus reuteri 1808) and its O3 atom interacts with Ser791. These residues are all conserved in the other GS structures (Leemhuis et al., 2013a). In contrast, Trp624, located at the foot of the long loop B2 is a distinct feature of DSR-M when compared to other dextransucrases (Claverie et al., 2017), and is in our structure observed in a stacking interaction with Glc+3 (Figure 37). Sugar-aromatic stacking interactions are a recurring theme in carbohydrate-protein complexes, and contribute roughly 6 kJ/mol to the binding energy (Asensio et al., 2013). According to the crystal structure, the chain length should at least attain three residues before this stacking interaction can contribute to the affinity of the incoming chain for the catalytic site, justifying our adoption of 3 units as the limit between “short” and “long” chains in the theoretical model (vide infra).

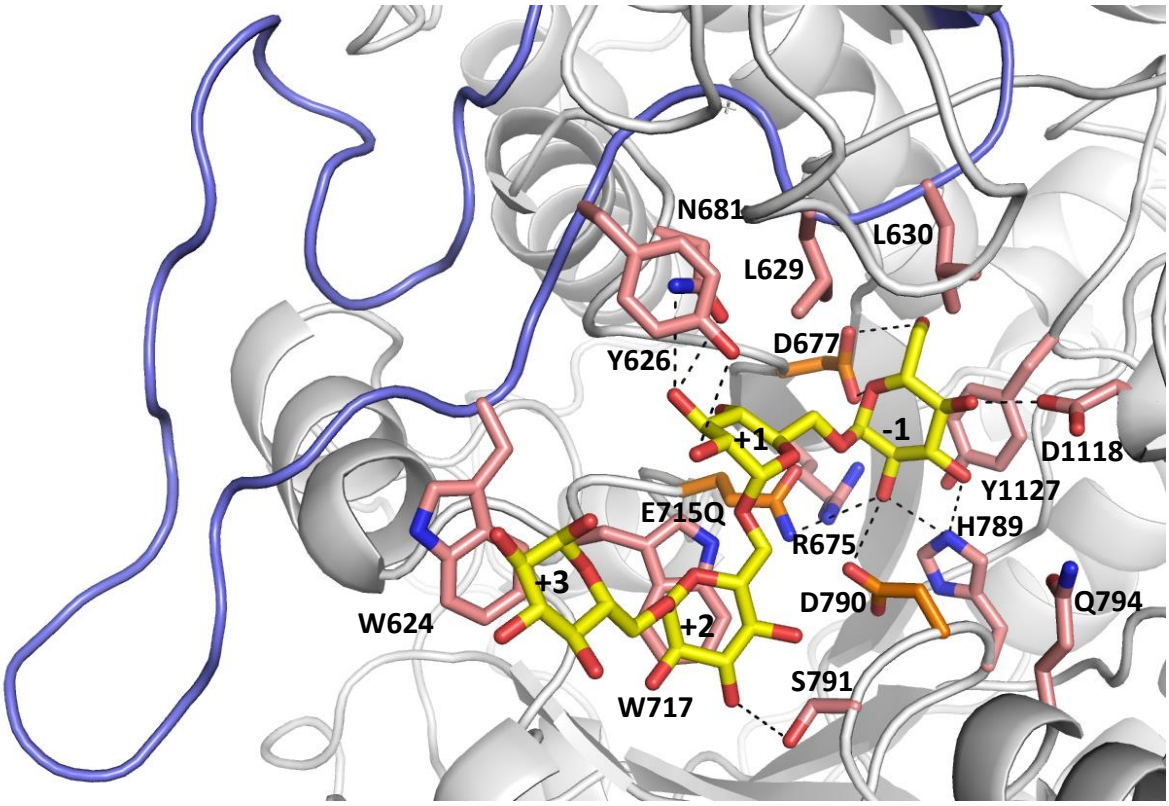


Figure 37: **Crystal structure of DSR-MΔ2-E715Q with an isomaltotetraose**. Indole rings form a platform for the nascent dextran chain with the two glucosyl units Glc+3 and Glc+2 stacking against the aromatic rings of Trp624 and Trp717, respectively. Catalytic residues are shown in orange sticks; residues participating in the catalytic site are shown in salmon sticks; Loop B2 (600-632) is presented in purple, helix H1' (807-817) has been removed for clarity.

II.2. The W624A mutation changes the reaction rate and final length distribution.

To better isolate the effect of this anchoring point on both the reaction rate and final length distribution of the resulting dextrans, further experiments were performed with the DSR-MΔV construct (called ΔV in what follows), which has similar catalytic activity as the Δ2 construct (Claverie et al., 2017) but lacks the sugar binding pockets of domain V susceptible to compete for binding to the isomalto-oligosaccharides. With the ΔV construct, we expect the active site to act as the major sugar binding area. We mutated Trp717 (equivalent to Trp1065 of GTF180-ΔN) or Trp624 (unique to DSR-M) into Ala, and analyzed the enzymatic reactions under identical conditions as the wt ΔV. The W717A mutation severely impacted the enzymatic activity (loss >99% compared to wt ΔV) and abolishes the ability of the mutant to synthesize polymers. This is in agreement with the severe loss of transglycosylase activity observed for all the mutants of Trp1065 in GTF180-ΔN (Meng et al., 2017), except for mutant W1065F. The replacement of Trp with Phe enabling to maintain a stacking interaction with Glc+2 obviously critical for chain elongation (Meng et al., 2017). The ΔV W624A

mutant is still capable of dextran synthesis, but its catalytic rate as measured by the fructose release is twice as slow as that of the wt enzyme. Obtaining an equivalent rate at the start of the reaction (Figure 38) indeed required a 2.3 fold larger amount of mutant enzyme compared to the wt. However, even with this larger amount of mutant enzyme, when the rate acceleration sets in after 15 minutes, the fructose release catalyzed by the wt enzyme gains another factor of 1.8, making the wt fourfold more efficient than the W624A mutant (Figure 38). When we analyze the resulting chains after full consumption of sucrose, dextrans produced by the W624A mutant are sixfold shorter (with a mean MM of 2.4 kg/mol) than those produced by the ΔV enzyme (with a previously determined average MM of 16 kg/mol (Claverie et al., 2017)) (Figure 38, bottom). These results hence indicate that the interaction in subsite +3 is critical for both the enzymatic rate constants and for the final product distribution.

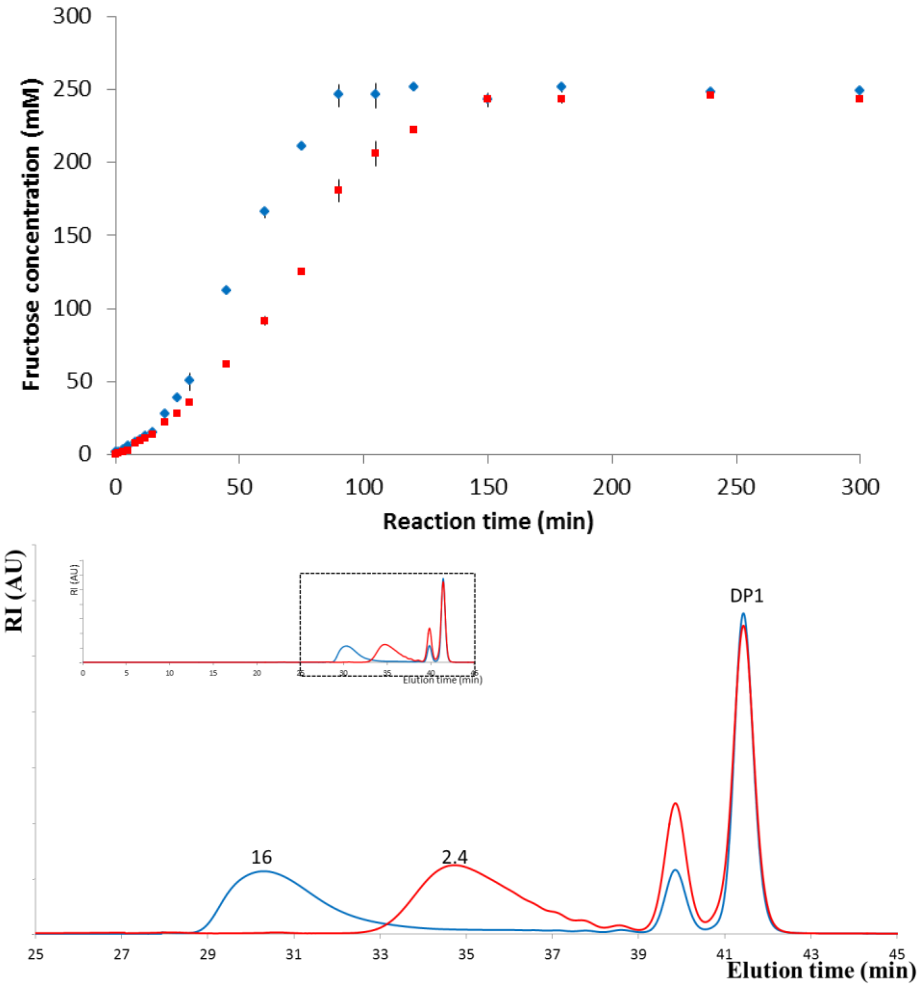


Figure 38: Effect of the W624A mutation on the enzymatic reaction. (top) Reaction progress on 100 $g.l^{-1}$ of sucrose for the wt DSR-MΔV (blue diamonds) or DSR-MΔV W624A mutant (red squares) with a concentration of the mutant 2.3 times higher than the wt. (bottom) Final distribution of the resulting chains shows that the mutant (red) makes smaller chains than the wt enzyme (blue). Annotation: molar mass at peak apex in $kg.mol^{-1}$.

II.3. Beyond a simple stacking platform: the W624A mutation equally influences the dynamics of the catalytic site.

When analyzing in detail the oligosaccharides at the time point of the reaction acceleration as measured by the fructose release (15 minutes, Figure 38), we observed a distribution already skewed to longer oligosaccharides for the wt enzyme than that generated by the W624A mutant (Figure 39, top). Whereas all chains generated by wt DSR-MΔV start with a sucrose disaccharide, a non-negligible fraction of nascent polymers produced by the W624A mutant corresponds to pure isomalto-oligosaccharides (starting with a glucose moiety at their reducing end) (Figure 39 top and Figure S10 and Figure S11). This behavior of ΔV suggests that the DSR-M active site accommodates only sucrose in the early stage of the reaction while protecting the glucosyl-enzyme from water interception, in agreement with our previous observations on DSR-MΔ2 (Claverie et al., 2017). In contrast, the cleavage of the β-D-glucosyl covalent intermediate by a water molecule seems more prominent in the W624A mutant, thereby increasing the pool of glucose as initial acceptors.

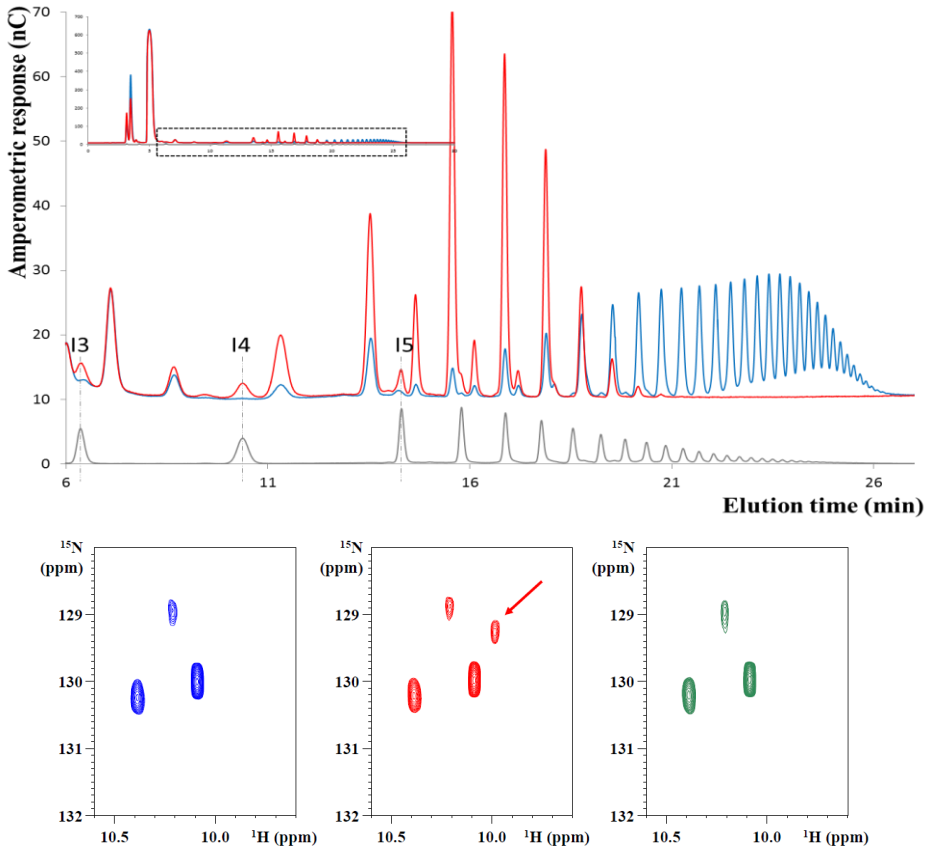


Figure 39: **Dynamic changes in the active site following the W624A mutant**. (top) HPAEC-PAD profiles of reaction medium after 15 min reaction using 292 mM sucrose. Blue: DSR-M ΔV, Red: DSR-M ΔV-W624A, grey: commercial dextran 1500. (bottom) Tryptophan indole region of the ^{1}H, ^{15}N TROSY spectrum of DSR-M ΔV (blue, left), DSR-M ΔV-W624A (red, middle) and DSR-M ΔV-W624A/W717A (green, right). The peak labeled with a red arrow is assigned to the indole ring of W717.

In order to gain further insight in this additional factor of the catalytic process, we attempted structural characterization of the mutant enzyme. In the absence of a crystal structure of the mutant, we produced a uniformly ^{15}N labeled sample of ΔV. Structural integrity of the protein could be assessed by characteristic methyl resonances found at -0.9 and -0.2ppm (Figure NMR1-Figure NMR3 - Figure NMR5, p146). We then recorded a TROSY spectrum at 800 MHz of the different proteins. Despite the large size of the enzyme (104 kDa for the ΔV isoform we used), we did observe some signals with amide proton chemical shifts in the 7.5-8.5 ppm range, probably corresponding to unstructured loops in the enzyme (Figure NMR2, p147). We moreover detected at least three signals with proton chemical shifts around 10.4-10.5 ppm, corresponding to indole Hε protons of Trp. Biosynthesis of the same protein with an indole specific labeling scheme (Schörghuber et al., 2017) confirmed that these signals belong to Trp residues. Their number is far inferior to the number of Trp residues in ΔV (15 Trp including the strep tag one, see Supplemental Information), but line broadening for immobile residues is expected for a protein of this size without deuteration (Pervushin et al., 1997; Xu and Matthews, 2013). To see whether one of these signals could correspond to the Trp624, we produced the ΔV W624A mutant with the same labeling scheme. Amazingly, at the level of the indole rings, no signal disappeared but a novel signal at 9.98/129.2 ppm appeared (Figure 39). One possible candidate is Trp717, at the base of the B2 loop and forming a π-π network (Chourasia et al., 2011) with Trp624 in the crystal structure (Figure 37). We produced the double W624A/W717A mutant, and the novel peak in the spectrum of the W624A single mutant disappeared, strongly suggesting that it indeed belongs to the W717 residue (Figure 39). NMR hence confirms the rigid platform formed by the Trp624 and Trp717 indole rings, and shows that their mutually restricted movement is released upon removal of the Trp624 side chain. As the Trp717 stacks against the Glc+2 ring, its increased floppiness in the W624A mutant enhances futile interactions of the incoming acceptor, and thereby directly affects the kinetic control of the initial step of sucrose elongation reaction (Figure 38).

Moreover, the loss of rigidity of the stacking platform in the W624A mutant and the lack of the extended hydrophobic patch facilitates access of water to the active site and concomitant glucose release (hydrolysis). As the released glucose units will be effective acceptors for further elongation, this factor equally leads to a higher fraction of pure isomalto-oligosaccharides (Figure 39). Furthermore, the lack of stacking with Glc+3 (or Fru+3 when sucrose was the initial acceptor) contributes to enhance the shorter chain formation at the detriment of longer ones thus affecting the control of the size of the final products.

II.4. Monte Carlo model of the chain elongation.

As a first step in the elongation reaction, an incoming sucrose molecule is cleaved to release a fructose unit while its glucose part forms a β-D-glucosyl covalent intermediate with the side chain of D677. This glucosyl moiety can be transferred onto an incoming acceptor molecule (water, sucrose, glucose, or oligosaccharidic chain) that hence is elongated by one unit, or the potential acceptor can dissociate again without elongation, qualifying this event as a "futile encounter" (Bar-Even et al., 2015). Irrespective of the molecular mechanism underlying non-productive elongation, be it hydrolysis of the covalent intermediate, an off-rate that exceeds the catalytic rate or an incorrect positioning of sugar acceptors (sucrose, glucose or oligosaccharides), the same scenario repeats itself with a novel incoming sucrose molecule until transfer leading to elongation happens. If successfully elongated, the resulting molecule (a trisaccharide if the acceptor was a sucrose molecule) will return in solution, and compete again for binding to the enzyme active site with all other sugar molecules from the pool. We model this reaction process through a simple Monte Carlo (MC) simulation, whereby we correlate the probability of elongation (the "productive cycle") with the time spent in the active site of the enzyme. The additional binding energy of chains larger than three glucose units through stacking of the Glc+3 ring against the W624 indole ring (Figure 37) is interpreted in terms of a slower off-rate for the larger chains, thereby increasing the probability that those longer chains are further elongated. At the level of the MC algorithm, this is encoded by a probability p(s) for small chains to be elongated (and hence 1-p(s) for a non-productive encounter), probability that changes to p(l) when the incoming chain exceeds a certain length l. Details of the algorithm are given in the Supplemental Information. When starting with a given concentration of sucrose, the initial rate of elongation (measured by the fructose release) is dominated by the slow p(s). Once a certain number of chains have grown beyond the threshold (that we set here to 3 units, *vide supra*), the larger residence time translates in a higher probability for a productive encounter with transfer of a novel glucose moiety. Based on this model, sucrose consumption is predicted to accelerate when this threshold is reached (Figure 40), in agreement with the initial slow phase and experimentally observed acceleration of the DSR-M enzyme (Figure 38).

Most interestingly, our model shows that the relative value of elongation rates for long and short chains not only determines the biphasic reaction rate but equally the final distribution of chain lengths. When the long chains are truly favored in terms of their elongation probability (large p(l) value), those that reach a size of 3 or larger become effective seeds, and become longer and longer to the detriment of the many shorter ones (see SI for the details of the simulation). However, when the ratio decreases, short chains are not that disadvantaged anymore. They too become possible candidates for elongation, leading to a distribution with more and shorter chains in the end (Figure

40). The MC simulation with variable rate of elongation depending on the length of the incoming chain hence captures two salient features of the enzymatic reactions of the wt and W624A mutant DSR-M enzymes.

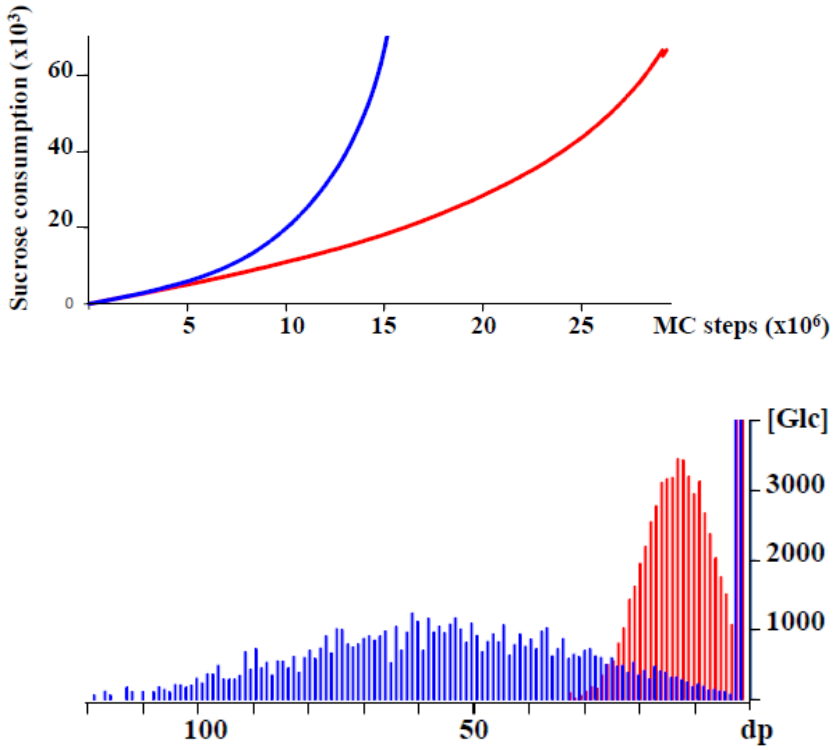


Figure 40: **Monte Carlo simulation of the enzymatic reaction**. The pool consists of 100,000 sucrose molecules, with an initial probability p(s)=0.001 for elongation, that increases to 0.5 for the wt enzyme (blue), or only 0.05 (red) for the DSR-M W624A mutant, once the chain exceeds three glucose molecules. The top panel shows the reaction rate, the bottom panel the length distribution of the final products shown as the Glucose concentration for a given degree of polymerization (dp).

III. Conclusion

Our crystal structure of the truncated α-transglucosylase DSR-MΔ2 in complex with an isomaltotetraose shows that the Trp624 indole ring is directly involved in anchoring the incoming chains to be elongated. Removal of this single indole ring leads to a decreased reaction rate and equally a final distribution of smaller dextran chains. NMR spectroscopy of thereby shows that removal of the Trp624 indole ring leads to a destabilization of the stacking platform, but also increases access of water to the active site (concomitantly promoting the formation of pure isomalto-oligosaccharide chains) and contribute to shorter chain formation. Directly related to the notion of futile enzymatic cycles, modeling the elongation process by a simple Monte Carlo simulation of the reaction whereby the probability of elongation increases when the chains reach a certain length can capture both salient features. Importantly, our results show that enzyme

engineering through the introduction or removal of discrete anchoring sites next to the active site can be used to control the length distribution of dextrans.

IV. Complementary work (not part of the publication)

IV.I. Further NMR analysis

In the characteristic zone of the ^{1}H, ^{15}N HSQC spectrum where we expect the Trp side chain resonances, we observe next to two intense resonances three less intense peaks (Figure NMR7, p152), suggesting some degree of mobility of the corresponding indole rings. The Trp in the streptavidin tag of the construct and the C-terminal Trp1290 with maximal accessible surface area in the crystal structure are plausible candidates for the intense peaks (Figure NMR8, p153). When we add dextran molecules of MM 11.3 kg/mol to the NMR sample, three novel indole resonances appear and become more intense upon increasing dextran concentration. In the crystal structure, Trp624 at the basis of the B2 loop and Trp717 (just before helix α5) directly stack against the sugar moiety, and hence are plausible candidates (Figure 41). Trp723 (in helix α5 and unique to DSR-M) is already further away, but anchors the B2 loop to the rest of the protein via a stacking interaction with Tyr619. Two more Trp residues, Trp648 and Trp1268, are at roughly the same distance of the sugar rings in the crystal structure, and hence might also correspond to one of the novel signals. Importantly, when we added the same molar concentration of glucose under the form of trioses (dp3), only one single of those resonances starts appearing (Figure NMR7, p152). All suggests that binding of the dextran to be elongated not only induces a displacement of the entire H1-H2 subdomain (Claverie et al., 2017) but also other dynamical and/or conformational changes when a minimal length of the product is reached.

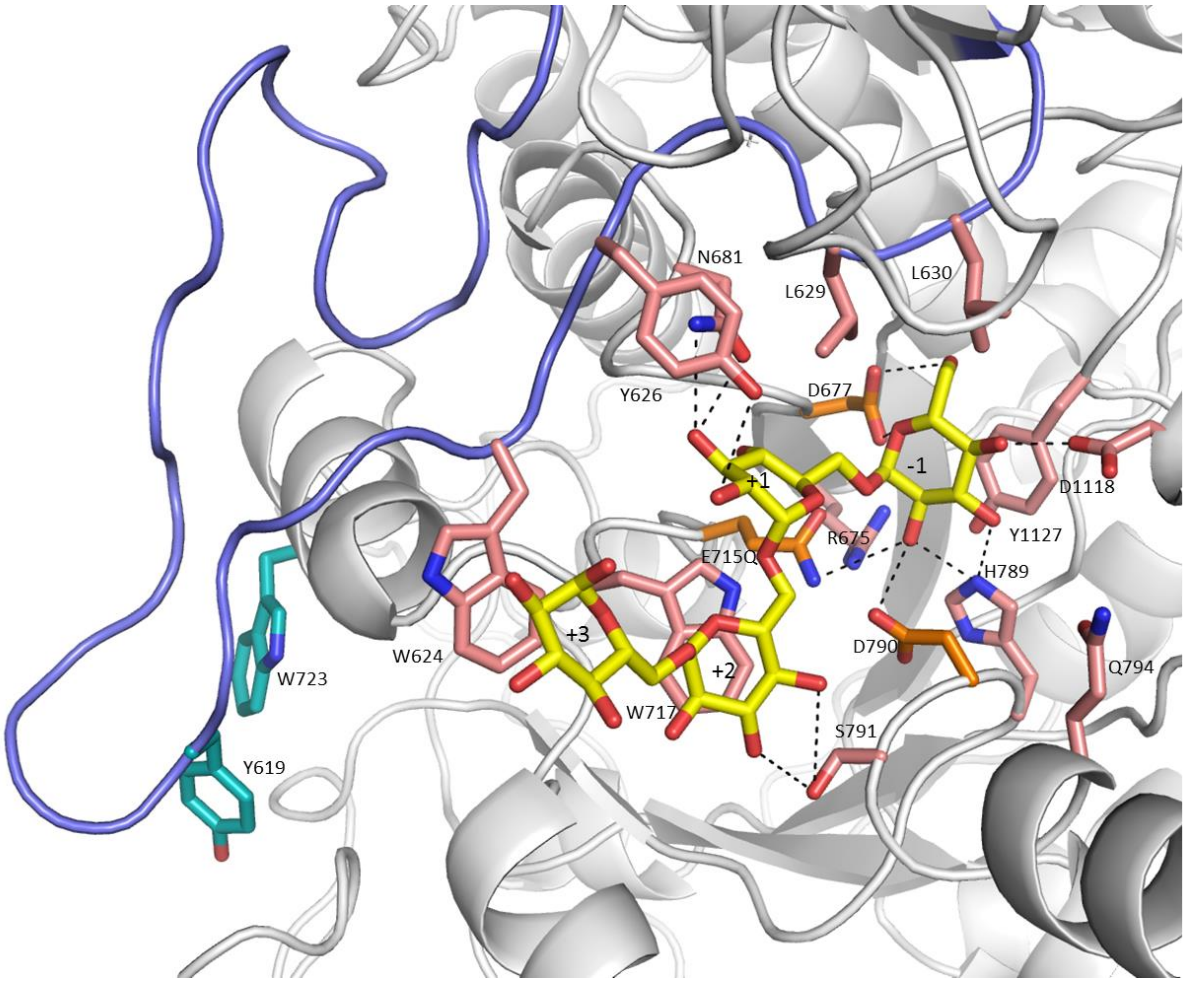


Figure 41: **Crystal structure of DSR-MΔ2-E715Q with an isomaltotetraose**. Indole rings form a platform for the nascent dextran chain with the first two glucose units stacking against the aromatic rings of Trp624 and Trp717. Catalytic residues are shown in orange sticks; residues participating to the catalytic site are shown in salmon sticks; Tyr619 and Trp723 mentioned in the text are shown in light blue sticks.

In order to explore the dynamic of loopB2, we mutated Tyr619 into Ala. This latter residue anchors the tip of the B2 loop to the structured core of the protein, notably through an aromatic stacking with the indole ring of Trp723 and through favorable H-bonds with the backbone atoms of Leu766, Leu767 and Asp764. We expect that this mutation could realese the tip of the loop that would gain mobility.

Thus, in an attempt to decipher the role of loop B2 mobility and the contribution of the Trp residues in interaction with oligodextrans as revealed by our NMR experiments, we further compared the reaction kinetics and product profiles obtained with ΔV, W624A and Y619A in the presence of sucrose alone, or sucrose and I2. When we monitored the fructose release as a function of time for the reaction with sucrose alone, a biphasic behavior was clearly visible for ΔV or Y619A mutant, with an initial slow fructose release which can be related to the low sucrose consumption rate (V_1)

followed after a couple of minutes by an accelerated one (V_2) (Figure 42, left). This behavior is similar to that reported for Δ2 (Claverie et al., 2017). Regarding the W624A mutant, the acceleration appears to be less pronounced. In the presence of I2, the biphasic behavior is noticeable for ΔV and Y619A mutant whereas the sucrose consumption rate remains constant for the W624A mutant. The observed increase of reaction rate in the presence of I2 acceptor can be due to stacking interactions with Trp717 (when I2 occupies +1 and +2 subsites) leading to I3 formation, whose accommodation as acceptor will be further facilitated by additional stacking with Trp624 in +3 subsite. Accordingly, when Trp624 is mutated, the reaction rate is constant (Figure 42) indicating that the interaction in subsite +3 is critical for rate acceleration. To check this point, acceptor reaction with I3 and I6 were performed and similar results were obtained. Whereas the slow initial phase disappeared and sucrose consumption rates nearly tripled for ΔV or its Y619A mutant, the W624A mutant strictly behaved as in the absence of acceptor. We attribute this to the positive and immediate contribution of the Trp624 indole ring in anchoring the acceptor when this latter is large enough to bridge from subsite +1 to subsite +3. The duo Trp717 and Trp624 is operational to position the acceptor in a proper configuration for glucosylation and subsequent product release. When the size of the acceptor goes up to a dextran 6 kg/mol, the rate of sucrose consumption further increases for ΔV or its Y619A mutant, but remains constant and low for the W624A mutant suggesting two comments. First, these effects were not observed with I6, showing that in the presence of oligosaccharide acceptors of DP higher than 6, additional interactions may occur in subsite +7 or higher. Among the candidates involved in these interactions, we could propose Trp723, which mobility was affected upon dextran 11.3 kg/mol addition. However, we cannot exclude other candidates in the proximity of Trp624 (ie: Tyr619, Tyr1201 or Tyr626) that could also act as supplementary anchorage points upon elongation. Second, abolishing the anchoring point (W624A) hence is sufficient for ineffective binding of larger acceptor. In agreement with this, communication with the domain V in the Δ2 forms of the enzyme seems equally broken in the W624A mutant: whereas reaction products are of larger size for Δ2 compared to ΔV, the W624A mutation leads to oligosaccharide chains of the same length for Δ2 or ΔV (Figure 43).

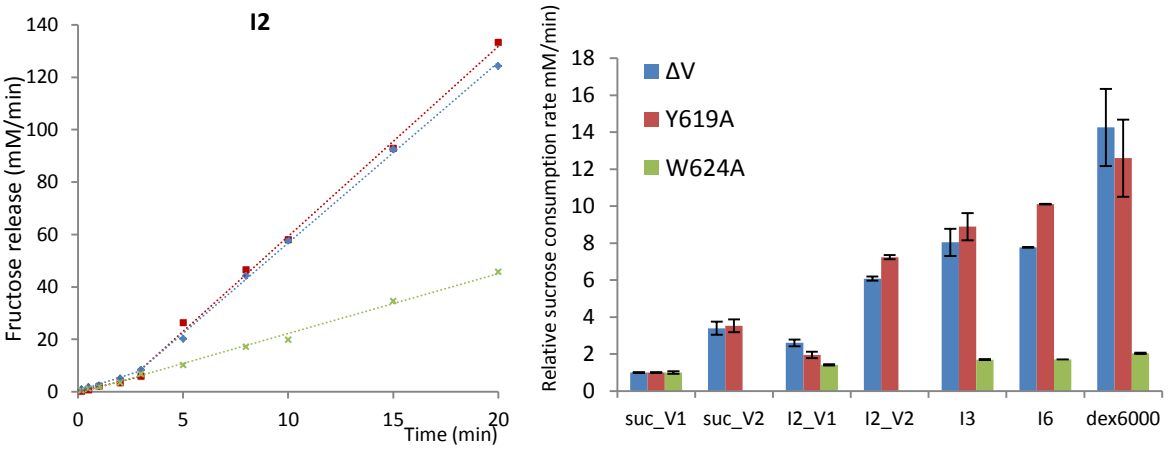


Figure 42: **Acceptor length determines the reaction speed for DSR-M and Y619A, but not for W624A.** (left) When sucrose alone or sucrose and I2 (shown) are given as initial substrate, sucrose consumption starts off low but increases threefold after 3 minutes for the DSR-MΔV or its Y619A mutant, but not for the W624A mutant where it remains low (right) Sucrose consumption rates

IV.2. The DSR-M W624A mutant produces short oligosaccharides in good yield

When characterizing the final products of the different enzymes after complete sucrose consumption, we previously found that the ΔV enzyme produces short dextran molecules with an average MM of 16 kg/mol. As mentioned before, the W624A mutant leads to even shorter dextrans, with a mean MM of 2.4 kg/mol, whereas the Y619A mutant yielded dextran chains of similar size as the wt ΔV. As already shown for the wt enzyme, a significant part of the mutant products consists of elongated sucrose and sucrose isomers, which are mixed with true oligodextrans (only composed of α-(1→6) glucosyl residues (Figure S14)). As the production of very LMM dextrans is of interest for industrial applications, we also performed acceptor reaction on glucose leading to the production of IMOS of DP2 to 12 in good yields (>52% and >47% for the reactions on 100 and 300 g/L glucose, respectively). The elongation of sucrose or sucrose isomers is disfavored due to the excess of glucose in the media, leading to the production of chains exclusively composed of α-(1→6) linked glucose molecules.

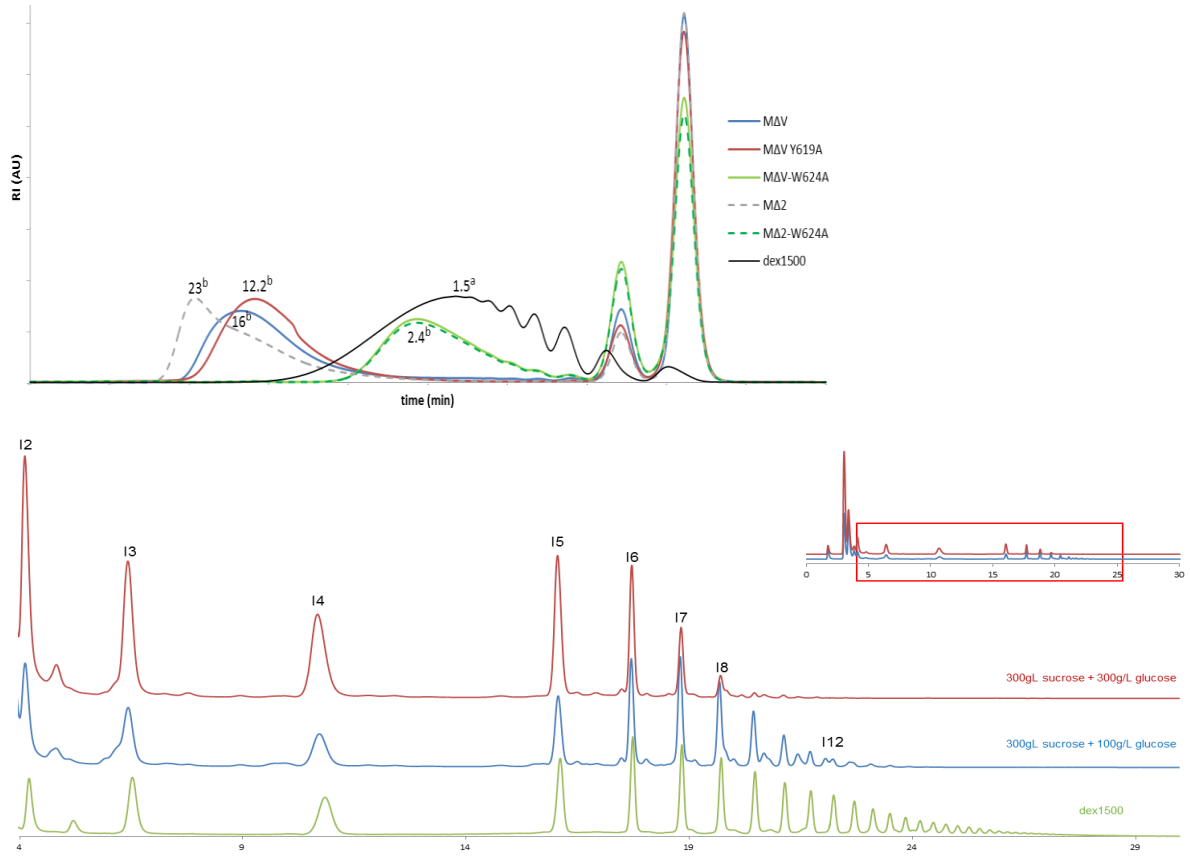


Figure 43: **DSR-MΔV W624A mutant produces LMM oligosaccharides.** (top) HPSEC profiles of the reactional medium after depletion of 292 mM sucrose (legend on graphic). Absence of product size difference between DSR-MΔ2 and ΔV W624A mutants, whereas wt Δ2 dextrans are longer than wt ΔV ones. Tyr619A mutant yielded dextran chains of similar size as the wt ΔV. (a) molar mass of commercial standards in kg/mol, (b) determined molar mass at peak apex in kg/mol. (bottom) HPAEC-PAD profiles after total sucrose depletion (legend on graphic) Glucose acceptor reactions lead to the synthesis of LMM IMOS fractions of interest for potential food applications.

IV.3. Discussion

The recently disclosed structure of DSR-M revealed that the domain V of this enzyme plays an important role for the production of low molar mass dextrans of around g/mol. The mode of synthesis was shown to be distributive and our analysis also revealed that, in the absence of domain V, the enzyme was still capable of synthesizing dextrans of around 16 kg/mol. The enzyme 3D structure revealed the presence of a loop, named B2, abnormally long in DSR-M compared to other GH70 enzymes of known structure, and suspected to play an important role for low molar mass dextran synthesis.

To investigate in further details the implication of this loop, soaking in the presence of higher amounts of I6 were performed and the X-ray structure, although solved at a low resolution, allowed the modelling of an I4 residue, a transitory product of the reaction, in the enzyme active site. This allowed to provide a map of the residues in interaction with the oligosaccharide in subsites -1, +1, +2

and +3. In particular, two Trp residues (Trp717 and Trp624), are proposed to be in stacking interactions with the oligosaccharide in subsite +2 and +3, respectively. Of note, if Trp717 is conserved in other GH70 enzymes, Trp624 is unique to DSR-M.

The role of these Trp residues, together with that of loop B2, was examined by NMR experiments using ^{15}N DSR-MΔV, a form devoted of the domain V to avoid any interference due to binding ability if this domain. These experiments, conducted for the first time with transglycosylases of such a high molar mass, were really informative. First, they revealed that one Trp residue shows an enhanced mobility upon I3 addition. Considering the size of the oligosaccharide, we propose to assign the corresponding signals to Trp717 located in subsite +2. Mutation of this residue quasi abolishes DSR-MΔV activity testifying of its crucial role in the elongation process, as previously shown for Trp1065, the equivalent conserved residue found in GTF180-ΔN (Meng et al., 2017). Adding longer dextrans resulted in the arising of two other Trp resonances that could be attributed to Trp624 located in loop B2, and Trp723 at the end of helix α5.

Mutation of Trp624 to Ala residue caused a mobility of Trp717 accompanied by a severe decrease of the dextran product size, which average molar mass drops from 16 kg/mol down to 2.4 kg/mol. Eliminating the anchorage point in +3 subsite is highly detrimental for elongation above 2 kg/mol. This demonstrates how important the synergy between Trp717 and Trp624 is. To support this, a correct positioning associated to a controlled mobility of loop B2 are also likely to be necessary for an optimal elongation. The control of this mobility is also mediated by another aromatic residue, Tyr619 (at the extremity of loop B2) in H-bond interactions with several residues and possibly in stacking interaction with Trp723. Compared to mutation of Trp624, mutation of Y619A also impacted but much less severely the size of the products, which reaches 13 kg/mol. All these conclusions are supported by the kinetic experiments performed in the presence of IMOS of increasing size, the wild type enzyme as well as the Y619A mutant being both activated by addition of IMOS whereas the W624A mutant is insensitive.

Finally, we also noticed an activation of the reaction rate in the presence of oligodextrans of 6 kg/mol. In parallel, the addition of 11.3 kg/mol to ΔV revealed the increased mobility of a third Trp, possibly Trp723, indicating that additional rearrangements may occur when oligosaccharides of DP higher than 6 are synthesized and interact with the active site. In this scenario, stacking interactions, possibly with Trp723 or Tyr619 itself, could provide additional anchorage points for defining a furrow for dextran binding or release assisted by the mobility of B2 loop. To identify the furrow in which longer dextrans bind, it would be necessary to isolate pure and longer isomaltooligosaccharides and perform characterization of other mutants.

IV. Complementary conclusion

We report herein the first 3D structure of a GH70 glucansucrase in complex with a transitory natural product bound in the active site. In parallel and to start tackling the question of the dynamics in the process of elongation of these enzymes, we also report the first series of ^{1}H ^{15}N NMR experiments, performed with a 100 kDa truncated form of the enzyme and several mutants. Completed by kinetic and product analyses, the study allowed unveiling the key role of a duo of Trp residues as mediator of binding and release of the products, probably accompanied by a controlled mobility of one loop. In addition, the study enabled us to generate efficient mutants synthesizing controlled MM dextrans and oligosaccharides of DP ranging from 1 to 12 directly from sucrose or with sucrose and glucose. Finally, it also provides new strategy to guide the insertion of Trp residues at correct positions to favor sugar elongation and turn tranglycosylases into efficient polymerases by protein engineering.

V. Material and methods

V.1. Protein expression and purification

Transformed *E. coli* BL21 star DE3 cells were used for protein expression as in (Claverie et al., 2017) and two-step protein purification (streptavidin affinity and size exclusion, see SI for details) was performed to obtain the preparation of pure enzyme used for biochemical and crystallographic studies. Construction of the plasmids used for enzyme production of DSR-MΔV and DSR-MΔ2, is presented in (Claverie et al., 2017).

V.2. Activity assays

Activity was assayed using the dinitrosalicylic method (Miller, 1959). One unit of DSR-M variant is defined as the amount of enzyme that catalyzes the production of 1 µmol of fructose per minute, from 292 mM sucrose, in 50 mM sodium acetate buffer pH 5.75, at 30°C.

V.3. Enzymatic reaction

Standard conditions: enzymatic reactions were performed on 292 mM sucrose, at 30 °C, using 1 U/mL of purified enzyme, in 50 mM sodium acetate buffer at pH 5.75. After total sucrose consumption, reactions were stopped by 5 min incubation at 95 °C, and samples were stored at -20 °C until further analyses.

V.4. Kinetic analysis of acceptor reactions

Acceptor reactions were performed in standard conditions with 292 mM sucrose and 4.6 mM of isomaltose (IM2), IM3, IM6, IM9, 6-kg/mol or 11.3-kg/mol dextran with 1 U/mL of purified enzyme. Samples were taken at regular intervals, and reaction was stopped by 10 min incubation at 95 °C.

V.5. Acceptor reactions on glucose with DSR-M W624A mutants

Acceptor reactions were performed in standard conditions with 292 mM sucrose and 100 or 300 $g.L^{-1}$ glucose with 1 $unit.mL^{-1}$ of purified enzyme. After total sucrose consumption, reactions were stopped by 5 min incubation at 95 °C, and samples were stored at -20 °C until further analyses. The yields

$$Y_{GOS} = \frac{[Glucosyl]_{mol}}{[Sacch]_{init_{mol}} + [Gluc]_{init_{mol}}} = \frac{\left([Sacch]_{init_{mol}} - [Sacch]_{final_{mol}}\right) + \left([Gluc]_{init_{mol}} - [Gluc]_{final_{mol}}\right) - [Leuc]_{final_{mol}}}{[Sacch]_{init_{mol}} + [Gluc]_{init_{mol}}} \ (\%_{mol})$$

V.6. Product characterization

The reaction medium was analyzed by HPLC as described in (Claverie et al., 2017) to determine the average molar mass of the products and assay glucose, fructose, and leucrose production as well as sucrose depletion by determining glucose production rates (hydrolysis activity), and fructose production rates that reflect both transglucosylation and hydrolysis activity.

V.7 Crystallization and Data collection

Crystals of the inactive mutant DSR-MΔ2-E715Q were obtained with a protein preparation at 6 mg/mL supplemented with 10 mM IM6 and 5 mM $SrCl_2$ in the following conditions: 0.2 M Sodium citrate, 0.1 M Bis Tris propane pH 6.5 and 20% (w/v) PEG3350. Crystals were cryoprotected in reservoir solution supplemented with 15% (w/v) glycerol and 100 mM IM6 after 1H soaking; they were cryo-cooled in a gaseous nitrogen flux at 100 K.

Diffraction data were collected at beam line ID30-B at the European Synchrotron Radiation Facility (Grenoble, France). The diffraction data were collected at 100K. The data sets were indexed and integrated using XDS and scaled using SCALA from the CCP4 suite (Winn et al., 2011)).

V.8. Structure determination

The structure of DSR-MΔ2-E715Q in complex with I6 was determined by molecular replacement with PHASER (McCoy et al., 2007) using the previously solved DSR-MΔ2 structure (Protein Data Bank entry 5O8L) as a search model. The high resolution structures of glucansucrases in complex with maltose or sucrose (PDB: 3KLL and 3HZ3, respectively) have been used as structural templates to guide the manual docking of an I4 molecule into the electron density. Manual rebuilding cycles using COOT (Emsley et al., 2010) alternated with restrained refinement using REFMAC5 (Murshudov et al., 1997) resulted in an incomplete model which lacked an important part (204 residues) of the N-terminal GBD. Final Rwork and Rfree values are 0.210 and 0.232, respectively. Coordinates and structure factors of the final model have been deposited at the PDB under the accession code 6HTV.

V.9. Mutagenesis studies

Mutants W624A, W717A and Y619A were constructed by inverse PCR (oligo-mediated introduction of site-specific mutations) using *dsrm-Δ2* and/or *dsrm-ΔV* plasmid as template and the primers described in Table 6. *E. coli* BL21 star DE3 cells (Invitrogen) were used as hosts for mutant productions. The reaction products from 292 mM sucrose were analyzed by HPSEC and HPAEC-PAD and compared with those obtained with the wild-type DSR-MΔV enzyme.

Table 6: **Primers used for the construction of the variants of this study (deletion or mutagenesis)**

Primer Name	Nucleotide Sequence	Main features
Y619Afor	caaccggccagcggagac	Mutation of position 619
Y619Arev	cgctggccggttggcttg	
W624Afor	ggagacgatgcgggatatgaac	Mutation of position 624
W624Arev	catatcccgcatcgtctccgc	
W717Afor	gaagatgcgtctcctaatgac	Mutation of position 717
W717Arev	ggagacgcatcttctaaaatgtg	

V.10. ^{15}N protein expression

Transformed *E. coli* BL21 star DE3 cells (Invitrogen) were grown on M9 minimal medium containing 5 mM KH_2PO_4, 10 mM Na_2HPO_4, 9 mM NaCl, 20 mM $^{15}NH_4Cl$, 1 mM $MgSO_4$, 0.1 mM $CaCl_2$, MEM Vitamin solution (Sigma Aldrich) and 2% glycerol. After about 5 h incubation at 37 °C to reach OD_{600} around 0.8-1.0, protein expression was induced by addition of 0.4 mM IPTG and cultures were incubated for 17 hours at 20 °C. Cells were then harvested by centrifugation, resuspended in purification buffer (20mM phosphate sodium buffer, pH 7.4, 500 mM NaCl, 20 mM imidazole) supplemented with EDTA free anti-protease tablets (Roche) and disrupted by sonication. After centrifugation, recombinant enzymes were recovered in the soluble fraction of the crude cell extract, ready for purification.

V.11. ^{15}NTrp protein expression

Transformed *E. coli* BL21 star DE3 cells (Invitrogen) were grown on M9 minimal medium containing 5 mM KH_2PO_4, 10 mM Na_2HPO_4, 9 mM NaCl, 20 mM NH_4Cl, 1 mM $MgSO_4$, 0.1 mM $CaCl_2$,MEM Vitamin solution (Sigma Aldrich) and 2% glycerol. After about 3 h incubation at 37 °C to reach OD600 around 0.4; 12 mg/L of ^{15}N labeled anthranilic acid kindly furnished by Roman J. Lichtenecker (Schörghuber et al., 2017) were added to the medium. After one additional hour incubation at 37 °C, protein expression was induced by addition of 0.4 mM IPTG and cultures were incubated for 17 hours at 20 °C. Cells were then harvested by centrifugation, resuspended in purification buffer (see SI) supplemented with EDTA free anti-protease tablets (Roche) and disrupted by sonication. After

centrifugation, recombinant enzymes were recovered in the soluble fraction of the crude cell extract, ready for purification.

V.12. NMR analysis of DSR-M variants

Pure ^{15}N enzyme preparations in MES buffer (30 mM MES pH 6.5, 100 mM NaCl, 0.05 g/L $CaCl_2$) at around 100 µM were analyzed by NMR on a Bruker Avance III 800 MHz spectrometer equipped with a TCI cryogenic probe head. Analyses were performed at 293K. Spectra were acquired with the TROSY pulse sequence as a 2D matrix of 1k x 128 complex points in the ^{1}H and ^{15}N dimension, respectively. Spectral windows were 14ppm centered on the water resonance for ^{1}H, and 40ppm centered at 118ppm for ^{15}N. A total of 256 scans per increment with a relaxation delay of 1.6s led to a total acquisition time of 16hours per spectrum. Spectra were apodized with a square sine function shifted by $\pi/2$ in both dimension, and Fourier transformed using the Bruker Topspin 3.0 software. Details of the NMR spectra are given in the NMR analyses section (p144).

V.13. Monte Carlo simulation.

Simulations were written in python, and executed on a standard PC. Input of the program demands the total number of molecules (50000), the probability of elongation for the short (0.01) and long (0.1-0.4) chains, the transition length from short to long (3). Details of the algorithm are given in the SI.

VI. Supplementary information

VI.1. Protein purification procedures

After growth, cells were harvested by centrifugation, resuspended in buffer A (PBS 1X, 280 mM NaCl, pH 7.4) at a final OD600nm of 200 and disrupted by sonication. The recombinant enzymes were recovered in the soluble fraction after centrifugation of the crude cells extract (15,000 × g, 30 min, 4 °C). Purification was then performed using the AktaXpress system (GE Healthcare, Vélizy-Villacoublay, France) at 8°C. Proteins were injected at 1 mL min^{-1} into a 5 mL Strep Trap HPTM column (prepacked with Strep Tactin Sepharose High PerformanceTM, GE Healthcare, Vélizy-Villacoublay, France), formerly equilibrated with buffer A. Elution was realized with a gradient of D-Desthiobiotin (from 0 to 2.5 mM) in buffer A over 10 column volumes at 4 $mL.min^{-1}$. Affinity purification was followed by a size exclusion step on superose 12 column 16x60 (GE Healthcare) from which the protein preparation is eluted with buffer B (30 mM MES pH 6.5, 100 mM NaCl, 0.05 $g.L^{-1}$ $CaCl_2$). Protein purity was verified by SDS-Page gel electrophoresis. Protein concentration was estimated by spectroscopy at 280 nm, using a NanoDrop instrument (Wilmington, DE, USA).

Theoretical molar extinction coefficients and molecular weights were calculated using the ExPASy ProtParam tool (http://web.expasy.org/protparam).

VI.2. Recombinant DSR-MDV protein sequence

Trp residues are colored in red, purification tags are highlighted in green for His tag and orange for Strep tag, and linkers are in grey.

MAHHHHHHVT SLYKKAGQNS VATSHNAAKS YDTKSFTNVD GFLTANSWYR PTDILRNGTK

WEPSTETDFR PLLMTWWPDK EVQANYLNYM SALGLGDQKI YTGASSQLDL NNAALIVQEA

IEKKISLEKS TKWLDDSIKS FIKSKRKDIQ GNLVDTNPGW TIDSETGSTN HLQNGAFIFT

624

NSPLVPEANA AEGNRLINRT PSQQTGNHIS YASQPYSGDD WGYELLLGND VDNSNPIVQA

EQLNWIHYLM NFGTITAPQD PDAHLANFDS IRIDAVDNVD ADLLQIAGDY FKAAYQVGEN

717

DKNANQHIHI LEDWSPNDVW YNQQVNGNSQ LTMDATMQNQ LLASLTRPIT SRDSMKSFTK

DALLVHRTAD NSYNQAVPNY SFIRAHDSEV QTIIAKIISD KHPDLYPTVD KALLAKDSAL

YDEAFTEYNA DMQKISSQKQ YTHNNMPSAY AILLTNKDTV PRVYYGDLFT DNGEYMANKT

PYYDAITSLL TARTKFVSGG QSLSVDKNDV LTSVRYGKGA LSATDNGSSD TRNQGIGVIV

SNNPNLDLNN DKVTLSMGIS HAHQAYRPLL LTNSQGIVAY ATDSEVPQNL YKTTNDKGEL

TFDASEIKGY DTVQTSGYLA VWVPVGASDE QDARTIASTE KNNGNSVYHS NAALDSQLIY

EGFSNFQTVP SKNASADEYA NVIIAKHAAD FNKWGVTSFQ MAPQYRSSTD GSFLDAVDTV

QNGYAFTDRY DLGFNAADGS KNPTKYGTDE DLRNAIKSLH AQKTYDGSSI QVMADFVPDQ

LYNMPLEQAV SVIRTDKYGV NSENPDIQNI IYAANIKSSG TDYQSIYGGK YLAELQKNPL

FKSLFDRIQI STKKTIDPNT RITQWSAKYF NGSNIQGKGI NYVLKDWASN KYFNVSSNDD

MYSRLPKQLM NQPAFLYKVV SAWSHPQFEK

VI.3. Electron density map and structural data statistics

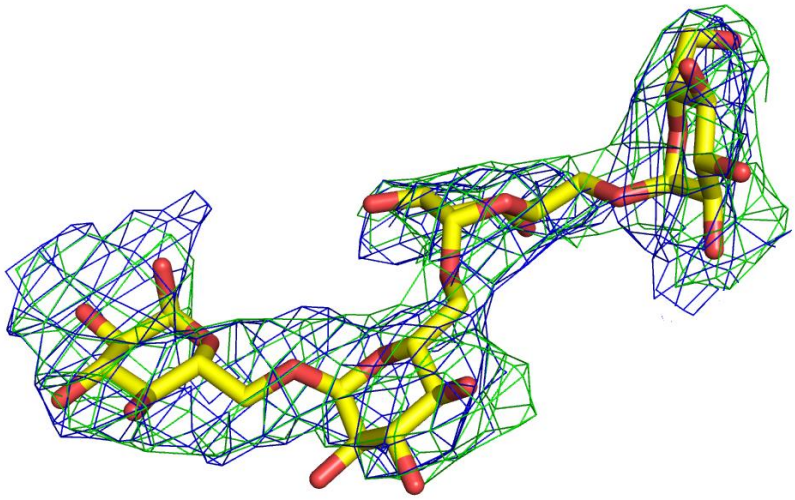


Figure S9: **Difference Fo-Fc map, contoured at 3σ (green) and difference 2Fo-Fc map, contoured at 1σ (blue), for the I4 molecule bound in the active site (yellow sticks).** PDB entry: 6HTV. A fifth glucose is probably present at the reducing end but the poor quality of the maps does not allow the modelling.

Table S4: **Data collection and refinement statistics**

	DSR-MΔ2 E715Q I4 Complex
Data collection	
Space group	$P3_121$
Unit-cell parameters (Å)	a=b=185.3 c= 150.8
	γ=120°
No. of molecule/AU	1
Matthews coefficient ($Å^3$/Da)	4.98
Solvent content (%)	75.3
Wavelength (Å)	0.9762
Resolution range (Å)	50.00-3.90 (4.11-3.90)*
No. of observed reflections	124806 (16205)*
No. of unique reflections	26662 (3826)*
Completeness (%)	97.3 (95.9)*
Multiplicity	4.7 (4.2)*
I/σ(I)	5.2 (1.2)*
R_{merge}	0.135 (0.658)*
CC ½	0.99 (0.71)*
Wilson B ($Å^2$)	85.6
Refinement	
Rwork/Rfree	0.200/0.228
RMSD Bond lengths (Å)	0.011
RMSD Bond angles (°)	1.36
Rama Favored (%)	96.6
Rama Allowed (%)	3.0
Rama Outliers (%)	0
Mean B; all atoms ($Å^2$)	113.2
Molprobity score/percentile	13.47/97
PDB Code	To be deposited

VI.4. Analysis of reaction medium after 15min reaction before and after invertase digestion.

Fifteen minutes after enzyme addition, reactions were stopped by 10 min incubation at 95 °C. Resulting samples were analyzed by HPAEC-PAD directly and after invertase digestion (β-Fructofuranosidase from Sigma Aldrich specific to sucrose hydrolysis). These digestions were performed by incubation of 15 µL reaction medium with around 300 units of invertase over night at 37 °C. Comparison of the products profiles before and after invertase digestion reveals a pic offset resulting from the release of a fructose molecule at chain extremities meaning that sucrose is the preferred acceptor at the beginning of the reaction for both wt (Figure S10) and W624A DSR-MΔV

(Figure S11) enzymes. However, concerning the mutant, a non-negligible population of pure isomato-oligosaccharides is present before invertase digestion reflecting a higher propensity for hydrolysis increasing the pool of glucose as initial acceptors.

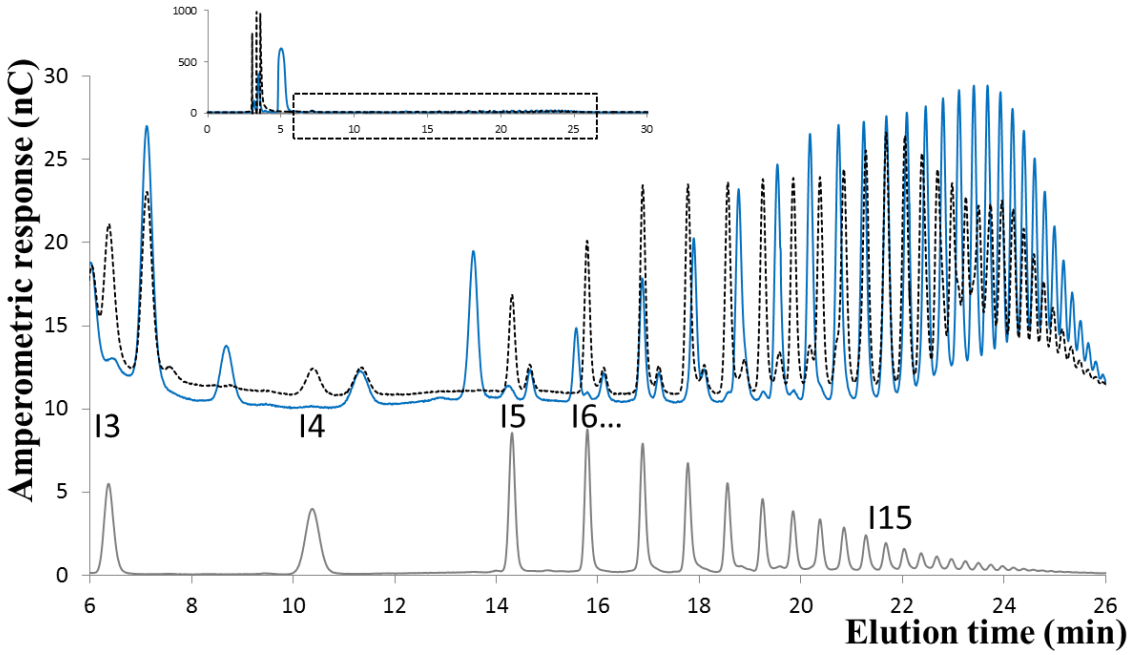


Figure S10: **HPAEC-PAD profiles of reaction medium at 15min reaction using DSR-MΔV on 292mM sucrose.** Blue: Before invertase digestion, Black dots: after invertase digestion. Grey: commercial dextran 1500. Reaction products are mainly elongated sucrose.

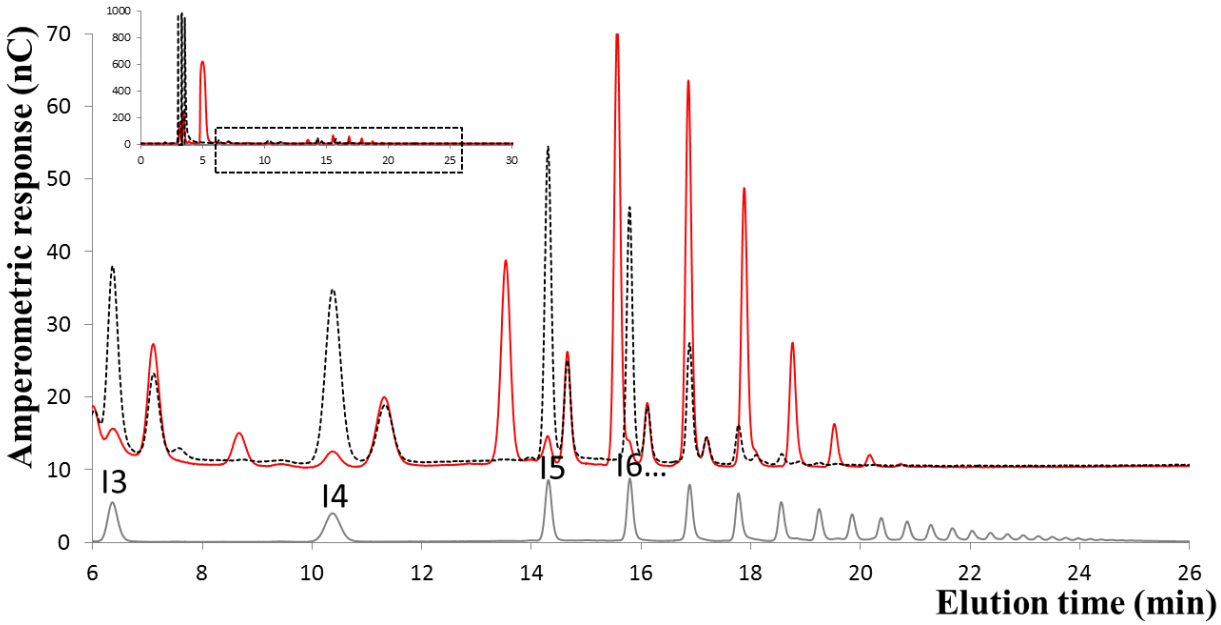


Figure S11: **HPAEC-PAD profiles of reaction medium at 15min reaction using DSR-MΔV-W624A on 292mM sucrose.** Red: Before invertase digestion, Black dots: after invertase digestion. Grey: commercial dextran 1500. Reaction products are mainly elongated sucrose but pure isomaltooligosaccharides are present in major amount than the wt enzyme.

VI.5. Monte Carlo model of the chain elongation.

The enzymatic process is encoded by a python program that basically uses two random numbers to (1) select a random molecule to be elongated from the pool of molecules that are not cleaved, and (2) decide whether this molecule will be elongated or rather diffuse back from the active site in solution, without any modification.

The starting population of n molecules (default value : n=50000) is represented by an initial array aa of this size, with all values set to 1 (= number of glucose units in the molecule, 1 for every molecule in the initial sucrose pool). A random molecule is chosen, and its elongation (or not) is decided upon by comparing a random number with a predefined threshold. The value of the molecule will increase by one when the second random number is smaller than the threshold, but at the expense of another sucrose molecule whose value will be set to 0, as it was the donor of the glucose unit (Figure S12). This ensures that the total glucose concentration remains constant.

A random molecule with index b is chosen (by a random number generator, selecting this number b between 1 and the size of the available acceptors, 50000 at the first step). This molecule will be elongated (= its value set to 2) if a random number (a) between 0 and 1 is smaller than p(s), here set to 0.001. If this is not the case, the step counter is increased by one and a novel molecule is chosen. If the random number is smaller than p(s), the value of aa[b] will be increased by 1. At the start of the reaction, this means it will go to 2. Because this requires the cleavage of a sucrose as donor, the algorithm looks in the set of sucrose molecules (array of aa with value of 1) for the one with the lowest index, and puts it to zero. This molecule hence will not be a potential acceptor anymore. Consumption is defined as the number of molecules set to zero, and is stored as a function of reaction steps (as a proxy for time).

When the reaction proceeds, a molecule with size 3 or larger can be chosen. If such is the case, this molecule will be elongated (= its value set to its current value + 1) if a random number (a) between 0 and 1 is smaller than p(l). Because this value is larger than p(s), the probability of elongation increases. A direct result is that the consumption of sucrose is expected to accelerate, as can be seen in Figure 38. A more indirect result is that these molecules can deplete the pool of available donors all the more efficiently as they tend to be elongated more efficiently. In other words, the larger is the ratio between p(l) and p(s), those that have reached the critical length l become more and more effective seeds. In the extreme case, imagine $p(s)=10^{-6}$ and p(l)=1, the single molecule that reaches by chance the length 3 will be elongated every time it is selected (with a chance of 1/50,000), whereas any other molecule has only 1 chance out of a million to be elongated. As a result, one can expect a single molecule of length ~50000 at the end of the simulation.

The end of the simulation is determined by the depletion of the pool of potential donors. This is not exactly the same as the total consumption, as can be seen on the following small example of 5 molecules (Figure S13). Indeed, in the extreme case when only sucrose molecules are elongated to a value of two (and never beyond), the available donor pool will decrease at double the rate of the consumption.

The final parts of the routine produce the histograms both of the chain length and of the glucose concentration corresponding to the different chain lengths and finally of the consumption.

Monte carlo model code:

```
import numpy as np
import argparse
def regulate(output_dir='.', n=5000, s=0.001, l=0.4, t=3):
    """generate 3 files results.
    Args:
        output_dir (str): directory path where to write results
        n (int): number of molecules
        s (float): small size molecules probability
        l (float): large size molecules probability
        t (int): minimal length to transit from short to long
    """
    i = 0
    aa = np.ones(n)  # sucrose molecules are both donors and acceptors
    xx = np.ones(n)  # Counter for the donors
    c = 0  # consumption of sucrose start -  will be updated during the algorithm
    outfile = open(output_dir+'/AutoconsumptionLength' +
                   '_'+str(l)+'_'+str(s)+'_'+str(n), 'w')
    r = n  # molecules available for elongation – will be updated during the algorithm
    while c < n:
        i = i+1
        bb = np.nonzero(aa)
        cc = bb[0]
        a = np.random.random(1)
        # index of the random molecule selected for elongation
        b = np.random.randint(len(cc))
    # selected from the entire list except for the cleaved molecules
        q = l
        if aa[cc[b]] < t:
            q = s
        if a < q:  # If true, the molecule becomes elongated by 1
            aa[cc[b]] = aa[cc[b]]+1
            xx = np.where(aa == 1)[0]
            # but this implies also that a sucrose molecule has been cleaved
            aa[xx[0]] = 0
            r = r-1
        c = n-r
        if i % 1000 == 0:
            e = str(c)
            m = int(np.max(aa))
            print('Cycle = ', i, 'Consumption = ', c, 'Max distribution = ', m)
            outfile.write(e+"\n")
        if (len(xx) < 0.1*n):  # We stop at 10% of monomers
            c = n
    print('Number of cycles = ', i)
    m = int(np.max(aa))
    print('Maximum of distribution = ', m)  # Used to construct a histogram
    kk = np.arange(m)  # array from 0 to (m-1)
    yy = np.histogram(aa, bins=kk)
    zz = yy[0]
    outfile = open(output_dir+'/AutohistogramLength' +
```

```
            '_'+str(l)+'_'+str(s)+'_'+str(n), 'w')
    i = 0
    while i < m-1:
        a = str(zz[i])
        outfile.write(a+"\n")
        i = i+1
    outfile.close()
    outfile = open(output_dir+'/AutohistogramLength_glucose'+'_' +
            str(l)+'_'+str(s)+'_'+str(n), 'w')
    i = 0
    while i < m-1:
        a = str(zz[i]*i)
        outfile.write(a+"\n")
        i = i+1
    outfile.close()
if __name__ == "__main__":
    parser = argparse.ArgumentParser(argument_default=argparse.SUPPRESS)
    parser.add_argument(
        "output_dir", help="dir path where to write the result files ")
    parser.add_argument("-n", type=int, help="number of molecules")
    parser.add_argument(
        "-s", type=float, help="small size molecules probability")
    parser.add_argument(
        "-l", type=float, help="large size molecules probability")
    parser.add_argument(
        "-t", type=int, help="minimal length to transit from short to long")

    args = parser.parse_args()
    regulate(**vars(args))
```

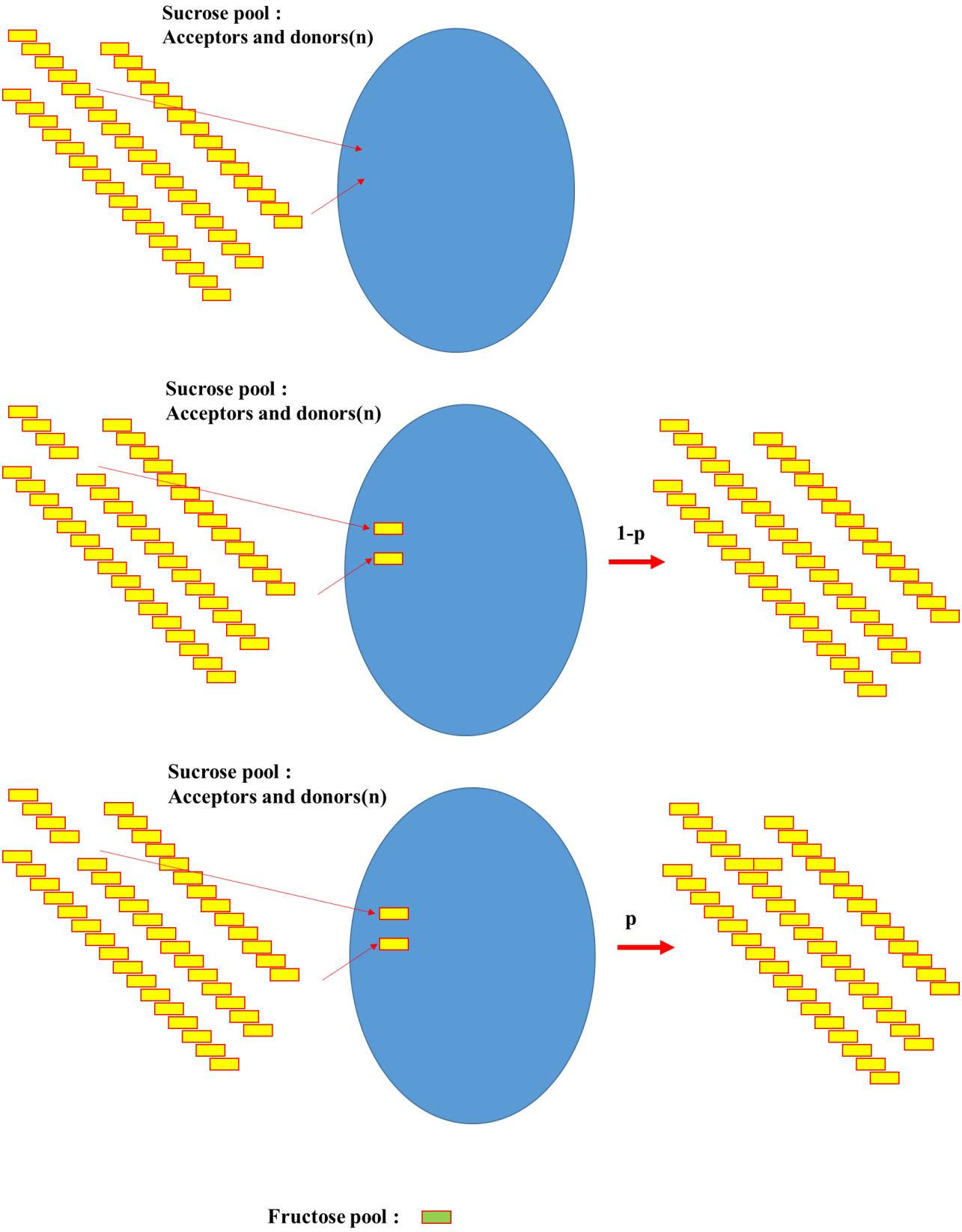

Sucrose pool :
Acceptors and donors(n)
Sucrose pool :
Acceptors and donors(n)
1-p
Sucrose pool :
Acceptors and donors(n)
p
Fructose pool :

Figure S12

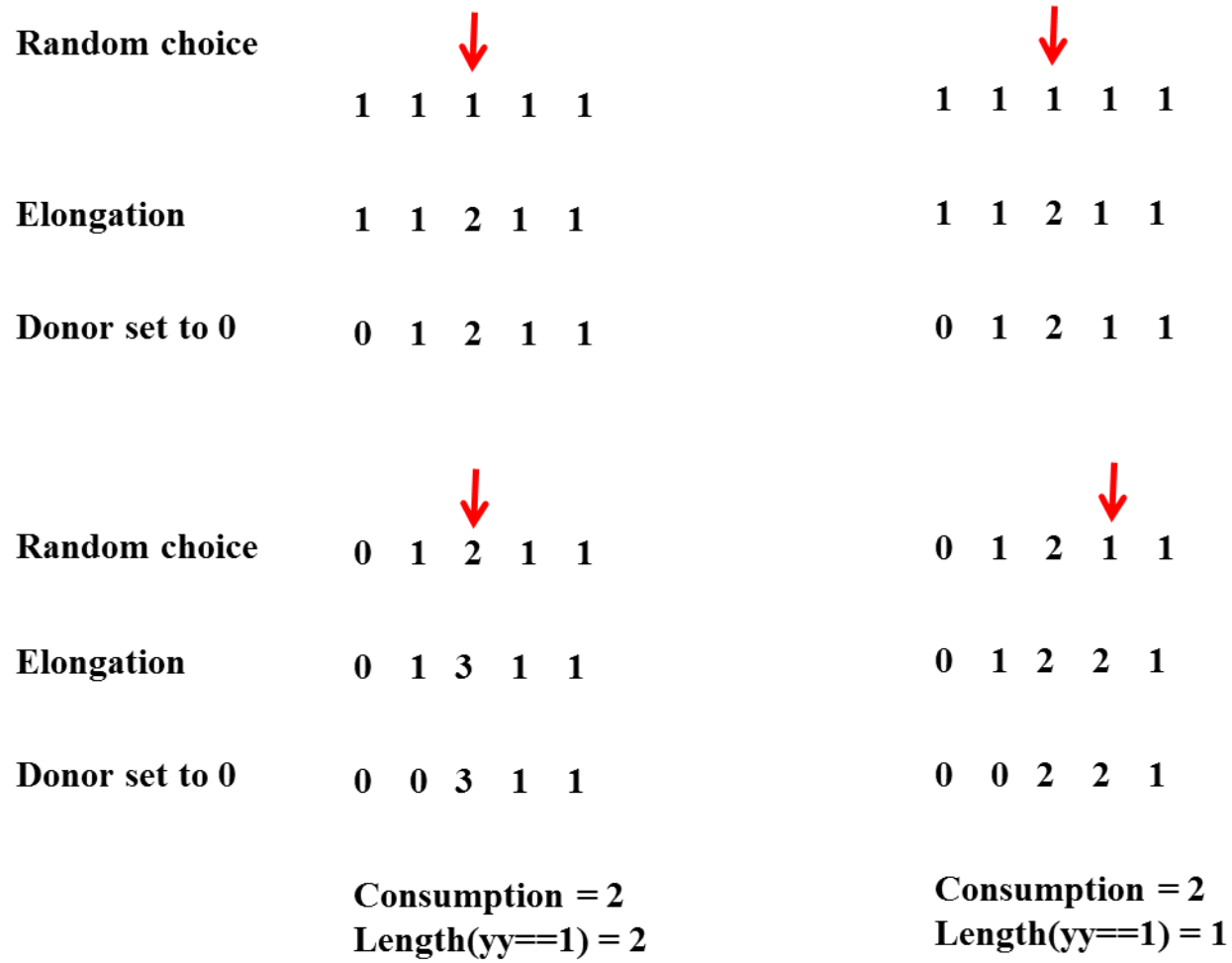

Problem of ending the algorithm
Random choice
1 1 1 1 1
Elongation
1 1 2 1 1
Donor set to 0
0 1 2 1 1
Random choice
0 1 2 1 1
Elongation
0 1 3 1 1
Donor set to 0
0 0 3 1 1
Consumption = 2
Length(yy==1) = 2
1 1 1 1 1
1 1 2 1 1
0 1 2 1 1
0 1 2 1 1
0 1 2 2 1
0 0 2 2 1
Consumption = 2
Length(yy==1) = 1

Figure S13

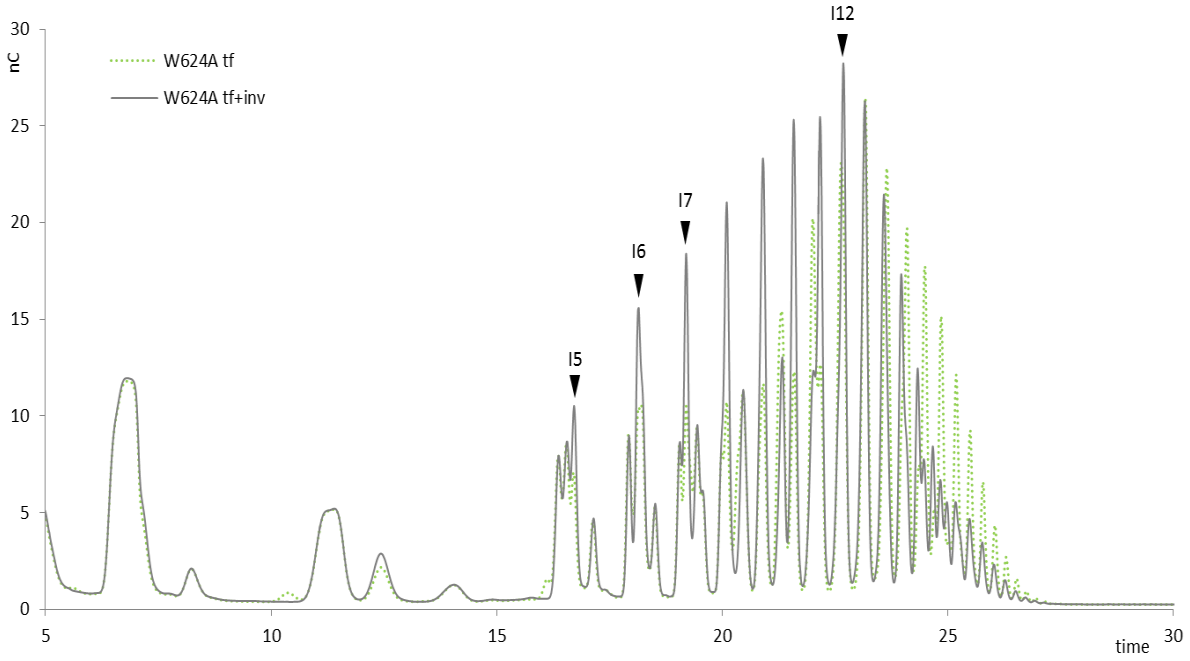


Figure S14: **Invertase digestion**. HPAEC-PAD chromatogram of products formed at the end of reaction on 292 mM sucrose by DSR-MΔV W624A, before (green dash) and after (gray) partial digestion with invertase. The majority of mutant products consists of elongated sucrose and sucrose isomers, which are mixed with true oligodextrans.

VII. NMR analyses

After purification by gel filtration and concentration to ~100 µM, uniformly ^{15}N labeled enzyme preparations in MES buffer (30 mM MES pH 6.5, 100 mM NaCl, 0.05 $g.L^{-1}$ $CaCl_2$) were analyzed by NMR on a Bruker Avance III 800 MHz spectrometer equipped with a TCI cryogenic probe head. All analyses were performed at 293K.

The one-dimensional proton spectrum was acquired using the watergate W5 pulse sequence with gradients and a double echo for water suppression (Liu et al., 1998). The recycle delay was set to 1s, whereas the delay of the binomial water suppression was set to 80µs. The spectral window was 14ppm sampled with 2k points, and 256 scans were acquired for a total measurement time of 8 minutes. Spectra were transformed after apodization with a square sine function shifted by $\pi/3$ and zero filling to 8k using the Bruker Topspin 3.0 software. The resulting 1D spectra were used to assess the integrity of the protein samples, notably through the detection of the methyl signals at -0.9 and -0.2ppm (Figure NMR 1-3-5).

Two dimensional spectra were acquired with the TROSY pulse sequence (MEISSNER et al., 1998; Pervushin et al., 1998) (trosyf3gpphsi19.2) as a 2D matrix of 1k x 128 complex points in the ^{1}H and ^{15}N dimension, respectively (Figure NMR 2-4-6). Spectral windows were 14ppm centered on the

water resonance for the ^{1}H dimension, and 38ppm centered at 118.5ppm for the ^{15}N dimension. A total of 256 scans per increment with a relaxation delay of 1.6s led to a total acquisition time of 16hours per spectrum. Spectra were apodized with a square sine function shifted by $\pi/2$ in both dimensions, and after zero filling to a 2k x 1k matrix Fourier transformed using the Bruker Topspin 3.0 software.

For the selectively labeled samples obtained with the ^{15}N labeled anthranilic acid, after a first spectrum with the same parameters as above, the nitrogen carrier was placed at 128ppm and the spectral limit in the ^{15}N dimension limited to 8ppm.

NMR Figures

To allow a better visibility on the NMR spectra, large figures are presented in this section.

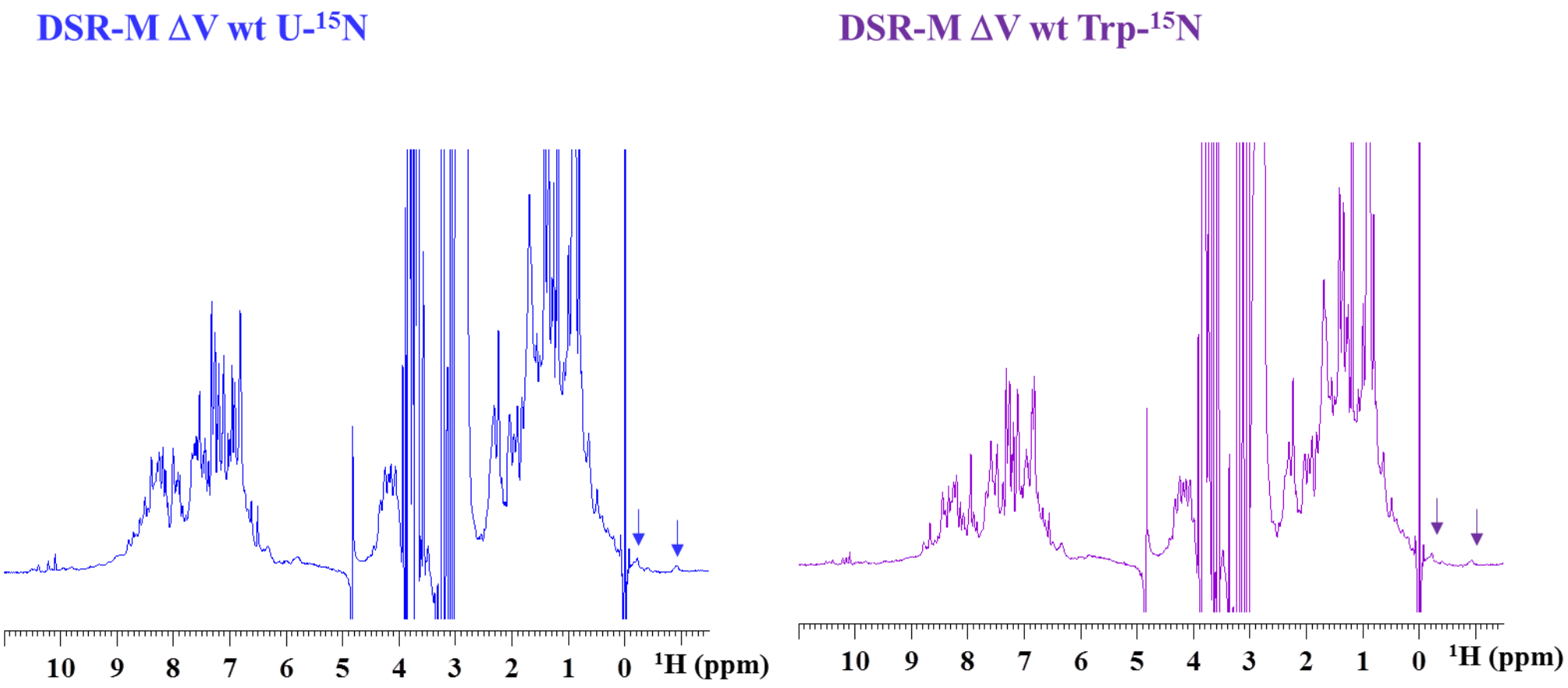


Figure NMR1: **1D proton NMR spectra of DSR-MDV wt U-^{15}N labeled (left) and DSR-MDV wt Trp-^{15}N labeled (right)**

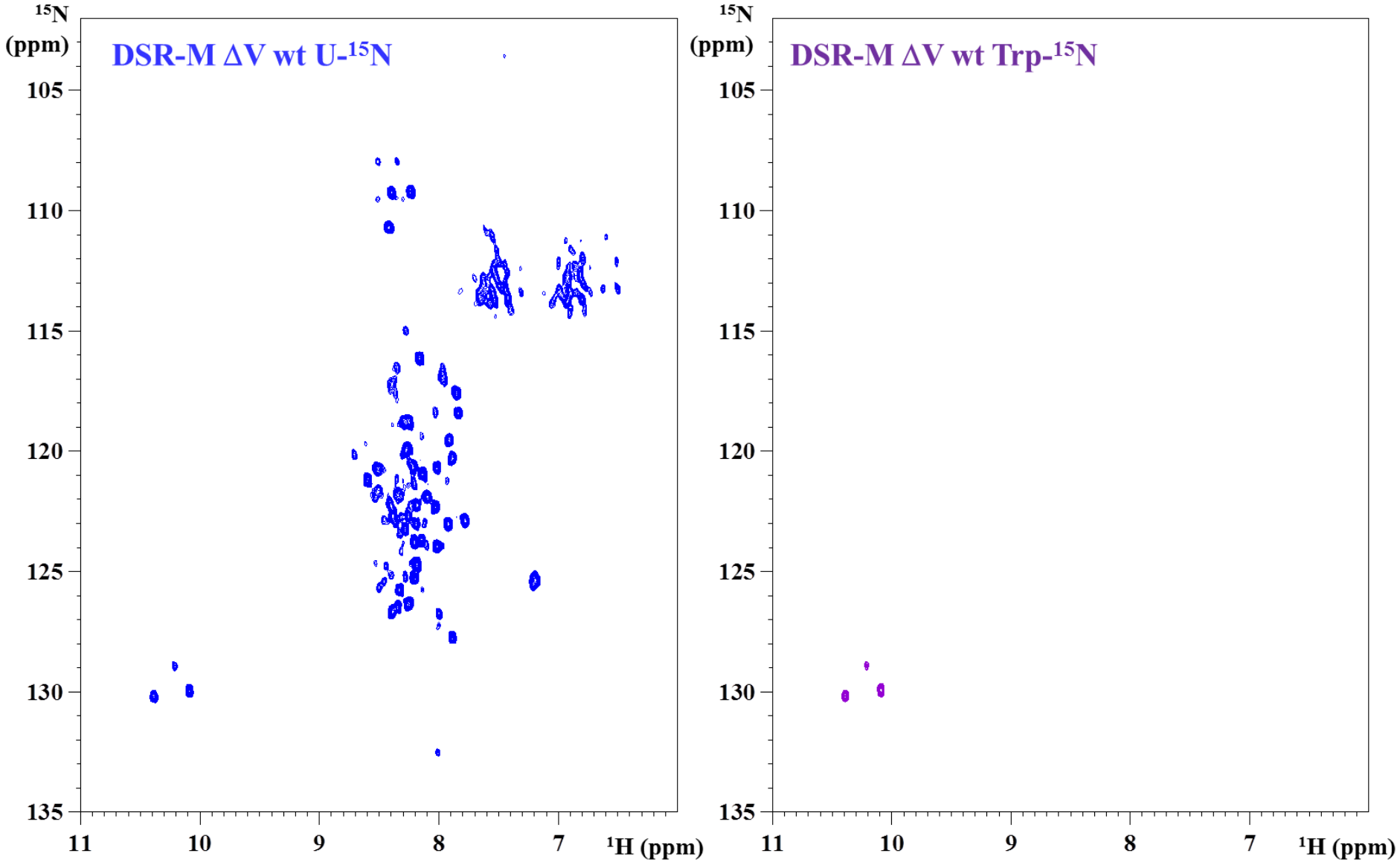


Figure NMR2: **2D ^{1}H, ^{15}N TROSY spectra of DSR-MDV wt U-^{15}N labeled (left) and DSR-MDV wt Trp-^{15}N labeled (right)**

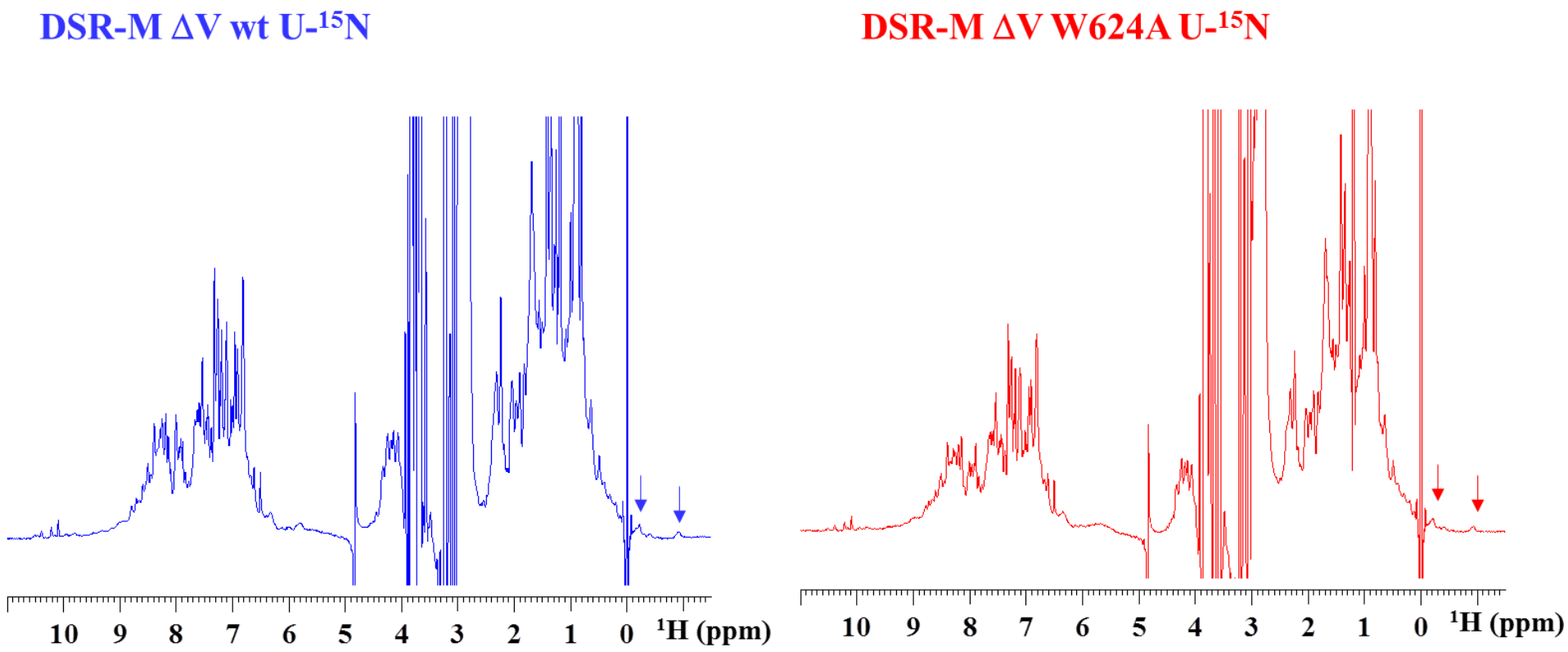


Figure NMR3: **1D proton NMR spectra of DSR-MDV wt U-^{15}N labeled (left) and DSR-MDV W624A U-^{15}N labeled (right)**

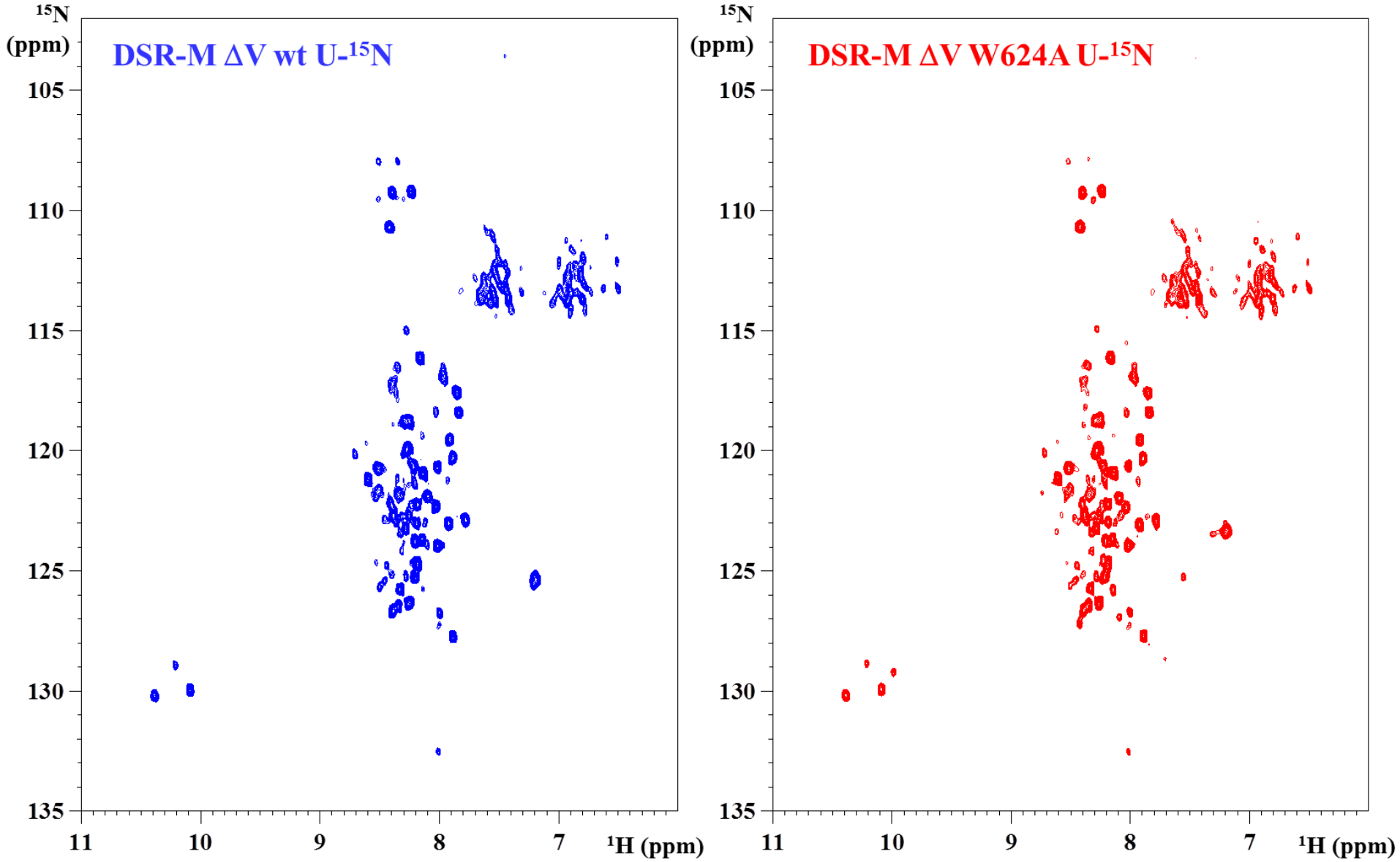


Figure NMR4: **2D ^{1}H, ^{15}N TROSY spectra of DSR-MDV wt U-^{15}N labeled (left) and DSR-MDV W624A U-^{15}N labeled (right)**

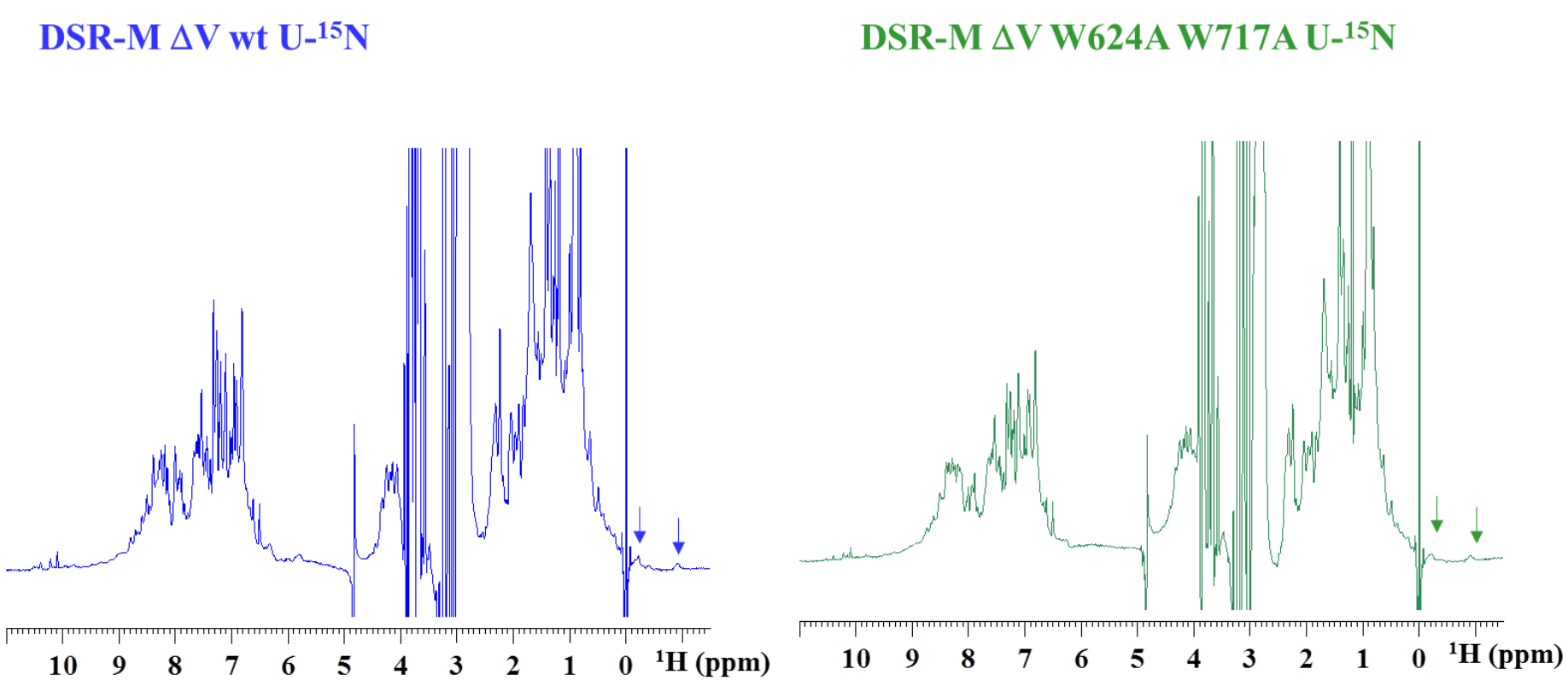


Figure NMR5: **1D NMR spectra of DSR-MDV wt U-^{15}N labeled (left) and DSR-MDV W624A W717A U-^{15}N labeled (right)**

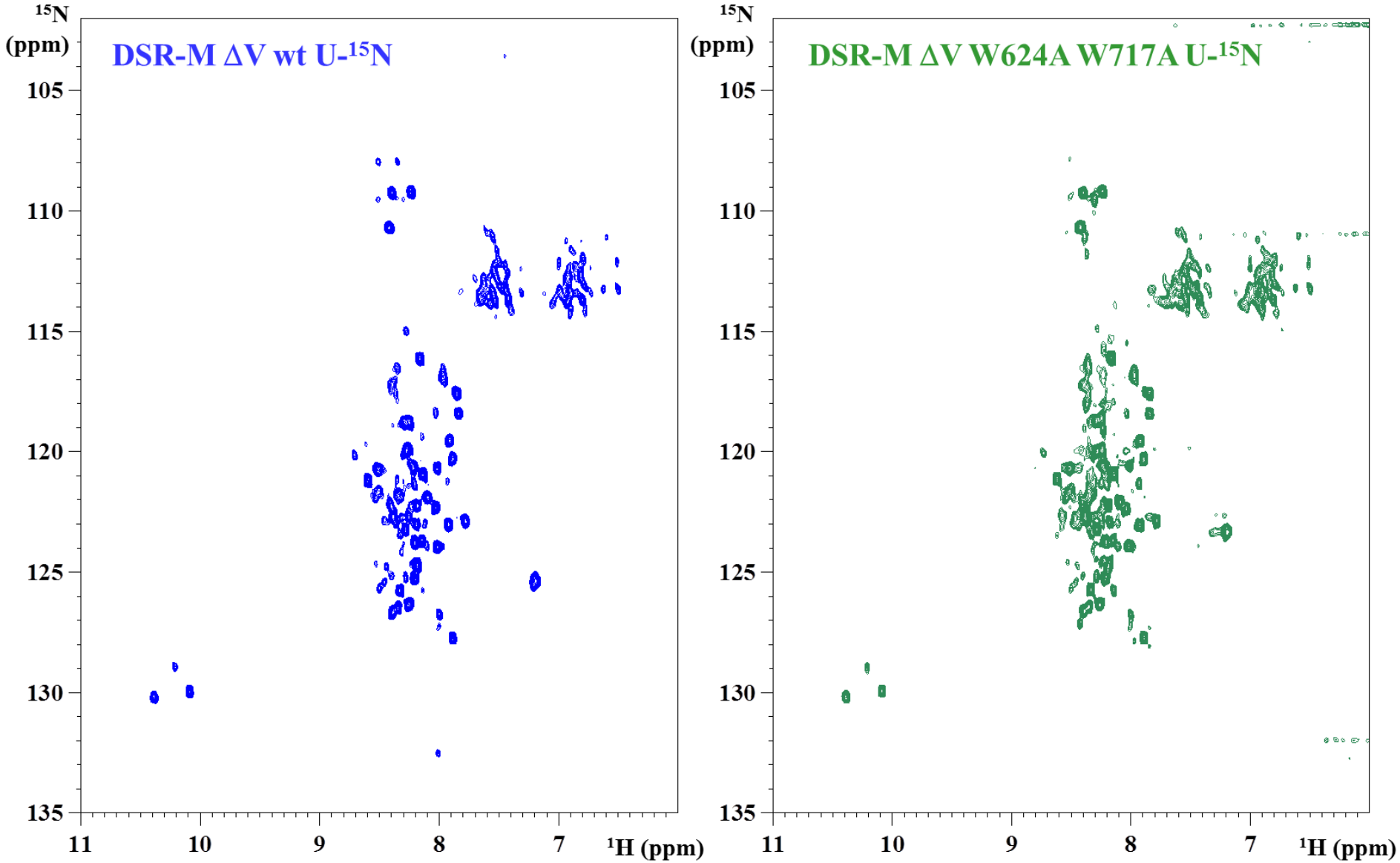


Figure NMR6: **2D ^{1}H, ^{15}N TROSY spectra of DSR-MDV wt U-^{15}N labeled (left) and DSR-MDV W624A W717A U-^{15}N labeled (right)**

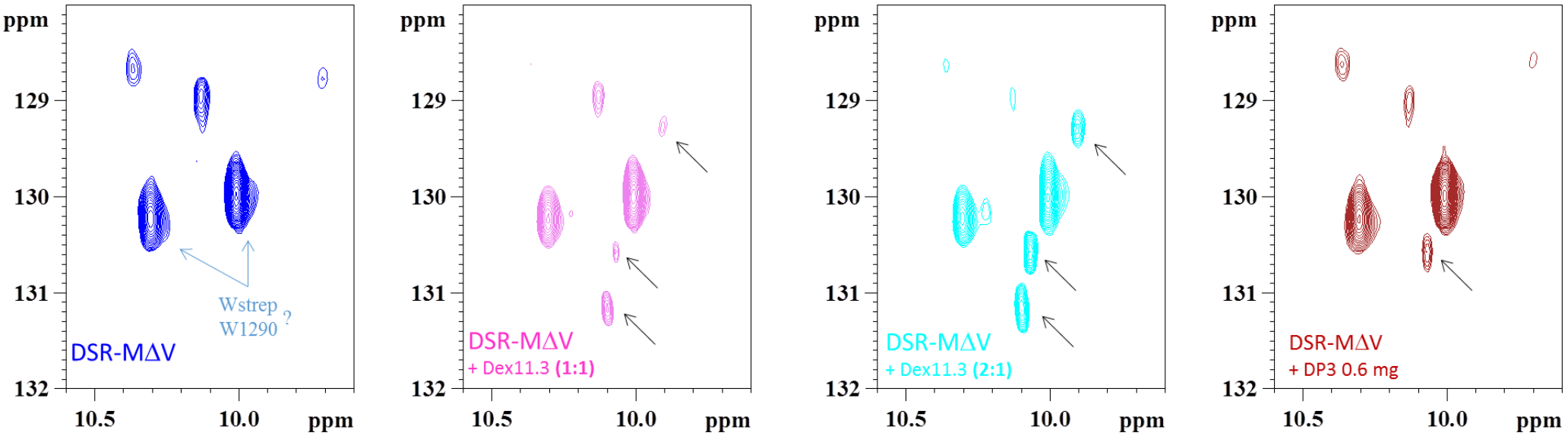


Figure NMR7: **Zooms of the ^{1}H, ^{15}N HSQC spectra of DSR-M ΔV centered on the Trp indole side chain**. Addition of 11.3 kg/mol dextran allows the observation of three novel signals (blue: 100µM ΔV, pink: 100µM dextran + 100µM ΔV, light blue: 200µM dextran + 100µM ΔV), whereas the same amount of glucose added as dp3 oligosaccharides (brown) has a more limited effect on the spectrum.

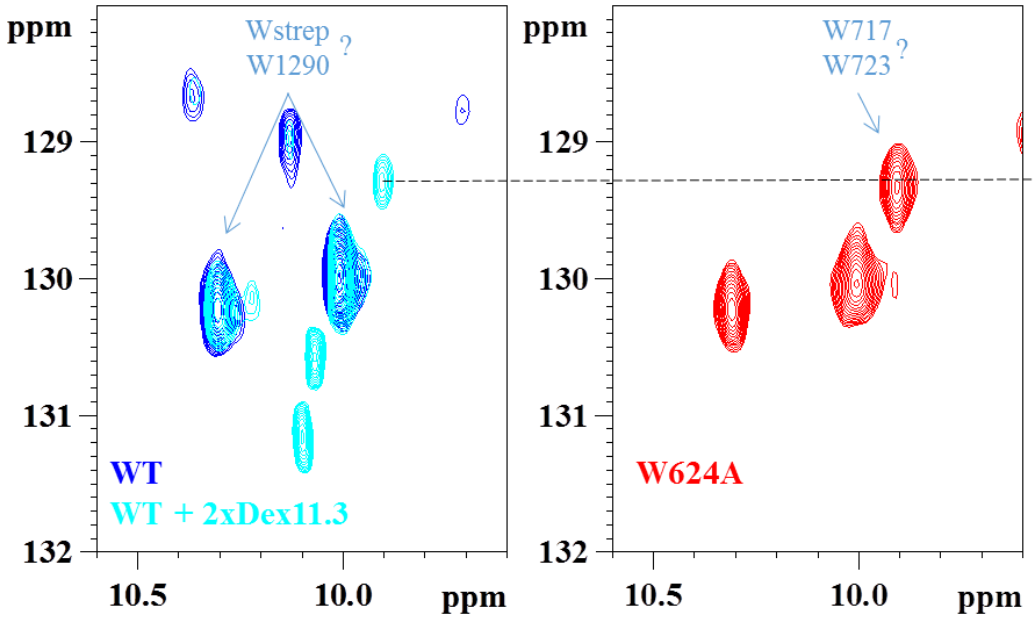


Figure NMR8: Zooms of the ^{1}H, ^{15}N HSQC spectra centered on the Trp indole side chain. Blue: ΔV superposed with light blue: addition of a 2:1 ratio of 11.3 kg/mol dextran, red: ΔV W624A, magenta: ΔV Y619A; green: ΔV Y619A-W624A.

Chapter IV

High molar mass dextran synthesis by DSR-OK dextransucrase from *Oenococcus kitaharae* DSM 17330, role of the domain V?

Marion Claverie, Gianluca Cioci, Marlène Vuillemin, Pauline Bondy, Jérémy Esque, Magali Remaud-Simeon and Claire Moulis.

In preparation for publication

I. Introduction

Glucansucrases (GSs) of the family GH70 (Lombard et al., 2014) are bacterial enzymes producing homopolysaccharides of D-glucosyl units from sucrose, a cheap and abundant agroresource. Among them, dextransucrases synthesize α-glucans comprising more than 50% of α-(1→6) osidic linkages called dextrans, which usually display very high molar mass (HMM) higher than 10^6 g/mol. However, few representatives were reported to produce mixed populations of dextrans of low (LMM) or high molar mass (Côté and Robyt, 1982; Joucla et al., 2006; Moulis et al., 2006b) and recently, the first natural dextransucrase synthesizing only LMM dextran of 20 to 30. 10^3 g/mol was described (Passerini et al., 2015; Claverie et al., 2017). HMM dextrans can be used for their texturing, thickening and emulsifying properties, especially in food and cosmetic (De Belder, 2003; Kothari et al., 2014; Vettori et al., 2012), for instance. But most of the applications concern fractions of 1, 10, 40 or 70. 10^3 g/mol, usually obtained by partial acid hydrolysis and solvent-fractionation of the native polymer produced by fermentation of the lactic acid bacteria *Leuconostoc mesenteroides* NRRL B-512 F on sucrose carbon source (Naessens et al., 2005). The resulting dextran displays a high α-(1→6) linkage content (>95%) and presents a low antigenicity and high water solubility of interest for medicinal uses (Lindberg and Svensson, 1968). After hydrolysis and fractionation, these clinical dextrans serve as infusion fluids, volume expanders, or as iron-carriers, anti-coagulants or vaccines adjuvants after derivatization, for instance (De Belder, 2003). They could also present interesting prebiotic properties (Grimoud et al., 2010; Tingirikari et al., 2014).

Considering that dextran applications largely depend on their size, it would be of great interest to develop enzyme-based processes allowing the production of controlled molar mass dextrans in one step from sucrose, and avoiding the use of acids and solvents. For that purpose, a better understanding of the dextransucrase structural determinants dictating polymer elongation appears necessary. In the last decades, sequence analyses and X-Ray 3D-structure characterizations allowed pointing some amino acid residues involved in linkage specificity. However, less is known on the structural determinants controlling glucan molar mass. To date, six 3D-structures of recombinant and truncated GH70 enzymes were solved: GTF180-ΔN from *Lactobacillus reuteri* 180 (Vujicic-Zagar et al., 2010a), GTFA-ΔN from *L. reuteri* 121 (Pijning et al., 2012), GTF-SI from *Streptococcus mutans* (Ito et al., 2011), the ΔN123-GBD-CD2 of DSR-E from *Leuconostoc citreum* NRRL B1299 (Brison et al., 2012) and more recently Gtf-B-ΔNΔV from *L. reuteri* 121 (Bai et al., 2017) as well as DSR-MΔ2 from *L. citreum* NRRL B-1299, the protein synthesizing only LMM dextrans directly from sucrose (Claverie et al., 2017). All these GSs have in common a 5 domain organization with a global U-shape folding. Domains A, B and C are also found in related GH13 enzymes, whereas domains IV and V are unique

to the GH70 family. The recent structural analysis of DSR-MΔ2 in complex with isomaltooligosaccharides, completed by biochemical characterization and mutagenesis studies allowed us to propose a scenario explaining how LMM dextrans are produced, via a distributive mode of synthesis. In particular, the function of specific structural features found in both the catalytic domain A and domain V were highlighted. Notably, residue Trp624, located in the exceptionally long loop B2 of DSR-MΔ2, was shown to be crucial for the elongation of glucooligosaccharides of molar mass higher than 2 Kg/mol and up to 16 Kg/mol. To synthesize dextrans of even higher mass (up to 23 Kg/mol), the assistance of two sugar binding pockets promoting interactions between the growing polymer chain and domain V was found to be required. In these pockets defined by a network of conserved amino acids, Tyr residues were proved to be essential for functionality (Claverie et al., 2017). These sugar binding pockets are widespread in the domains V of GH70 enzymes and their number varies depending on the enzymes (Brison et al., 2016). In DSRM-Δ2, the distributive and LMM dextran synthesizing enzyme, only two of them are functional (Claverie et al., 2017). As the sugar binding pockets play a role in chain extension, we suspect that a high number of functional binding pockets could be related with the ability of synthesizing HMM dextrans by promoting a shift between distributive and processive mode of action. Notably, truncations in the domain V of DSR-S vardel Δ4N from *L. mesenteroides* NRRL B-512F, a HMM synthesizing GS, led to the synthesis of lower molar mass dextrans and to the conclusion that this enzyme adopts a semi-processive mechanism (Moulis et al., 2006b). Similar effects were observed upon truncations of the domain V of GTF180 (Meng et al., 2015a). Notably, the extensive truncations also often induced a drastic loss of catalytic efficiency.

To test our hypothesis on the role of the sugar binding pockets, we have studied herein in details the mode of polymer formation as catalyzed by the dextransucrase DSR-OK recently identified in *Oenococcus kitaharae* DSM 17330, and producing a dextran showing the highest MM reported to date (>10^9 g/mol) with physical properties never observed before (Vuillemin et al., 2017). Deep structural and biochemical characterizations of DSR-OK were undertaken. Sequence analysis enabled the identification of seven putative sugar binding pockets in the domain V and their roles were investigated through different approaches including pocket deletions, mutagenesis targeting the conserved aromatic residues shown to be essential for functionality, and construction of chimeric enzymes.

II. Results & Discussion

II.1. Design and characterization of DSR-OK Δ1, a model of study

Using the disorder prediction algorithm implemented in the RONN program, the N-terminal extremity of DSR-OK sequence (first 83 amino-acids) was predicted to be disordered. Fearing that this disorder could disturb crystallization, this part of the N-terminal domain V was truncated, resulting in a 160 kDa form called DSR-OKΔ1 (residue 84 to 1484, Figure 44A). Heterologous production of DSR-OKΔ1 in *E. coli* BL21 star DE3 was performed in the same conditions as those used with the whole size enzyme (Vuillemin et al., 2017), yielding to equivalent expression level of soluble enzyme at 20,000U/$L_{culture}$. Purification was based on a single step of affinity chromatography (using Strep-tag) followed by gel permeation, resulting in a DSR-OKΔ1 preparation with a specific activity estimated at 90 U/mg versus 120 U/mg for the whole size enzyme purified in similar conditions. Even if the truncation had a non-negligible impact on the enzyme activity, we preferred to use a purified preparation of this truncated form -still keeping a very good activity- for enzyme characterization and crystallization trials.

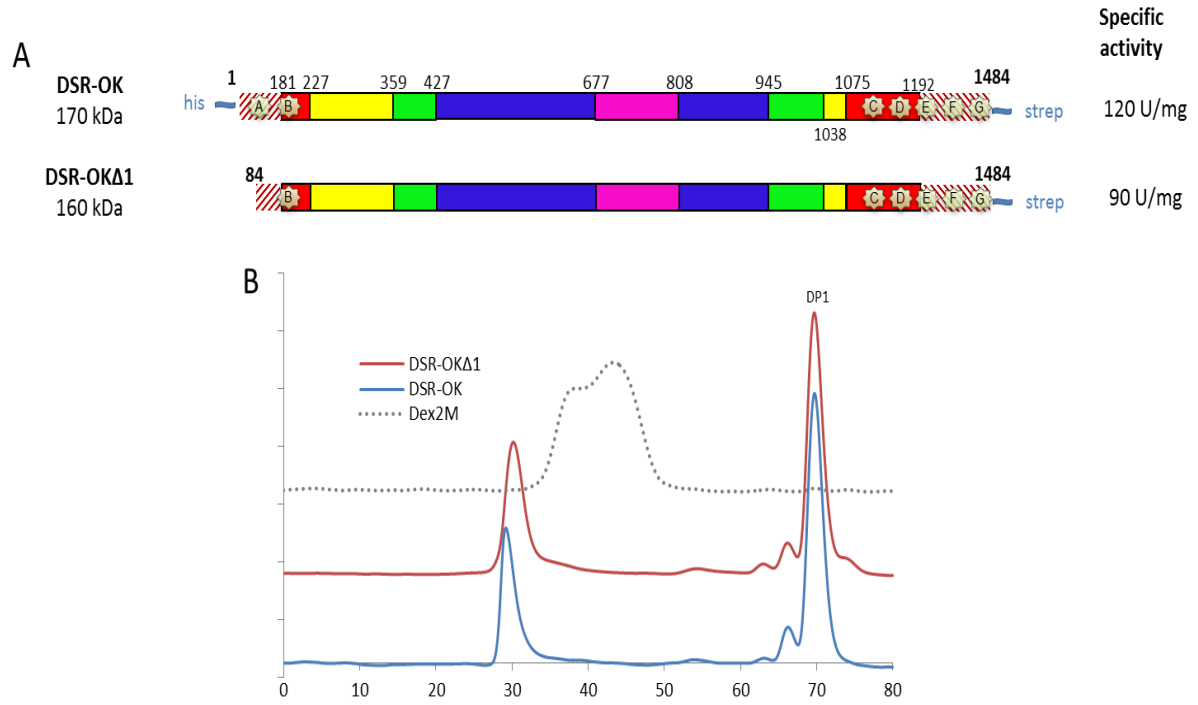


Figure 44: **A. Schematic structural organization of DSR-OK and the truncated variant DSR-OKΔ1, based on amino acid alignment with GTF-180-ΔN and structural analysis**. Hatched red lines correspond to non-aligned zones. Glucansucrases are divided into five structural domains: i) domain V in red, ii) domain IV in yellow and the three core domains: i) domain A in blue, ii) domain B in green and iii) domain C in purple. Putative sugar binding pockets (A, B, C, D, E, F, G) are represented by pale-yellow stars. **B. HPSEC profile of the final medium after depletion of 292 mM sucrose by DSR-OK (blue) and DSR-OKΔ1 (red).** Commercial dextran $2x10^6$ g/mol (Dex2M) (dashed line).

From 100 g/l sucrose, more than 70% of the transferred glucosyl units were incorporated in HMM dextran (Figure 44B), 18% in oligosaccharides of intermediate size, 8.6% in leucrose and only 3.5% in free glucose.

Deletion of the first 83 amino acids slightly impacted the polymer size, as shown by the SEC profiles (Figure 44B). However, both DSR-OK and DSR-OKΔ1 dextrans are HMM polymers, and are higher in size than the largest commercial dextran available as standard (2. 10^6 g/mol). Indeed, the molar mass of DSR-OK dextran as determined by AF4-MALLS was estimated at 1. 10^9 g/mol (Vuillemin et al., 2017). In addition, both polymers display equivalent rheological properties (data not shown). Here, as our main objective was to identify determinants differentiating dextransucrases producing HMM polymers (around 10^8-10^9 g/mol) from those, like DSR-MΔ2, producing only LMM dextran of around $3x10^4$ g/mol, DSR-OKΔ1 was considered as a good model of study.

II.2. Monitoring of dextran synthesis by DSR-OKΔ1

Kinetics of polymer syntheses from 292 mM sucrose using 1 U/ml of DSR-OK Δ1 or DSR-MΔ2 were followed by HPAEC-PAD analyses. Three product profiles obtained from samples taken at 10 min, 30 min and final reaction time are presented in Figure 45.

The product profiles obtained with DSR-OKΔ1 and DSR-MΔ2 are clearly different. After 10 min reaction, both enzymes produce one series of oligosaccharides resulting from the successive transfer of glucosyl units onto sucrose (Figure S15), which is the primer preferentially used as it was previously shown for DSR-MΔ2 (Claverie et al., 2017) and for other polymerases such as GTF-A (Dobruchowska et al., 2013). However, the degree of polymerization (DP) of the oligosaccharides synthesized with DSR-OK does not exceed DP6, whereas oligosaccharides of DP ranging from 3 to 30 are produced with DSR-MΔ2 (Figure 45A). As sucrose consumption is equivalent in both reaction medium, this indicates that DSR-OKΔ1 might retain growing chains which are not released in the medium or cannot be detected, their size being too high with regard to their quantity. Similarly, after 10 min reaction, HMM polymers are not detectable by HPSEC chromatography (See Figure 45D). The difference between the two enzyme modes of action is confirmed after 30 min reaction and after total substrate consumption. Additional series of oligosaccharides are detected, which probably result from transfer onto glucose, fructose or leucrose that accumulate in the reaction medium during the reaction course. Again, these oligosaccharides are much shorter in size for DSR-OKΔ1 than for DSR-MΔ2, and present in lower amount. Notably, HPSEC analyses revealed the presence of one population of very HMM dextran after 30 min reaction, this population being preponderant at the end of the reaction (Figure 45D), without a clear progressive growth of its molar mass. In that respect, DSR-OKΔ1 seems to adopt a semi-processive mode of dextran synthesis similar to that

already proposed for DSR-S-Δ4N (Moulis et al., 2006b). At the beginning of reaction, a non-processive mechanism allows the release of short chain oligosaccharides in the reaction mixture. When they reach a degree of polymerization around DP-6, these oligosaccharides seem to be preferentially elongated to yield HMM dextran, which are probably anchored to the protein before being released in the medium. At the end of the reaction, the oligosaccharides content has increased and represents 18% of the transferred glucosyl units. Their formation may be attributed to the high concentration of fructose, an acceptor becoming, at this stage, a severe competitor of the dextran chains for glucosyl residue capture.

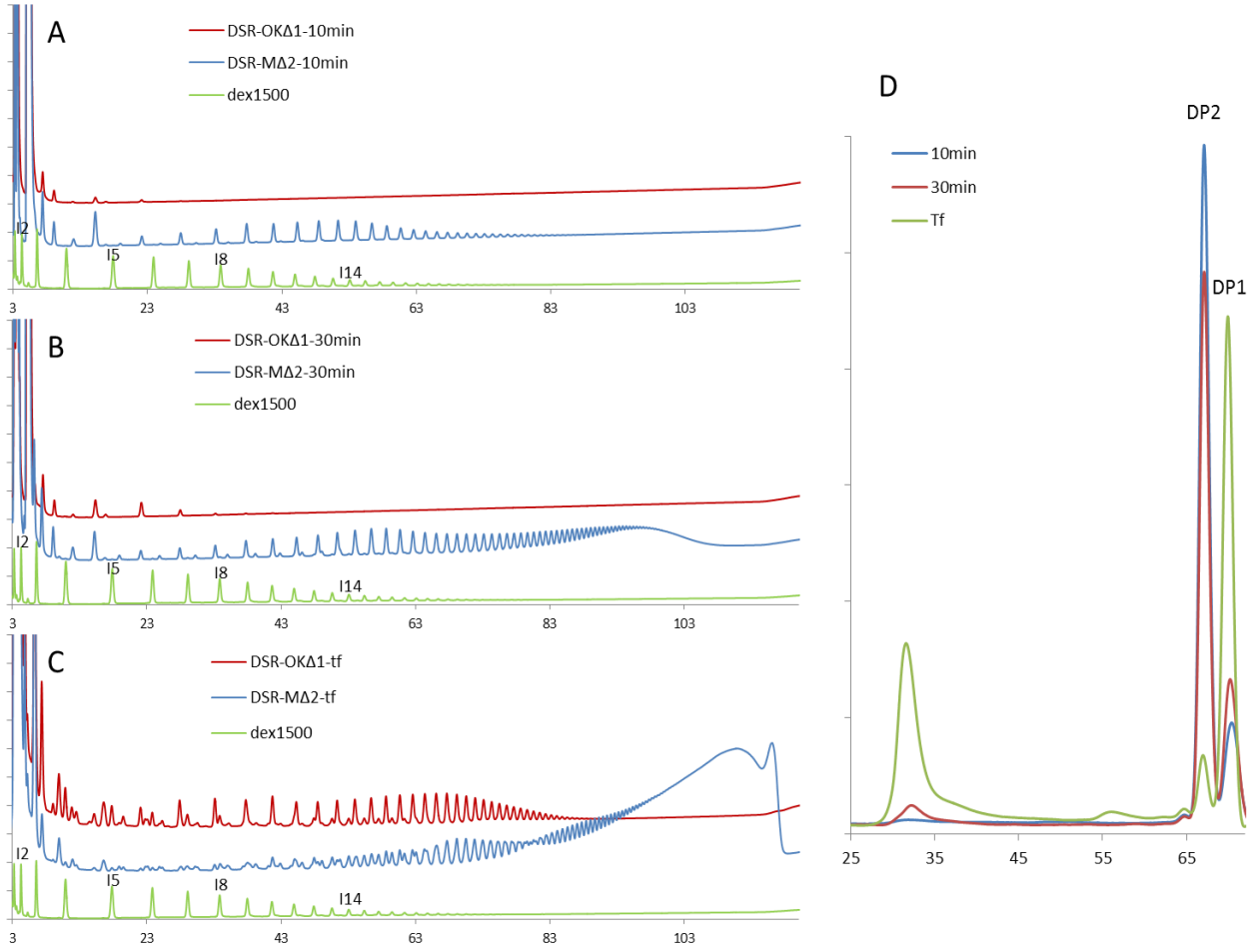


Figure 45: **(ABC) Comparison of HPAEC-PAD product profiles of DSR-MΔ2 (blue) and DSR-OKΔ1 (red) at different reaction times from 292 mM sucrose.** A: after 10 min; B: after 30 min; C: final reaction medium. Commercial dextran 1500 is shown as standard (green). **D HPSEC profiles of DSR-OKΔ1 at different reaction times from 292 mM sucrose.**

In order to identify which structural determinants in DSR-OKΔ1 would be involved in such an efficiency to elongate dextran chains until their final size crystallization trials were attempted to pursue the study at molecular level.

II.3. DSR-OK Δ1 structural insights

Crystallization assays with the entire form were not successful and that performed with the truncated DSR-OKΔ1 only resulted in very small and poorly diffracting crystals. Aiming to improve diffraction quality, construction of an inactive mutant DSR-OKΔ1 E500Q (acid/base catalytic residue targeted), was undertaken to attempt co-crystallization trials with sucrose donor or products. DSR-OKΔ1 (active and inactive form) crystallized only in the presence of I6 (isomaltohexaose), a natural product of the enzyme, which probably stabilizes the protein. After optimization, a first data set was collected at a maximal resolution of around 5 Å. The low resolution did not allow to solve the structure, but we could determine that crystals most likely belong to space group $P2_12_12_1$, with unit-cell parameters a=92.1 Å; b=120.7 Å; c=339.8 Å. Further crystal optimization led to better diffracting crystals, obtained with the inactive mutant and that diffracted at 3 Å resolution. The statistics of collected data are summarized in Table S5. The crystals also belong to space group $P2_12_12_1$, with unit-cell parameters a=68.4 b=101.1 c=240.0. The Matthews coefficient calculation suggests the presence of one molecule in the asymmetric unit, with a VM value of 2.6 $Å^3$ Da^{-1} and a solvent content of 53% (Matthews, 1968). Structure determination using molecular replacement with the reported structure of GTF180-ΔN (PDB entry 3KLL) as a search model did not allow solving the structure completely. It appears that only domain C and part of the catalytic domain A are visible in the electron density map. In contrast, the top of domain A and further domains B, IV and V seem disordered and there is a clear lack of main chain connectivity (R factors after molecular replacement are still very high). In particular, the calcium binding site (observed in all GH70 structures to date) and several key residues for the catalytic site formation (such as Leu370 equivalent to Leu940 of GTF180-ΔN) are absent in the density which might explain the disorganization of the rest of the protein. Part of the domain IV is also visible but seems to be misplaced compared to other GH70 structures (Figure 46). We recently suspected a role of Zn ions present in the crystallization conditions that might cause protein damages, even if they probably favored crystallization. Indeed, activity assays in presence of 1 mM of $ZnCl_2$ induced a drastic loss in enzyme activity (relative activity <1%), maybe due to modification of the enzyme conformation and folding. Seeding assays as well as other parameter variations were also tested but no improvement was observed yet. The poor reproducibility of crystallization conditions was also problematic, rendering these attempts even more difficult. Thus, despite the large number of tested conditions, our efforts to solve the structure of this HMM dextran polymerase remained unsuccessful.

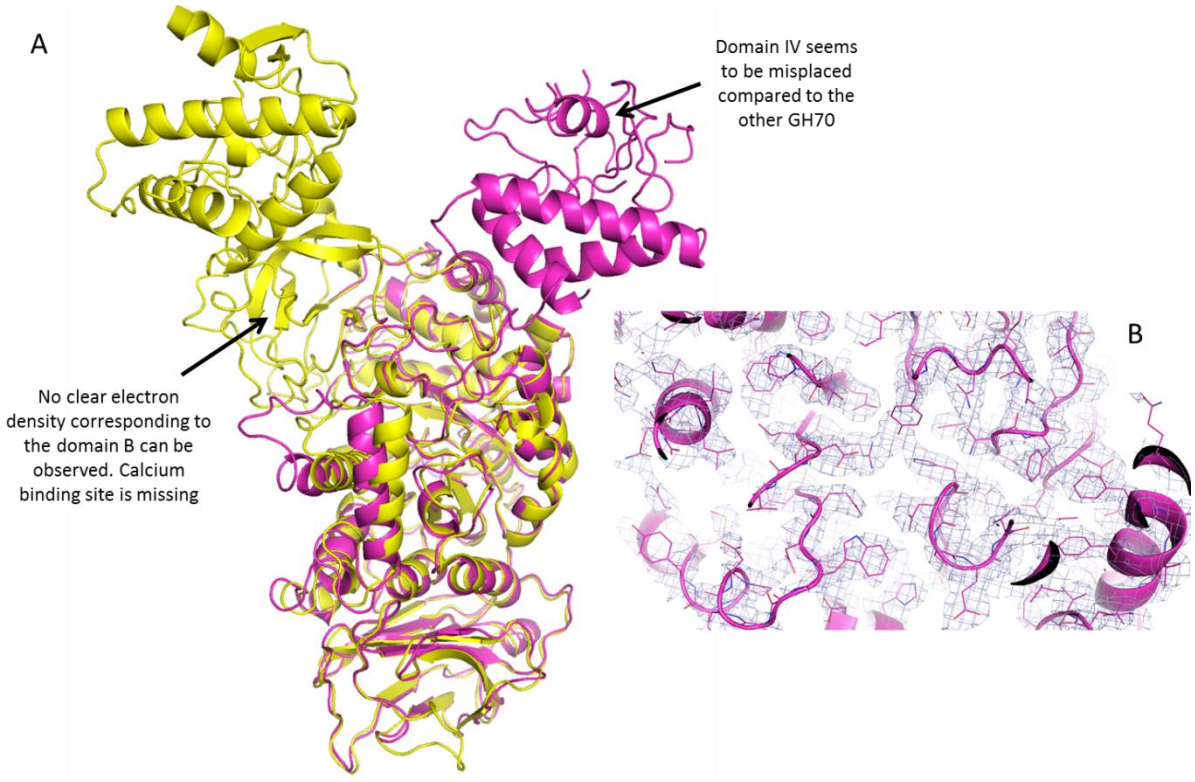


Figure 46: **Comparison of DSR-OKΔ1 structural data obtained so far with the structure of GTF180-ΔN (3klk).** A, yellow: GTF180-ΔN (3klk); magenta: DSR-OKΔ1 structure that resulted from the molecular replacement using the 3 Å data set showing a lack of main chain connectivity. B, Sample of the electron density map in the active site.

To overcome the absence of structural data, SAXS analyses of the protein in solution as well as structure modelling of the enzyme core were performed. SAXS curves revealed that DSR-OKΔ1 is well behaving in solution without aggregation and shows typical modulations of multi-domains proteins. In comparison with DSR-MΔ2, SAXS envelops revealed that DSR-OKΔ1 adopts a much more elongated shape in solution (Figure 47 and Figure 48). This feature is supported by further analyses: i) P(r) functions of DSR-MΔ2 show an almost bell-shape curve which nicely tends to zero at Dmax≈125 Å characteristic of compact particle while for DSR-OKΔ1, we observe an asymmetric curve which tends to zero only at very high Dmax ≈ 255 Å, characteristic of elongated and flexible proteins (Figure 48B); ii) dimensionless Kratky plot of DSR-MΔ2 revealed a well-defined maximum at √3 which is typical of globular, compact particle while for DSR-OKΔ1 the maximum is over √3, indicating extended, flexible particle and the shoulder observable at high qRg values indicates flexibility (Figure 48C). Thus, DSR-OK displays an extended shape (see Rg and Dmax values on Figure 48D) whereas DSR-MΔ2 has a domain V folded back onto the catalytic domain. However the functional relevance of this different positioning remains to be determined.

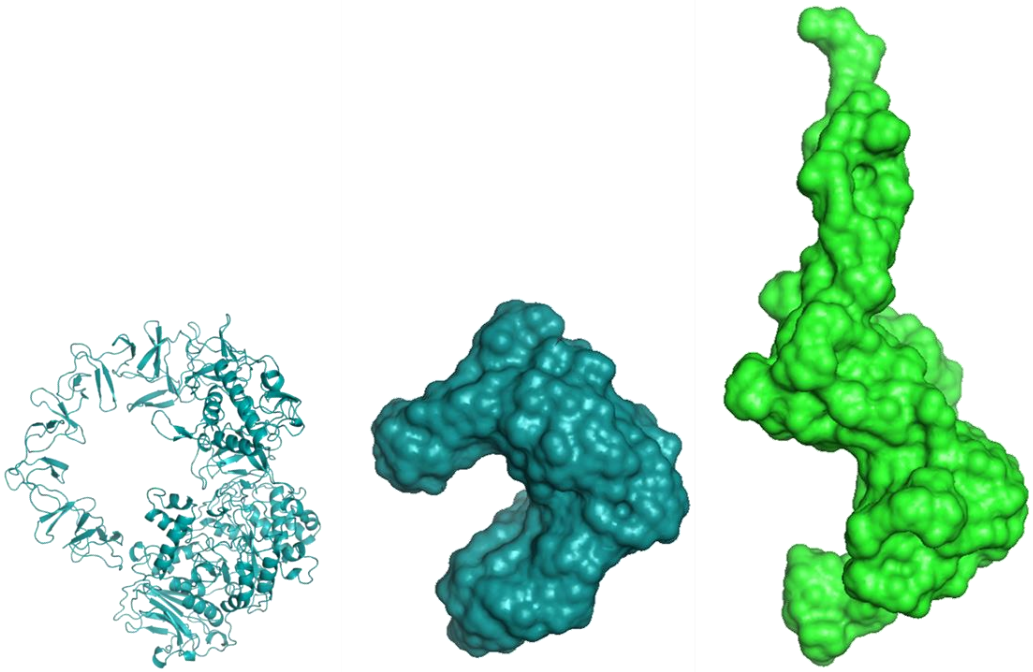
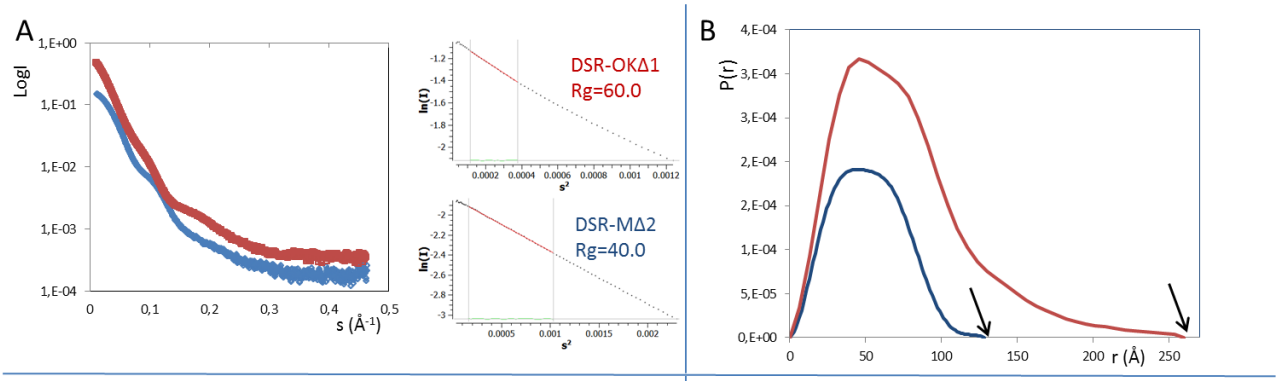


Figure 47: **SAXS analysis of DSR-OKΔ1 in comparison with DSR-MΔ2.** From left to right: DSR-MΔ2 structure; SAXS envelop of DSR-MΔ2; SAXS envelop of DSR-OKΔ1.



	DSR-MΔ2	DSR-OKΔ1
MW (kDa)	144	160
Rg Guinier	40.2	56.6
Rg reciprocal space	40.0	60.0
Rg real space	39.9	60.4
Dmax Å	~ 225	~ 260
Porod Volume (V)	190000	348000
Estimated MW* kDa	~ 118	~ 204

Figure 48: **SAXS analysis of DSR-OKΔ1 in comparison with DSR-MΔ2. A, SAXS curves. B, P(r) functions. C, Dimensionless Kratky Plot. D, SAXS values table.** (blue) DSR-MΔ2; (red) DSR-OKΔ1.

Several attempts to construct a model of the entire enzyme that could fit the SAXS data remained unsuccessful (data not shown). This reflects the difficulties to model the protein domain V without crystallographic data. This is probably due to its flexibility and/or its particular fold. For that reason, only the enzyme core (domains A, B and C) was modelled in order to be compared with the three other available structures of GH70 polymerases. The best model from Modeller, namely model 1 was used as a reference to perform rmsd calculations based on Cα (see Table 7). The low rmsd values (<1Å) indicate that no significant structural changes are observed between the best model and other models, for the aligned regions. But in Table 7, we can notice that not all the Cα were aligned, meaning that some regions are structurally different. For instance, loop A1 conformation varies in each model: models 3 and 5 display a short helix partially comparable with that of DSR-MΔ2 while others are just random coils (Figure 49). As this loop is already very variable between the X-ray structures used as templates, it is difficult to discriminate on what would be the best model. In contrast, there is no difference in the conformation of loop B2 in the five different models. Finally, DSR-OK model 1 was selected for comparison with the other GH70 structures. We first searched for residues that could be equivalent to the amino acids Trp624 (loop B2) and Trp717 (loop connecting β5 strand to α5 helix of the barrel) of DSR-MΔ2, which were shown to be implicated in oligosaccharide elongation (cf. Chapter III). There is no equivalent of Trp624 in DSR-OK despite the fact that loop B2 is quite similar. In contrast, Trp717 of DSR-MΔ2 corresponds to Trp502 of DSR-OK and is conserved in all GH70 enzyme 3D-structures except in the branching sucrase ΔN123-GBD-CD2 and was shown to be crucial for polymer synthesis in GTF180-ΔN (Trp1065). In our model, two aromatic residues (Tyr523 and Trp530) of helix α6 are unique to DSR-OKΔ1 3D-structure (Figure S16). Even if located at more remote distance from the active site than the anchoring patches (Trp624 and Trp717) of DSR-MΔ2, they could be involved in oligosaccharide binding. Upcoming mutagenesis should help to investigate their possible role.

Table 7 : **Rmsd values (PyMol values) of the different DSR-OKΔ1 structural models and GH70 polymerase structures with DSR-OKΔ1 model 1 as a reference.**

	Rmsd on Cα	Number of aligned residues /680
dsrok_model2	0.114	631
dsrok_model3	0.103	620
dsrok_model4	0.114	632
dsrok_model5	0.106	630
3klk/926-1605	0.511	611
3aie/372-1057	0.425	614
4amc/926-1609	0.462	606
5lfc/563-1282	0.573	602

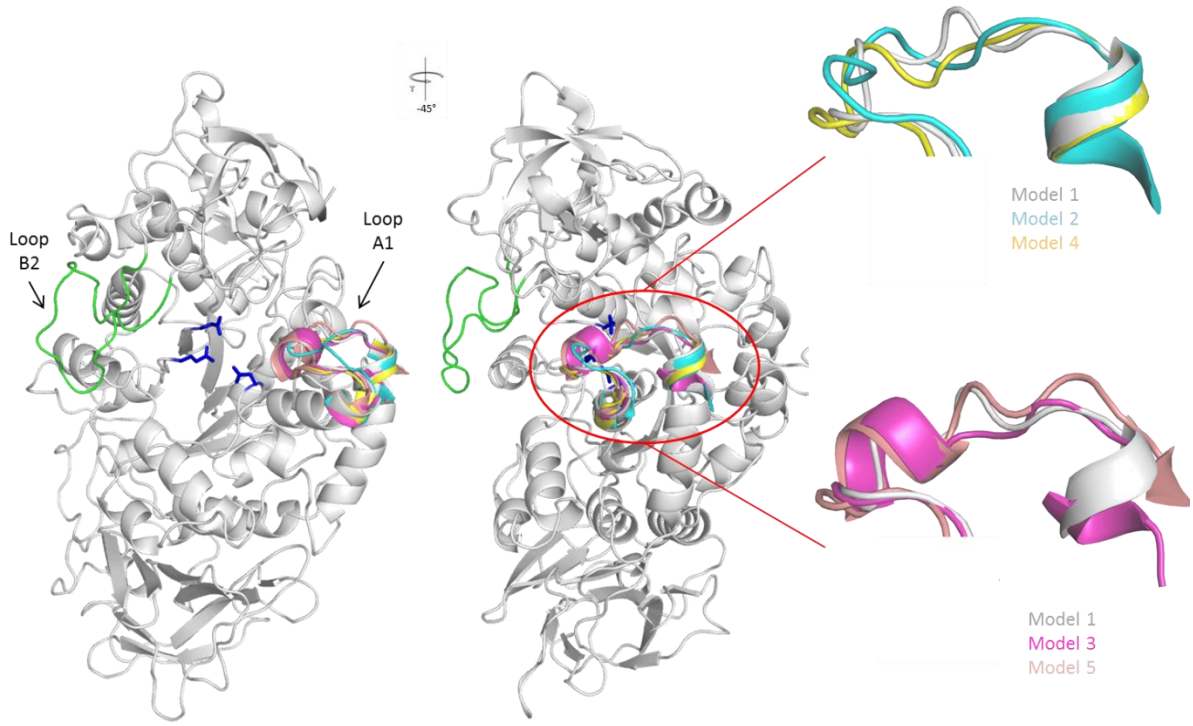


Figure 49: **Comparison of models of DSR-OKΔ1 core (domain A, B and C).** (Left panel) Gray: model 1 with loop B2 shown in green and superposition of loop A1 from the different models. (Right panel) zoom on the loopA1 region of the superposition of the different models.

II.4. Functional implications of domain V

Seven repeats putatively corresponding to sugar binding pockets were identified in DSR-OK sequence by sequence alignment with a set of repeats previously identified in ΔN123-GBD-CD2 and DSR-M (Brison et al., 2016; Claverie et al., 2017), (Figure 44 and Figure 50). Two of them (A, B) are located at the N-terminal extremity of domain V and the others at the C-terminal one. Repeats A, D, E, F and G present both an aromatic residue (tyrosine or phenylalanine) and the QxK motif proposed to be a defining signature of pocket functionality (Brison et al., 2016; Claverie et al., 2017). In repeats B and C, the aromatic residues are conserved but motifs MxK (pocket B) and ExQ (pocket C) replace the QxK motif. Based on the criteria used to assign pocket functionality (Claverie et al., 2017), DSR-OK would possess at least 5 functional sugar binding pockets. To note, the conserved tyrosine Tyr70 of repeat A was truncated in DSR-OKΔ1 (motif QxK was kept) suggesting that only four pockets would be functional in this variant (Figure 44B).

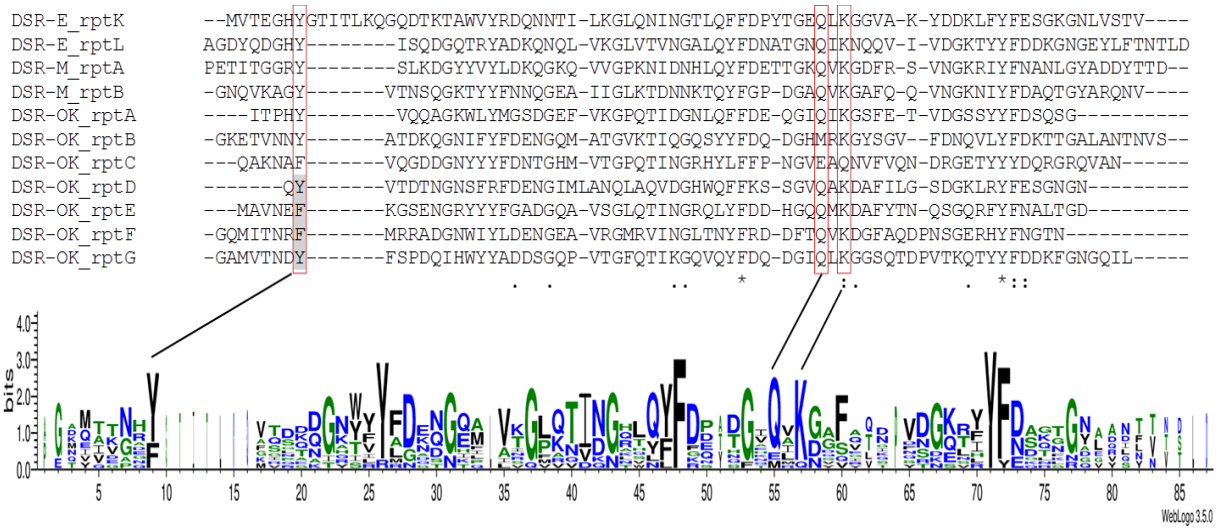


Figure 50: **Repeats of putative sugar binding pockets identified in DSR-OK.** Sequence alignment was performed against 4 repeats identified in domains V of ΔN123-GBD-CD2 (from DSR-E) and DSR-MΔ2. A LOGO sequence based on this alignment is also shown. The common aromatic stacking residue and the QxK motif locations are framed in red. Single point mutations were performed on the grey highlighted residues. DSR-OK repeats (aa numbering): A: 66-126, B: 146-218, C: 1093-1160, D: 1661-1222, E: 1223-1287, F: 1347-1412, G: 1413-1484.

II.5. Effect of sugar binding pocket deletions

Starting from DSR-OKΔ1, several truncated variants were constructed, in which sugar binding pockets were progressively eliminated (Figure 51). All these variants were expressed in quasi-equivalent amounts in soluble fraction of *E. coli* extracts (data not shown). However, none of them were active, except DSR-OKΔ2 that showed a 90% loss of specific activity compared to DSR-OKΔ1. At least 2 sugar binding pockets (ie pockets D and E) are necessary for enzyme activity. These findings are strongly contrasting with those obtained with DSR-MΔ2, for which a complete suppression of the domain V had no impact on specific activity. Activity loss may have several origins but we cannot exclude problems of folding as aggregates were observed during the purification steps of DSR-OKΔ3, DSR-OKΔ4 and DSR-OKΔV, suggesting a lower stability of these variants in solution, even if circular dichroism spectra of pure DSR-OKΔV and pure DSR-MΔV were similar (the active form of DSR-M deleted of its entire domain V (Figure S17). Indeed, some parts of the protein may be folded correctly even if the U shape required for the catalytic core formation is not correctly structured.

HPSEC profiles of Figure 51 further show that DSR-OKΔ2 kept the ability to synthesize HMM dextran of slightly reduced size compared to DSR-OKΔ1 but also produced a minor population of LMM dextran around 13 kg/mol. The synthesis of mix populations comprising both HMM and LMM dextrans was already shown to be catalyzed by variants of the DSR-S from *L. mesenteroides* NRRL B-512F progressively truncated of their C-terminal part (DSR-S vardel Δ3, DSR-S vardel Core, (Moulis et

al., 2006b)). This reveals that all the dextran chains were not elongated to the same extent and could be attributed to the simultaneous suppression of sugar binding pockets F and G. To note, the deletion of pocket A in DSR-OKΔ1 compared to DSR-OK also induced a slight decrease of HMM polymer size (Figure 44B).

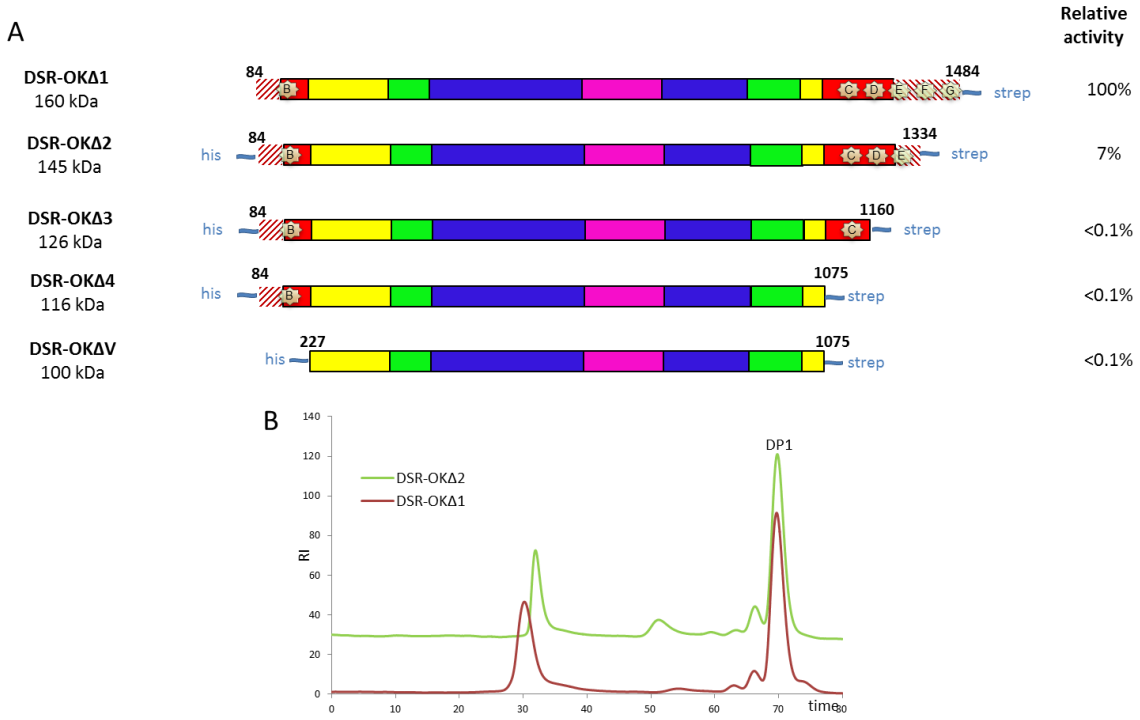


Figure 51: **A. Schematic structural organization of DSR-OKΔ1 and its truncated variants.** Hatched red lines correspond to non-aligned zones. Glucansucrases are divided into five structural domains: i) domain V in red, ii) domain IV in yellow and the three core domains: i) domain A in blue, ii) domain B in green and iii) domain C in purple. Putative sugar binding pockets are represented by pale-yellow stars. **B. HPSEC profile of the product obtained from 292 mM sucrose with DSR-OKΔ1 (red) and DSR-OKΔ2 (green).**

To evaluate the strength of dextran/enzyme interaction, affinity gel electrophoreses of pure DSR-OKΔ1, DSR-OKΔV and DSR-MΔ2 were carried out using gels prepared with either 6000 g/mol Dextran (Dex6000) or 68400 g/mol Dextran (Dex68400). As seen on Figure 52, an important migration delay is observed for DSR-OKΔ1 on the gels containing dextrans, while no delay is seen for DSR-OKΔV, confirming that the affinity of DSR-OKΔ1 for dextran is principally conferred by its domain V. Moreover, the migration delay is higher on the gel containing Dex68400 than on the gel prepared with Dex6000, and this difference is more pronounced than for DSR-MΔ2. Finally, we prepared gels with increasing concentrations of Dex68400 ranging from 0.15% to 0.45 %. We can see that the delay is already maximum at 0.15% concentration. We further decreased Dex68400 concentration down to 0.01% and also observed a maximum delay (data not shown), indicating that the affinity of DSR-OKΔ1 domain V for dextran is much higher than that of DSR-MΔ2. The binding strength of the HMM synthesizing polymerase was so strong that we could not determine any Kd in these conditions.

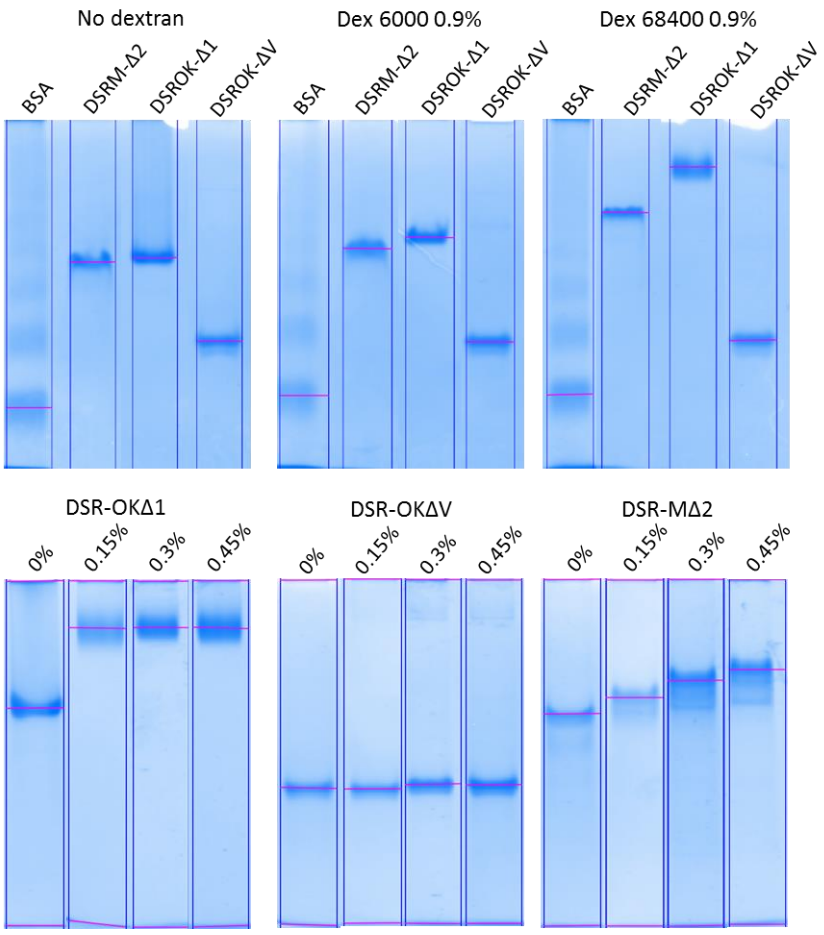


Figure 52: **Affinity gel electrophoresis of DSR-OK variants in gels containing variable amounts of dextrans**. Upper panel, affinity gels with various dextran contents (BSA: negative control, DSR-MΔ2: positive control). Lower panel, affinity gels for which the content in dextran 68.4 kg/mol (Dex68400) is expressed in %(w/v).

To take the analysis one step further, and because truncations of the putative binding repeats might have induced a destabilization of DSR-OK tridimensional structure, several point mutations were undertaken targeting the conserved aromatic residues in pockets D, E, F and G of DSR-OKΔ1 template.

II.6. Effect of mutations targeting the aromatic residues of the sugar binding pockets

All the mutants were similarly expressed and active (Table S6). It was verified by NMR analyses and dextranase digestion that the reaction products were mainly composed of α-1,6 linkages and that the mutations did not impact the linkage specificity (data not shown).

In repeats F and G, the putative stacking residues Phe1354 (repeat F) and Tyr1420 (repeat G) were mutated in alanine with no consequence on the product profile compared to the profile obtained with DSR-OKΔ1 (Figure 53A). The ability to synthesize HMM dextran is maintained and no LMM

products are formed in contrast to what was seen when both the repeats F and G were deleted in DSR-M Δ2. It could be interesting to couple those two mutations in a double mutant to see whether LMM products could be synthesized.

The single mutations of Tyr1162 and Phe1228 in respectively repeats D and E impacted more severely the product profile and the enzyme activity relative to the wild type, which was estimated at 49% for DSR-OKΔ1-Y1162A (pocket D) and 73% for DSR-OKΔ1-F1228A (pocket E). Both single mutants produced LMM dextran in addition to HMM one (Figure 53). The effect was even more pronounced for the double mutants DSR-OKΔ1-Y1162A-F1228A that retained only 20% of activity and almost totally lost its ability to synthesize HMM dextran to the profit of LMM dextran formation.

Kinetic monitoring of dextran synthesis with this double mutant was carried out (Figure 53B Figure 45). Interestingly, HPAEC-PAD profiles are very similar to those obtained with DSR-MΔ2 indicating that the double mutation induced a switch from a processive to a distributive mode of elongation. The two mutated residues are unable to exert a stacking with the glucosyl ring of bound dextran, thus lowering the strength of interaction. This demonstrates that binding to domain V mediated by the succession of sugar binding pockets is the main determinant that governs the "enzyme processivity" an elongation efficiency of HMM dextran synthesizing GSs.

Moreover, the role and functionality of each sugar binding pocket is clearly not equivalent. Indeed, mutations in pockets D and E have higher impact than those in pockets F and G, farther from the active site if we consider an extended structure of the domain V. We can suggest that the closest pockets to the active site are the most important for the correct positioning of the growing dextran chain. In DSR-MΔ2, the mutation of the stacking residue of pocket A had a higher impact than in pocket B. Pocket A is farthest from the active site in the primary sequence but considering the compact form of DSR-MΔ2, this pocket is structurally and spatially closest to the catalytic gorge than pocket B. Hence, the role of each pocket would also be very dependent on the orientation and possible motion of domain V. That is why further identification and characterization of sugar binding pockets in other GH70 enzymes would be of interest in order to deepen our understanding of their role.

Altogether, the ability to form HMM dextrans would be related to i) the sugar binding pockets that would provide anchoring platforms for dextran extension with an avidity increasing with their number ii) the positioning of domain V with regard to the active site.

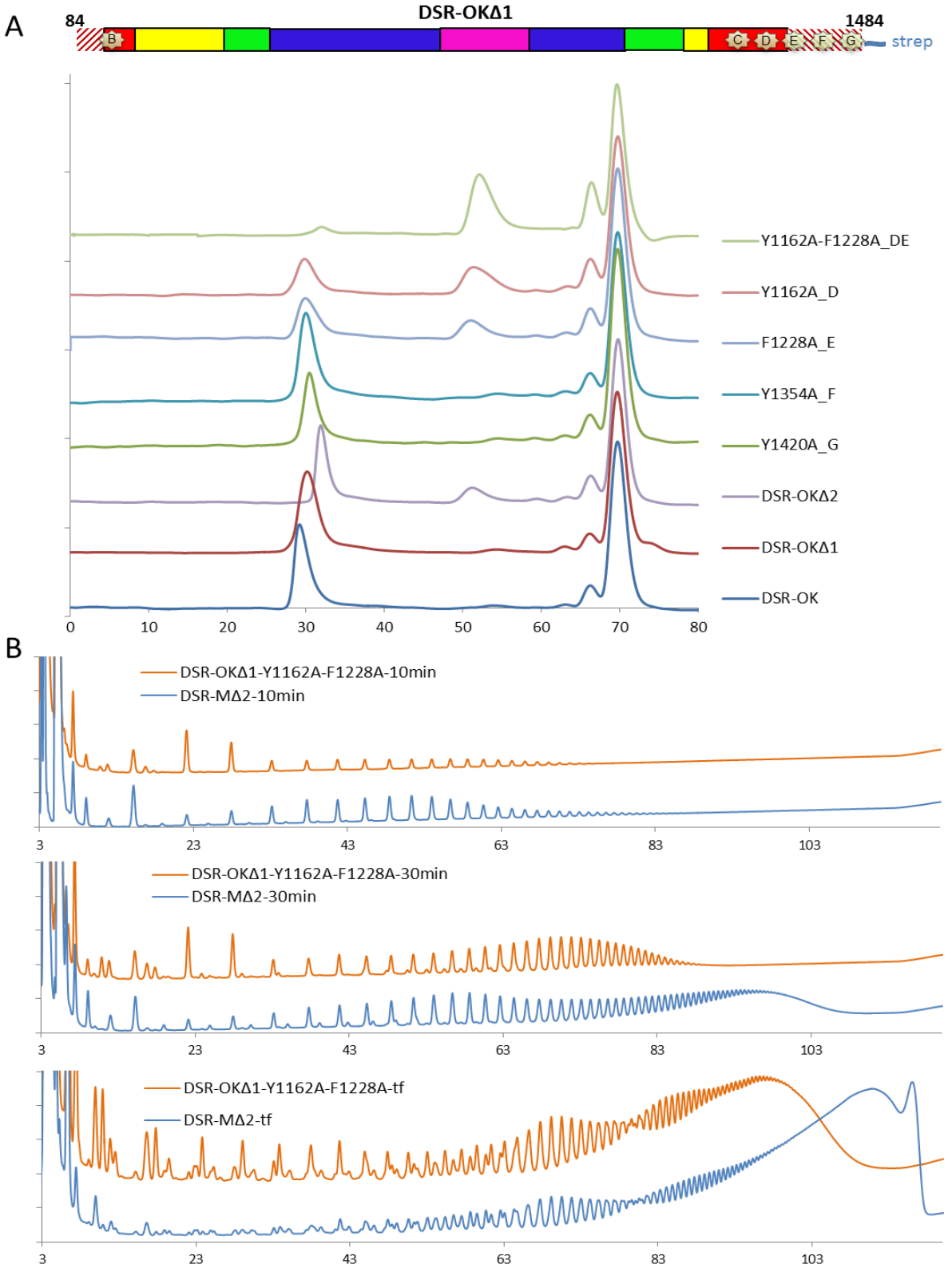


Figure 53: **A. HPSEC analysis of the products synthesized by DSR-OK domain V variants from 292 mM sucrose.** Blue, DSR-OK; red, DSR-OKΔ1; light purple, DSR-OKΔ2; green, DSR-OKΔ1-Y1420A (pocket G); turquoise, DSR-OKΔ1-Y1354A (pocket F); purple, DSR-OKΔ1-F1451A-K1459A (pocket G); orange, DSR-OKΔ1-Y1162A (pocket D); light green, DSR-OKΔ1-Y1162A-F1228A (pockets D et E). **B. Comparison of HPAEC-PAD product profiles of DSR-MΔ2 (blue) and DSR-OKΔ1- Y1162A-F1228A (orange) at different reaction times from 292 mM sucrose.** After 10 min, 30 min and final reaction medium (tf).

II.7. Construction of chimeric enzymes with domain V swapping

Surprisingly, dextrans produced by DSR-OKΔ1-Y1162A-F1228A (absence of stacking in pockets D and E), DSR-S core ΔA (deleted of its entire domain V) (Moulis et al., 2006b) and DSR-MΔV – all lacking functional binding pockets – display a comparable molar mass of around 10-13 kg/mol. Considering that the production of HMM dextrans is governed by the structure of domain V, the construction of chimeric enzymes with inter-exchange of DSR-OKΔ1 and DSR-MΔ2 domains V was undertaken. Swapping of domains IV and V was also performed to limit possible problems of chimera folding.

To determine at which position the domain should be swapped, we used sequence alignments and structural superimposition of the GH70 glucansucrase structures solved to date (Ito et al., 2011; Pijning et al., 2012; Vujicic-Zagar et al., 2010a). We noticed the presence of a helix in domain IV (aa424-aa429, DSR-M numbering), which seems to influence the more or less extended form of domain V and corresponds to the hinge region between domains IV and V previously described for GTF-180ΔN (Pijning et al., 2014). This helix was swapped together with domain V. The most conserved positions at the boarders of each domain were selected as swapping points (Figure S18). In addition, to limit possible problems of chimera folding due to the spatial proximity of domains coming from different enzymes, the concomitant swapping of domains IV and V was also performed. The resulting proteins namely chimeras 1, 2, 3 and 4 are shown in Figure 54.

Chimeras 3 and 4 were produced in the same amount as the parental enzyme, while chimeras 1 and 2 levels of expression and activity were lower than for DSR-OKΔ1 (Figure S19). However, the production yields were sufficient to incubate the enzyme sonicated extracts with 292 mM of sucrose, until total substrate depletion.

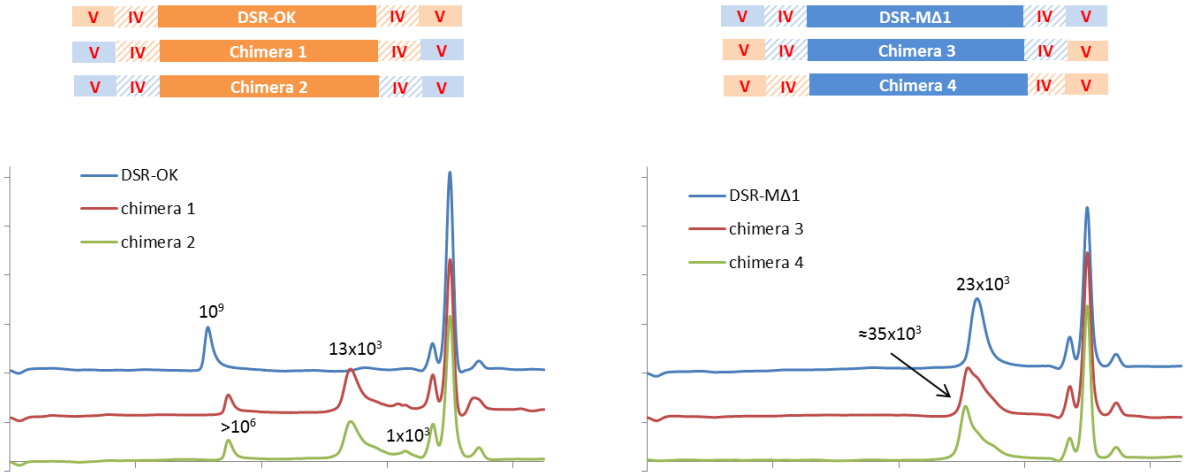


Figure 54: **Schematic representation of the chimeric enzymes and HPSEC profiles of the final reaction medium of the different enzyme forms.** Parts of DSR-OK are represented in orange, parts of DSR-MΔ1 in blue. Molar mass at peak apex are indicated in g/mol.

As shown on HPSEC chromatograms (Figure 54), when DSR-OK domain V (or IV and V) is replaced by that of DSR-MΔ2, resulting enzymes (chimeras 1 and 2) synthesized two main populations of dextrans: one of molar mass superior to 10^6 g/mol (but inferior to that produced by the parental enzyme) and a major population of molar mass once again estimated around 13 kg/mol. The origin of domain IV did not influence the dextran profile. However, the profile of products looks like the previous ones related to DSR-OKΔ2, DSR-OKΔ1 –Y1162A, DSR-OKΔ1 –F1228A and DSR-OKΔ1-Y1162A-F1228A (mutation of stacking residue of pockets D and/or E), with a bimodal dextran population. In fact, the interchange with DSR-M domain V, known to be of low binding efficiency (Claverie et al. 2017) has the same effect on DSR-OK polymerization process as an alteration of its own domain V abilities to anchor growing dextran chains.

Concerning chimeras 3 and 4, resulting from the exchange between domain V (or domains IV + V) of DSR-MΔ2 and DSR-OKΔ1 in a DSR-MΔ2 template, they both produced a population of dextran comprised between 10 and 40 kg/mol with a molar mass at peak apex of around 35 kg/mol. The synthesized products are higher in size than those produced by DSR-MΔ2. However, grafting DSR-OKΔ1 domain V in place of its own domain and introducing two additional sugar binding pockets did not result in the production of a HMM dextran superior to 10^6 g/mol. Again, stability and folding issues also need to be considered and swapping of entire domains at both enzyme extremities is certainly very damaging.

Consequently, although domain V inter-exchange impacted product size in the expected directions, it was not sufficient to completely change the mode of dextran elongation. We have previously shown for DSR-MΔ2 that the interplay between the catalytic domain and domain V governs the size of the polymer. Specific determinants in the catalytic domain of DSR-OK may also participate in the high processivity of this enzyme. In particular, we have pinpointed aromatic residues in helix α6 of DSR-OK that could deserve mutations to investigate their roles. Similarly, Leu370, the equivalent of Leu940 in GTF180 (located in domain B) could also be mutated to evaluate his role and determine if this residue impacts the size of the polymer like in GTF-180.

III. Conclusion

The recombinant enzyme DSR-OK, originated from *Oenococcus kitaharae* DSM17330, was demonstrated to synthesize a HMM and quasi linear dextran. This polymer presents exceptional rheological properties (viscosity, yield stress and a gel like behavior) never observed before for other dextrans. For these reasons, DSR-OK was selected as a model to identify the determinants involved in the control of polymer size. More specifically, we placed the focus on the role of the numerous putative sugar pockets covering the domain V of the protein.

Monitoring of dextran synthesis revealed that DSR-OK adopts a semi-processive mode of elongation comparable with that of DSR-S Δ4N. The first phase of chain initiation producing oligosaccharides is quickly followed by the synthesis of HMM dextran (>10^8 g/mol). This contrasts with the mode of dextran elongation previously described for DSR-MΔ2 enzyme (Claverie et al., 2017), which displays an original distributive mode of elongation responsible for the synthesis of LMM dextran of 30 kg/mol. In another hand, affinity assays revealed that the binding ability of DSR-OK towards glucan chains is high and mainly conferred by its domain V. In contrast, the binding affinity of DSR-MΔ2 domain V was previously shown to be low (Claverie et al., 2017). Such differences revealed that the polymerization mechanism (distributive vs semi-processive) could be determinant for the ability of dextransucrases to synthesize HMM dextran. Our hypothesis was that the domain V and particularly the sugar binding pockets could be an important promotor of the processivity thanks to strong interaction with the growing dextran chain.

By mutating only two stacking aromatic residues of two consecutive binding pockets in the variant DSR-OKΔ1-Y1162A-F1228A, we have switched the semi-processive mode of DSR-OK toward a distributive one, showing the influence of these pockets on the enzyme processivity. Elimination of the sugar binding pockets in DSR-OK (this study) and DSR-SΔ4N (Moulis et al., 2006b) also led to LMM dextran formation, but with a much higher decrease of specific activity. Targeting residues responsible for the functionality of the sugar binding pockets is thus a better mean to modulate glucansucrase dextran size. We have also shown that the number and position of the pocket relatively to the active site is important and that all the binding pockets do not play exactly the same role. Our assumption is that binding pockets spatially closer to the catalytic domain would be more involved in the control of processivity, but this will have to be confirmed by structural data. In addition, the avidity for dextran increases with the number of pockets interacting with the polymer: short oligosaccharides bind with lower affinity than dextran does, and one can imagine a cooperativity between the pockets when the glucan chain is long enough to interact with several binding pockets, a cooperativity that may extend to the catalytic site.

Indeed, as revealed by the behaviors of the chimeric enzymes constructed from DSR-OK and DSR-M in which domains IV and V were interchanged, transformation of a HMM synthesizing enzyme into a one dedicated to LMM dextran synthesis, and vice versa, is not trivial especially without losing too much catalytic efficiency. Activity loss also arises when the domain V becomes less efficient after mutation of binding pockets for instance, and can be attributed to a decrease of apparent affinity for dextran substrate and/or a perturbation of the concerted action between the catalytic domain and domain V. This highlights the fact that HMM synthesizing GSs have optimally evolved to ensure a perfect tuning between the catalytic domain and domain V. Their interplay may be due to continuity between the acceptor binding subsites located in the catalytic domain and the sugar binding pockets. This connection may well exist and some Trp residues whose role deserves to be investigated have been pinpointed. However, to have an image of the position of the dextran during chain extension and start touching at the dynamic of chain extension, efforts to solve the structure of GH70 polymerases in complex with a long dextran must be pursued.

IV. Material and Methods

IV.1. Construction of DSR-OKΔ1

Construction of DSR-OKΔ1 consisting in deleting the first 107 amino acids from the entire form pET-53/DsrOK plasmid DNA template (Vuillemin et al., 2017) including the his tag the was performed following the method described by Wang and Malcolm, 1999 (primers in Table 9). Briefly, this method consists of the use of a two-stage procedure, based on the QuikChange Site-Directed Mutagenesis protocol (Stratagene, La Jolla, CA), in which a pre-PCR, single-primer extension stage before the standard protocol, allows the deletion of a sequence of interest.

IV.2. Protein expression and purification

Proteins cloned in a pET55-DEST vector were expressed and purified as previously described in (Claverie et al., 2017). For the proteins cloned in a pET53-DEST vector, freshly transformed BL21 star (DE3) E. coli cells (Thermo Fisher Scientific) were grown overnight in LB medium supplemented with 100 µg/mL of ampicillin, at 37 °C. Then, 1 L modified ZYM5052 medium (Studier, 2005) (with the following changes: 1% lactose, 1% glycerol and no glucose added) was inoculated from the starter culture at an OD600 nm of 0.05, and incubated at 23 °C under agitation. After 23 h of culture, cells were harvested by centrifugation, re-suspended at a final OD600nm of 200 in purification buffer (PBS 1X, 280 mM NaCl, pH 7.4), and disrupted by sonication. Enzymes were then recovered in soluble supernatant extract, after centrifugation (20,000 rpm, 30 min). Protein purification was performed using the ÄKTAxpress (GE Healthcare) at 12 °C, with a first step consisting in Strep tag affinity

chromatography. Proteins were injected at 1 mL/min into a 5 mL Strep Trap HPTM column (GE Healthcare), pre-equilibrated with purification buffer. After binding, Strep tagged enzymes were eluted at 4 mL/min using a gradient of D-Desthiobiotin (from 0.05 to 2.5 mM) in 20 column volumes. Affinity purification was followed by a size exclusion step on superpose 12 column 16x60 (GE Healthcare) from which the protein preparation was eluted with crystallization buffer (30mM MES pH 6.5, 100mM NaCl, 0.05 g/L $CaCl_2$) or with 50 mM sodium acetate buffer, pH 5.75, 100mM NaCl, 0.05 g/L $CaCl_2$ for biochemical characterization. Protein purity was verified by SDS-Page gel electrophoresis. Protein concentration was determined by spectroscopy at 280 nm on Nanodrop ND-1000 (Thermo Fisher Scientific) and using the protein theoretical molar extinction coefficient and the molecular weight calculated by the ExPASy ProtParam tool (http://web.expasy.org/protparam).

IV.3. Activity assays

Activity was assayed using the dinitrosalicylic method (Miller, 1959). One unit of DSR-M variant is defined as the amount of enzyme that catalyzes the production of 1 µmol of fructose per minute, from 292 mM sucrose, in 50 mM sodium acetate buffer pH 5.75, at 30°C.

IV.4. Enzymatic reaction

All dextran syntheses were carried out on 292 mM sucrose at 30 °C using 1 U/mL of enzyme (pure enzyme preparations) in 50 mM sodium acetate buffer. After total substrate depletion reactions were stopped by 10 min incubation at 95 °C, and samples were stored at -20 °C until further analyses. (For kinetic studies, samples were taken at regular intervals until total sucrose depletion, and reaction was stopped by 10 min incubation at 95 °C).

IV.5. Product characterization

HPAEC-PAD (high performance anion exchange chromatography with pulsed ampero-metric detection) analyses were performed using a CarboPac TM PA100 analytical column (4 mm x 250 mm) coupled with a CarboPacTM PA100 guard (4 mm x 50 mm). Product separation was perofrmed using a sodium acetate gradient (6–500 mM) in 150 mM NaOH over 45 min (1 mL/min).

HPSEC analyses were performed using Shodex OH-Pak SB-802.5 and SB-805 columns (Showa Denko, Minato-ki, Tokyo, Japan) in series coupled with a Shodex OH-Pak SB-G guard column, and placed in a 70 °C oven. The samples were diluted to a maximum of 10 g/kg of total sugars in the eluent (0.45 M $NaNO_3$ and 1% ethylene glycol). Elution was performed at a flow rate of 0.3 mL/min. When possible, the weight-average molar masses of synthesized dextrans were determined using a calibration curve with standards of 10 g/kg of fructose, sucrose and dextrans of 11 300, 39 100, 68 400, $503x10^3$ and $2x10^6$ g/mol (Sigma-Aldrich). To note, dextrans produced by DSR-OK are very HMM polymers and

determining their molar mass necessitates more expansive methods such as AF4-MALS (Vuillemin et al., 2017). Consequently, for dextrans >2x10^6 g/mol HPSEC is used as a comparative method but is limited to discriminate slight differences. For that reason, rheological analyses were also realized to characterize the effect of the various mutations.

IV.6. SAXS measurements and processing

SAXS analysis and envelop calculation were realized as described in (Claverie et al., 2017).

IV.7. Building the DSR-OK core models

Modeller 9.17 (Webb and Sali, 2014) was used to build hundred models of the DSR-OK core (residue from 359 to 1038) from multiple templates, i.e. from residue 372 to 1057 for 3aie (Ito et al., 2011), from residue 926 to 1605 for 3klk (Vujicic-Zagar et al., 2010a), from residue 926 to 1609 for 4amc (Pijning et al., 2012), and from residue 563 to 1282 for 5lfc (Claverie et al., 2017). The residue range given for the templates corresponds to the sequence numbering of the corresponding PDB structure. In addition to the template structures, a multiple alignment (Figure S20) was generated using MUSCLE webserver with default values (http://www.ebi.ac.uk/Tools/msa/muscle/), and given as input to Modeller. Finally, all models were ranked from their DOPE score (Shen and Sali, 2006), and the five best models (lowest scores) were chosen for the comparison analyses.

IV.8. Circular dichroism analyses

Les expériences de dichroïsme circulaire sont réalisées sur un appareil J-815 CD spectrometer (JASCO). Les enzymes sont diluées dans le tampon acétate de sodium 50 mM ou phosphate de sodium 50 mM et les mesures sont réalisées en normalisant via la ligne de base de réponse du solvant utilisé. Les mesures sont prises dans des cuves de 1 mm ou 0.01 mm et un point est réalisé tous les 0.1 nm. Mesures prises tous les 0.1 nm dans des cuves de 1 mm dans l'acétate de sodium pour DSR-OKΔV et dans le phosphate de sodium pour DSR-MΔV.

IV.9. Chimera construction

Two DNA fragments corresponding to gene regions to be swapped were amplified by PCR using the primers and DNA templates detailed in

Table 8 and were then assembled into a single vector using Gibson Assembly (New England Biolabs) (Gibson, 2009). (Primers were designed using the NEB builder server: http://nebuilder.neb.com/). All chimera genes were thus cloned in pET55/DEST vector comprising a Strep and a 6xHis tag at the N- and C-terminal end respectively.

Table 8: **Primers used for chimera constructions**

		Primer name	Primer sequence	template	size of the fragment (bp)
Chimera 1	PCR 1	DSROK-IV Fwd	GTGTTGCTACAAGTCATAATGCCGTTTATTCGACC	pET55-*dsrok*	2538
		DSROK-IV Rev	TATTTGACTCTTGATTCATCAGCTGTTTAGGCAGAAATG		
	PCR 2	pET-55-DSRM-V Fwd	ATGAATCAAGAGTCAAATACTG	pET55-*dsrmΔ1*	6740
		pET-55-DSRM-V Rev	ATGACTTGTAGCAACACTG		
Chimera 2	PCR 1	DSROK Fwd	GACACCAACCCAGGGTGGAATGGGGAAAGTGAAGATCC	pET55-*dsrok*	2079
		DSROK Rev	GTCTTTTAAGACGTAGTTGATCCCACGGCCCTGGATGTT		
	PCR 2	pET55-DSRM-IV-V Fwd	ATCAACTACGTCTTAAAAGACTGG	pET55-*dsrmΔ1*	7231
		pET55-DSRM-IV-V Rev	CCCTGGGTTGGTGTCTAC		
Chimera 3	PCR 1	DSRM-IV Fwd	GATGATTTCACAGCTCATAATGCTGCAAAGTCTTATGATACCAAAAG	pET55-*dsrmΔ1*	2685
		DSRM-IV Rev	GCCGGGCTGATTAGTAAGCTGCTTGGGCAGACG		
	PCR 2	pET55-DSROK-V Fwd	ACTAATCAGCCCGGCGAA	pET55-*dsrok*	7166
		pET55-DSROK-V Rev	ATGAGCTGTGAAATCATCATTTTG		
Chimera 4	PCR 1	DSRM Fwd	AACCAGCAGCCCTCTTGGACGATTGATAGTGAAAC	pET55-*dsrmΔ1*	2194
		DSRM Rev	CGCGCAGCACATAATAAGCTCCTTTTCCCTGAATGTTAG		
	PCR 2	pET55-DSROK-IV-V Fwd	GCTTATTATGTGCTGCGC	pET55-*dsrok*	7625
		pET55-DSROK-IV-V Rev	AGAGGGCTGCTGGTTGAT		

IV.10. Mutagenesis study

IV.10.1. Construction of deletion mutants

For the construction of DSR-OKΔV-Cter and DSR-OKΔV; *dsrok ΔV-Cter*, and *dsrok ΔV* genes were amplified by PCR from pET-53/DsrOK plasmid DNA template (Vuillemin et al., 2017) using the primers described in Table 9. The addition of CACC sequence (underlined) to the 5'-forward primers allowed the correct insertion of genes into the pENTR/D-TOPO® vector (Life Technologies). From a positive

entry clone, LR recombination (Gateway® LR Clonase® II enzyme mix, Life technologies) was performed with pET-55-DEST and pET53-DEST destination vectors (Novagen). Expression clones were selected on LB agar plates supplemented with 100 µg/ml of ampicillin. Plasmids were extracted with Sigma-Aldrich GenElute HP Plasmid Miniprep kit, verified by restriction analyses, and the genes of interest were sequenced (GATC Biotech). *E. coli* TOP10 competent cells (Life Technologies) were used for all cloning experiments. Finally deletion of the sugar binding pockets in DSR-OKΔ1ΔFG and DSR-OKΔ1ΔDEFG were cloned in the pENTR/D-TOPO® vector (Life Technologies) using the In-Fusion® HD Cloning kit from Clontech. As for the previous constructions, LR recombination of positive entry clones were then realized with pET-55-DEST and pET53-DEST destination vectors (Novagen).

IV.10.2. Site-directed mutagenesis

Mutants E500Q (inactive mutant), Y1662A, F1228A, F1354A and Y1420A by inverse PCR (oligo-mediated introduction of site-specific mutations) as described in (Claverie et al., 2017) using pET53-*dsrok-Δ1* plasmid as template and the primers described in Table 9.

Table 9: **Primers used for the construction of the variants of this study (deletion or mutagenesis)**

Primer Name	Nucleotide Sequence	Main features
DSR-OKΔ1for	GGAGATATACCATGGCACATGGCGAGTTTGTTAAGG	Amplification of dsrokD1 gene
DSR-OKΔ1rev	CCTTAACAAACTCGCCATGTGCCATGGTATATCTCC	
E500Qfor	CTATTTTGCAAGACTGGAGCC	Mutation of position 500 - inactive mutant
E500Qrev	CCAGTCTTGCAAAATAGACAGG	
DSR-OKΔVfor	CACCCAAAATGATGATTTCACAG	Deletion of entire domain V
DSR-OKΔVrev	GCCGGGCTGATTAGTCAGC	
DSR-OKΔ4for	CACCATGGCACATGGCGAGTTTG	Deletion of domain V at Cterm end
DSR-OKΔ4rev	GCCGGGCTGATTAGTCAG	
DSR-OKΔ3for	GCCGCCCCCTTCACCATGGCACATGGCGAGTTTGTTAAGG	Deletion of pockets D, E, F and G
DSR-OKΔ3rev	GGCGCGCCCACCCTTATTAGCAACCTGCCGGCCTCTTTGATC	
DSR-OKΔ2for	GCCGCCCCCTTCACCATGGCACATGGCGAGTTTGTTAAGG	Deletion of pockets F and G
DSR-OKΔ2rev	GGCGCGCCCACCCTTTGTTTGAAAACCGGTCACAATCTGCC	
Y1162Afor	CTAATCAGGCTGTGACTGATAC	Mutation of Y1662 in pocket D
Y1162Arev	GTTCGTATCAGTCACAGCCTG	
F1228Afor	GTCAACGAAGCCAAGGGC	Mutation of F1228 in pocket E
F1228Arev	CGCTGCCCTTGGCTTCG	
F1354Afor	TGATCACAAATCGTGCCATGCGC	Mutation of F1354 in pocket F
F1354Arev	CCGGCGCATGGCACGATTTG	
Y1420Afor	GACGCCTTCTCCCCAGAC	Mutation of Y1420 in pocket G

Y1420Arev	GGAGAAGGCGTCATTTGTGAC	

V. Supplementary information

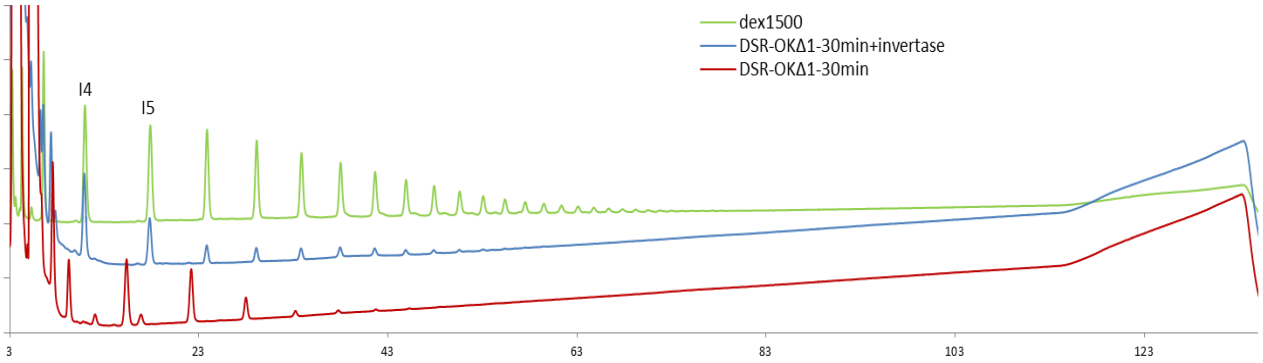


Figure S15: **Invertase digestion. HPAEC-PAD chromatogram of products formed after 30 min of reaction on 292 mM sucrose by DSR-OKΔ1, before (red) and after (blue) digestion with invertase.** Commercial dextran 1500 is shown as standard (green).

Table S5: **Data collection and refinement statistics**

	DSR-OK $ZnCl_2$
Data collection	
Space group	$P2_12_12_1$
Unit-cell parameters (Å)	a=68.40 b=101.06 c= 239.86
No. of molecule/AU	1
Matthews coefficient (Å^3/Da)	2.59
Solvent content (%)	52.55
Wavelength (Å)	0.9724
Resolution range (Å)	50.00-3.20 (3.37-3.20)*
No. of observed reflections	113629 (16327)*
No. of unique reflections	27683 (4006)*
Completeness (%)	98.2 (99.2)*
Multiplicity	4.1 (4.1)*
I/σ(I)	4.1 (1.7)*
R_{merge}	0.143 (0.433)*
CC ½	0.98 (0.87)*
Wilson B (Å^2)	39.4

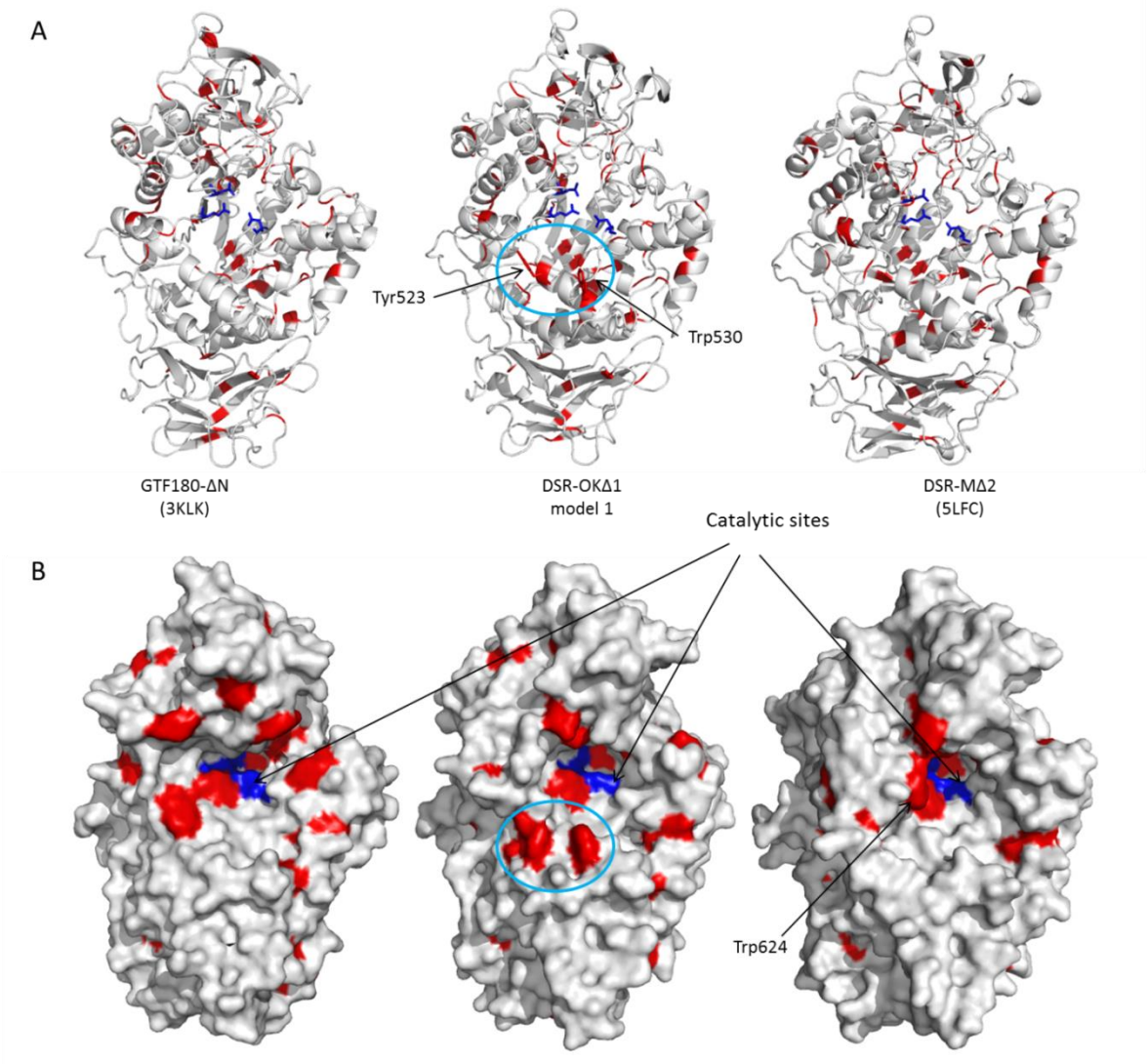


Figure S16: **Identification of aromatic residues in GTF180-ΔN, DSR-OKΔ1 model 1 and DSR-MΔ2.** Aromatic residues are colored in red, catalytic residues are colored in blue. Location of Tyr523 and Trp530 unique to DSR-OK is circles in light blue.

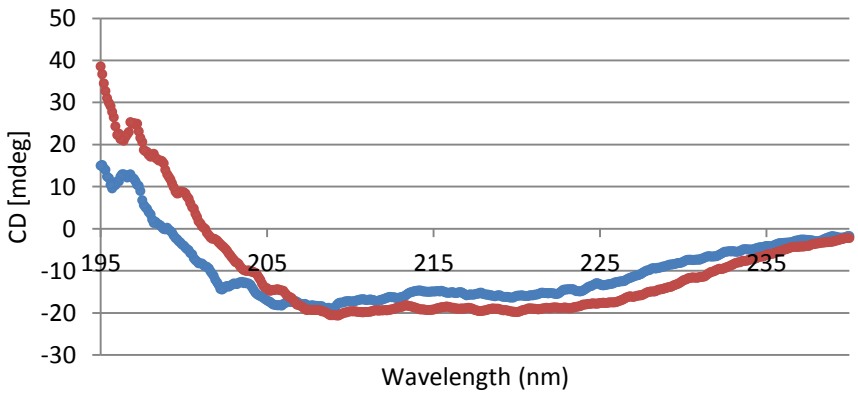
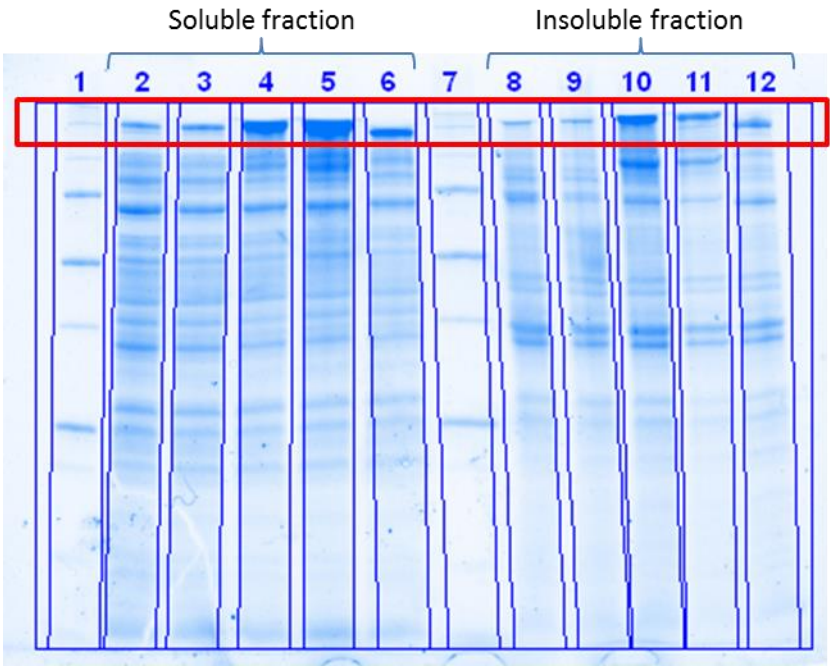


Figure S17: **Circular dichroism profiles of DSR-OKΔV (red) and DSR-MΔV (bleu).**

Table S6: **Expression levels in the soluble fraction of DSR-OKΔ1 variants**

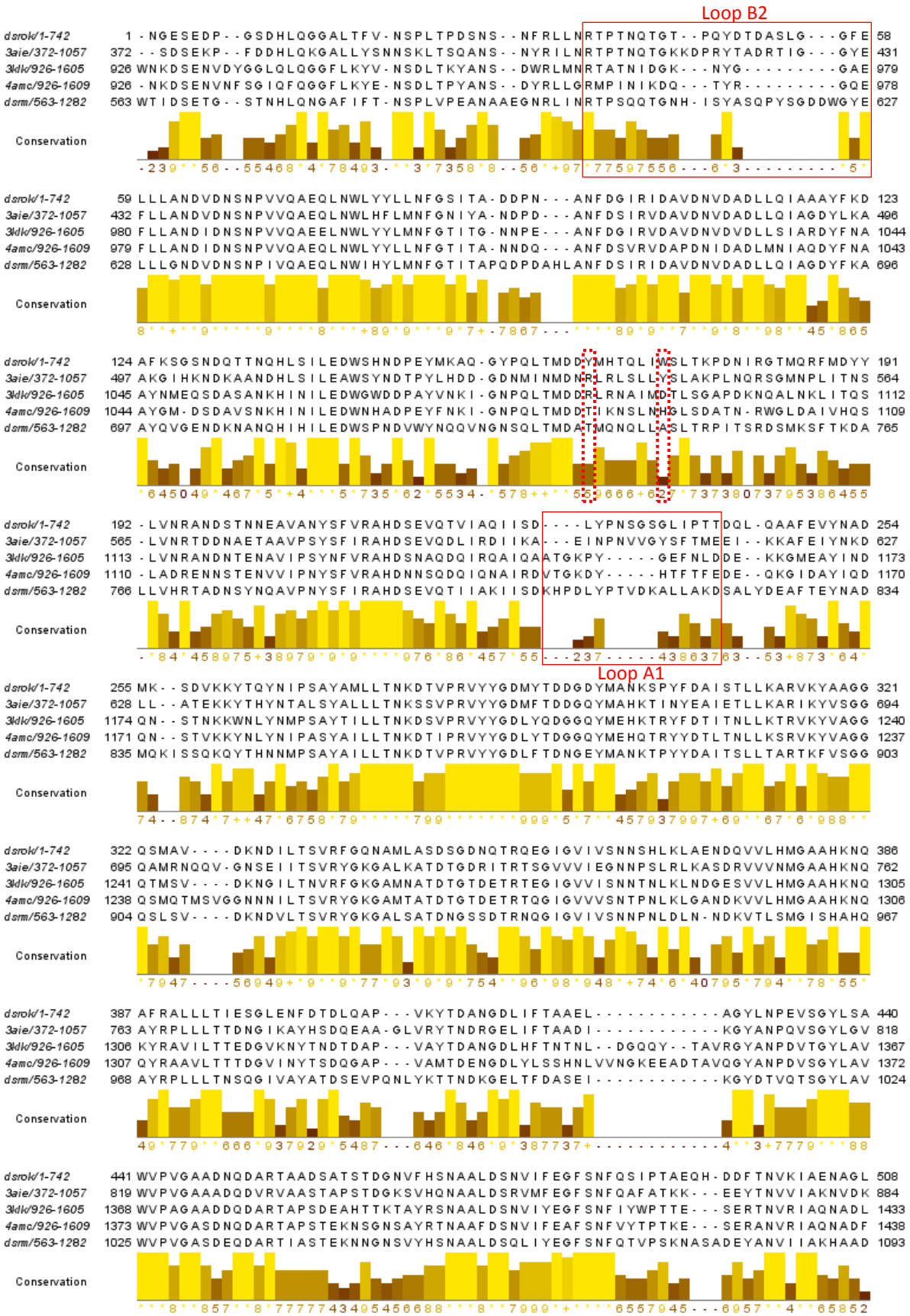
Enzyme	Expression level (U/Lculture)
DSR-OK	18962
DSR-OKΔ1	18709
Y1162A_D	5397
F1228A_E	8392
Y1162A-F1228A_DE	1276
Y1354A_F	13074
Y1420A_G	17750

Figure S18: **Sequence alignment of DSR-OKΔ1 and DSR-MΔ2 with GH70 glucansucrases whose structures have been solved (3KLK: GTF180-DN; 4AMC: GTFA-DN; 3AIE: GTF-SI).** Blue boxes correspond to the helix at the beginning of domain IV and red lines indicate the swapping areas. In red, domain V; in green, domain B; in yellow, domain IV.



Enzyme	$U/L_{culture}$
Chimera1	394
Chimera2	325
Chimera3	16175
Chimera4	11002
DSR-MΔ2	13250

Figure S19: **Expression levels of chimeric enzymes in comparison with DSR-MΔ2.** Lanes 1 and 7: PrecisionPlus Biorad standard, Lanes 2-6: soluble fractions, Lanes 8-12: insoluble fractions. Lanes 2 and 8: chimera 1, lanes 3 and 9: chimera 2, lanes 4 and 10: chimera 3, lanes 5 and 11: chimera 4, lanes 6 and 12: DSR-MΔ2.


Loop B2
Loop A1
dsrok/1-742
3aie/372-1057
3klk/926-1605
4amc/926-1609
dsrm/563-1282
Conservation

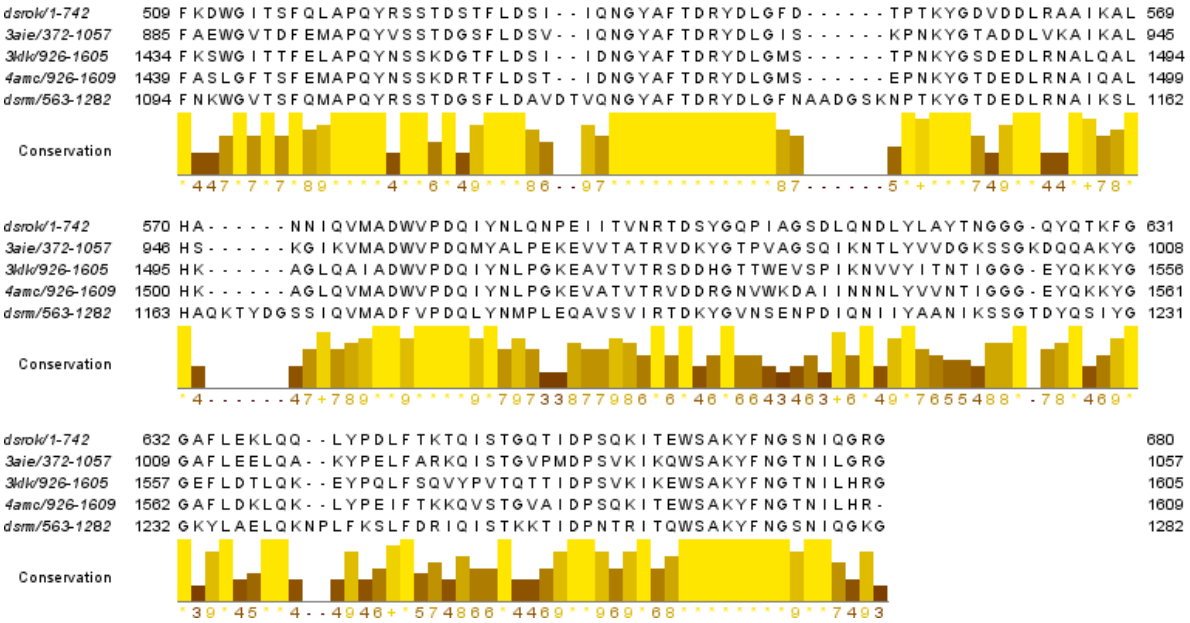


Figure S20: **Sequence alignment of DSR-OK core (domains A, B and C) with that of other GH70 polymerase structures.** Areas delineating loops B2 and A1 as defined in (Claverie et al., 2017) are framed in red, Tyr523 and Trp530 unique to DSR-OK are framed with red dashed lines. Alignment was realized using MUSCLE with default parameters. (1-742 of dsrok in the alignment corresponds to 359-1038 in real numbering).

Conclusion
&
Prospects

All along this thesis, we have seen that α-transglucosylases of the GH70 family are attractive tools for the synthesis of glucopolymers or glucoderivatives of industrial interest. They are very efficient transglycosylases, most of them using sucrose, an abundant and renewable agricultural resource, to synthesize a large spectrum of products. In our work, the focus was placed on one type of GH70 enzymes named dextransucrases and catalyzing the formation of α-glucans with a majority of α-1,6 linkages from sucrose.

These enzymes have been studied since the fifties. The literature counts number of publications related to their discovery, producing strains, biochemical characterization and applications. During the last decade, major progresses have been made in the understanding of their mechanism at the molecular level thanks to their structural characterizations. New subfamilies have been discovered and the evolutionary scheme linking GH70 to GH13 family is starting to be uncovered. However, despite all these available data, the polymerization mechanism adopted by dextransucrases as well as the dynamic of polymer elongation is not fully understood and elucidated. In particular, the identification of the molecular determinants governing the mode of polymer synthesis and by extension its size still requires investigations. Indeed, among dextransucrases, some synthesize only high molar mass (HMM) dextran polymers directly from sucrose substrate, others produce a mixed population of high and low molar mass (LMM) products, and one enzyme has recently been described as producing only LMM dextrans.

As most of the dextran applications concern products of controlled molar mass obtained after several chemical steps, identifying the molecular determinants involved in the tight control of dextran size would represent a considerable advantage from a biotechnological and applicative point of view to provide tailor-made dextransucrases synthesizing in one step dextran of controlled molar mass.

In this context, the objective of this thesis was to decipher the elongation mechanism adopted by two dextransucrases synthesizing polymers of either LMM or HMM. The following strategy was adopted:

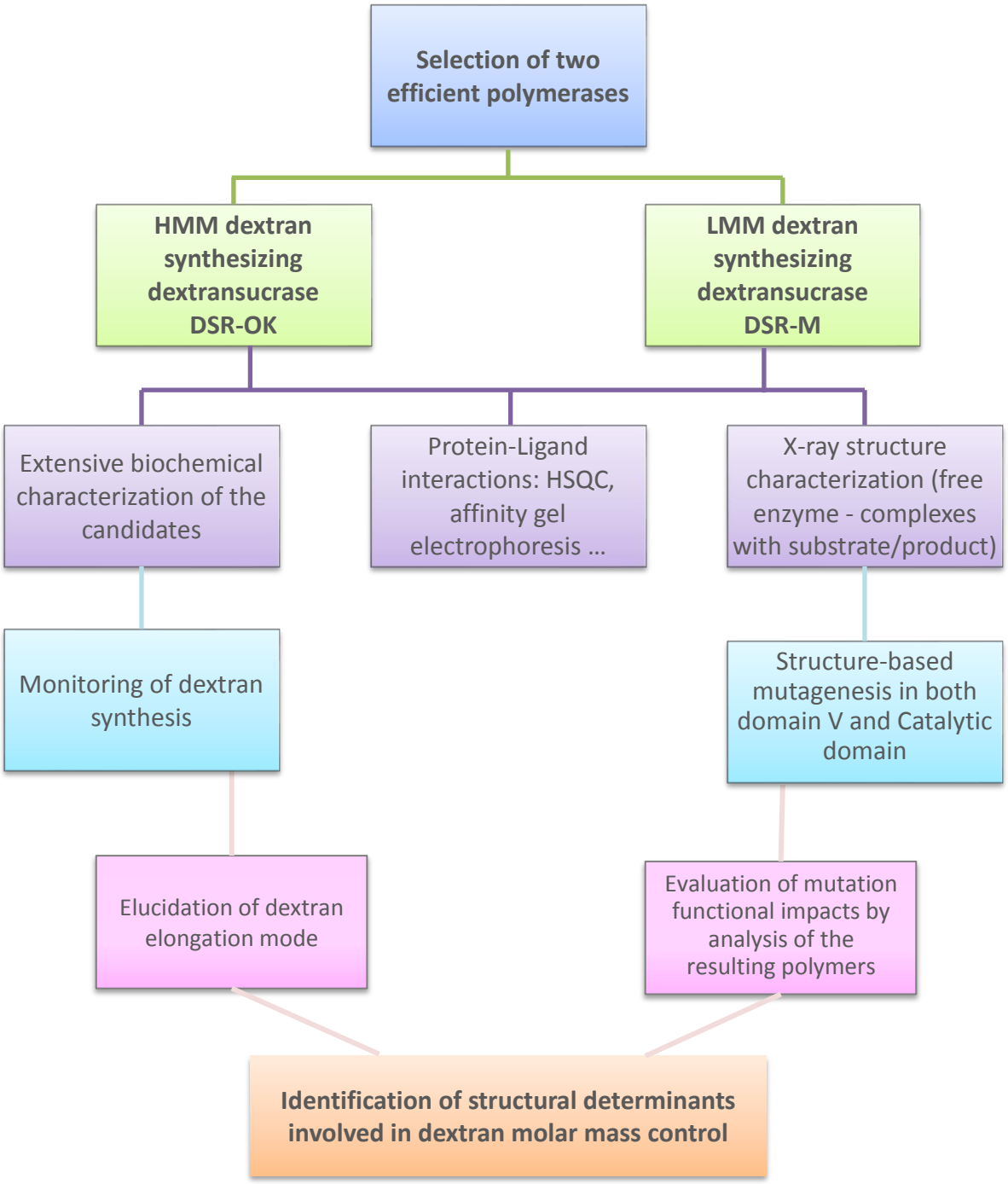


Figure 55: **Strategy of the current study**

We have hence selected two GH70 dextransucrases recently discovered in our group: DSR-M from *L. citreum* NRRL B-1299 that naturally synthesizes only LMM dextran and DSR-OK from *O. kitaharae* DSM17330 that produces a HMM polymer. The main results that we have obtained using the above strategy as well as perspectives for further investigations will be presented in what follows.

DSR-M: short-chain polymerase from *Leuconostoc citreum* NRRL B-1299

The full inventory of the GH70 enzymes encoded by this LAB strain allowed the identification of three novel α-transglucosylases in addition to those previously described. Among them, DSR-M was shown to be a very original dextransucrase as it is the first enzyme of the family described to naturally, exclusively and very efficiently produces only LMM linear dextrans (~30 kg/mol) from sucrose (Passerini et al., 2015). This made it a very interesting candidate to address the problematic of our study and identify determinants that could explain why this enzyme is dedicated to only LMM dextran synthesis while others produce HMM polymers (>10^8 g/mol).

As sequence analysis did not reveal any obvious discriminating features that could explain this specificity, the structural characterization of this enzyme was undertaken. To improve crystallization and further crystal diffraction, we had to construct shorter forms of the protein. With care to keep a maximum of the native protein integrity, only the predicted as disordered region at the N-terminal extremity and the C-terminal APY motifs (known to not interfere in the enzyme mechanism) were truncated in the DSR-MΔ2 variant. This form, displaying the same biochemical behavior as the parental enzyme, was also used to monitor dextran synthesis. During the first phase of the reaction, sucrose was shown to be the preferred initiator of dextran synthesis yielding a series of isomaltooligosyl sucrose. The sucrose consumption rate was initially low but increased and remained constant once the oligosaccharides produced had reached a certain DP indicating that, above this DP, DSR-M elongates oligodextrans irrespective of their size. Added to the fact that the size of the polymer produced by DSR-MΔ2 gradually increases with time, all these findings show that this polymerase adopts a chain elongation mode typical of distributive enzymes.

In parallel, we solved four X-ray structures of DSR-MΔ2 including the free enzyme, a sucrose complex and two isomaltotetraose (I4) complexes, for which the sugar was located either in the catalytic domain or in the domain V. The global structure of DSR-MΔ2 is organized in five domains as previously described for other GH70 enzymes. The originality of this enzyme comes from its compactness with its domain V folded back to the catalytic domain. The enzyme adopts a kind of "horseshoe" shape both in crystals and in solution (SAXS analysis). In one of the solved complexes with I4 (PDB entry: 5NGY), the sugar molecule found in the domain V is rather close to the catalytic domain, suggesting that these two domains may interplay during dextran synthesis.

This structure displays the largest part of a domain V reported so far for a GH70 enzyme, and allowed the first biochemical characterization of sugar binding pockets presumably involved in glucan binding during polymer synthesis in a dextransucrase (Figure 56). Three out of five putative pockets were shown to play a role in the product size specificity of DSR-M. These pockets, also described in ΔN123-

GBD-CD2 branching sucrase (Brison et al., 2016) display at their bottom a conserved aromatic residue (Tyr or Phe) that we proved to be critical for their functionality. We also observed that DSR-MΔ2 domain V displays a rather low binding affinity towards the growing dextran chains, which may participate in its incapacity to synthesize larger polymers. Thus, DSR-MΔ2 domain V could be considered as a "weak" glucan binding domain resulting in the synthesis of exclusively LMM dextrans. This is also supported by the fact that the enzyme devoid of its entire domain V kept its polymerase efficiency, with only a moderate decrease of the polymer size. One can imagine that the binding ability of the domain V is probably linked to the sugar binding pockets. Their location, their number and their binding efficiency might be influencing parameters.

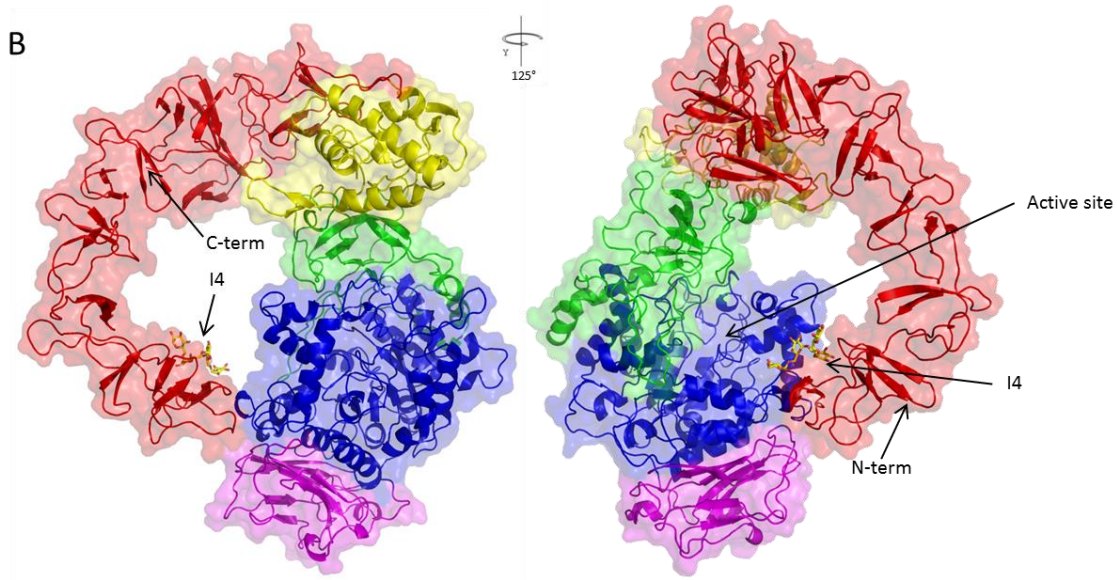


Figure 56: **Structure of DSR-MΔ2 in complex with I4 in the domain V**

Structural analyses also revealed other features that differentiate DSR-MΔ2 3D-structure from the other GH70 structures solved to date. Notably, specific loops around the catalytic site were identified, namely loops A1 and B2. Loop A1 presents the particularity to harbor a little helix inserted between the two helices of the H1-H2 subdomain. By comparing the sucrose-complex structure with that of the I4-complex, we have observed a motion of this whole subdomain resulting in either a close (sucrose complex) or an open (I4-complex) conformation of the catalytic cleft. This motion likely helps the accommodation or the release of both the substrate and product molecules. Another original feature of DSR-MΔ2 is its long loop B2. By solving the first 3D structure of a GH70 glucansucrase in complex with a transitory natural product (I4) bound in the active site, we observed that the Trp624 residue was interacting with the glucosyl unit in subsite +3. This residue, at the foot of loop B2, participates in the glucan binding through stacking interactions. Replacement of this residue by an alanine resulted in a drastic decrease of the polymer size. Further NMR analyses revealed its role in anchoring the whole loop to the core of the protein, allowing interaction with the

growing dextran chain. Finally, the generated mutant is quite interesting for potential applications. Indeed, it produces from sucrose only a very linear short dextran of only 2.4 kg/mol. Moreover, by adding glucose in the media, this mutant efficiently synthesizes controlled MM oligosaccharides of DP ranging from 1 to 12 that could be of interest for prebiotic uses for example.

Altogether, we established that i) Trp624 of loop B2 is crucial for the elongation of glucooligosaccharides of molar mass higher than 2 kg/mol and up to 16 kg/mol and that ii) assistance of two sugar binding pockets promoting interactions between the growing polymer chain and the domain V is required to synthesize dextrans of even higher mass (up to 23 kg/mol). However, DSR-M domain V presents a global low affinity toward the growing chain – probably linked to the sugar binding pocket number, location and functionality – that prevents the synthesis of higher MM polymers.

DSR-M product size specificity is then governed by structural determinants located in both the domain V and the catalytic domain of the enzyme. Based on these findings, we pursued our investigations on the second candidate of this study, trying to establish correlations or differences that could explain what confers the ability of DSR-OK to synthesize such HMM polymers.

DSR-OK: very long-chain polymerase from *Oenococcus kitaharae* DSM 17330

Mining of the *Oenococcus kitaharae* DSM17330 genome allowed the recent identification of the *dsr-ok* gene. The resulting recombinant protein is the first characterized GH70 enzyme from this LAB genus. In contrast with DSR-M, DSR-OK enzyme produces the dextran with the highest molar mass described to date ($>10^9$ g/mol), which presents singular rheological properties (Vuillemin et al., 2017). To identify the structural determinants responsible for such polymerization efficiency, we have also undertaken the structural characterization of this candidate. Unfortunately, our efforts to solve its crystallographic structure remain unsuccessful at this time.

A slightly shorter form of the protein, namely DSR-OKΔ1, was constructed for which the predicted as disordered N-terminal extremity was deleted. This form was used both for crystallization trials and for dextran synthesis monitoring. Now, the profile of products during kinetic follow-up with DSR-MΔ2 and DSR-OKΔ1 are totally different. For DSR-OKΔ1, the absence of accumulation of oligosaccharides/LMM dextran and the rapid detection of HMM polymers (since 30 min of reaction) revealed that DSR-OKΔ1 adopts a semi-processive mode of elongation comparable with that of DSR-S Δ4N , another dextransucrase also specialized in HMM dextran synthesis (Moulis et al., 2006b).

The domain V of DSR-S vardel Δ4N was previously reported to play a major role in dextran chain anchoring during the elongation process, as its partial or total truncation resulted in a drastic size

reduction for the dextrans produced by the mutants. It was also shown that DSR-S domain V had a strong glucan binding ability (Suwannarangsee et al., 2007), contrary to what we observed for DSR-MΔ2. However, at that time, sugar binding pockets were not described and only large domain truncations were performed. In the absence of structural data for DSR-OK, we decided to first focus our investigations on its domain V and in particular on its C-terminal part.

Based on the characterization of sugar binding pockets described in the domains V of DSR-MΔ2 and ΔN123-GBD-CD2, seven putative binding pockets were identified in DSR-OK domain V. Among them, five were localized in the C-terminal part of the domain (pockets C, D, E, F, G). Several mutants deleted of some pockets or of the entire domain V were constructed and characterized. Now, and contrary to what we observed for DSR-M, deletions induced a high impact on the enzyme specific activity and were even lethal when too large. However, affinity assays on both DSR-OKΔ1 and DSR-OKΔV (deleted of its entire domain V) revealed that the binding ability of DSR-OK towards glucan chains was very high and mainly conferred by its domain V. Now, as pockets D, E, F and G comprised the conserved residues proposed to confer binding functionality, namely the aromatic residue and the QxK motif, mutagenesis of these pockets was undertaken. Variants in which the aromatic stacking residue of pockets D or E was mutated were the most impacted, as they both synthesize a double population of dextran, one being of LMM (around 10 kg/mol). The combination of these two mutations had an even higher effect, resulting in a variant that almost completely lost its ability to produce HMM polymers in favor of the lower mass population. This highlights the essential role of these two residues on the enzyme size specificity. But above all, this mutant was shown to adopt a distributive mode of elongation close to that of DSR-MΔ2. These results confirm that the "processivity" of such efficient HMM dextransucrases would be only conferred by some of the sugar binding pockets present in their domain V, especially thanks to some aromatic residues indispensable for functionality.

Discussion and perspectives

Although important efforts have been made with both candidates, only the structure of DSR-MΔ2 was solved. However, it allowed us to identify sugar binding pockets – already described in the branching sucrase ΔN123-GBD-CD2 by Brison et al. – in the domain V of this dextransucrase and to start the investigations on their role. Such putative binding pockets were also predicted by sequence alignment in the DSR-OK domain V. For both candidates, the aromatic stacking residue localized at the bottom of these pockets was shown to play a crucial role on the size of the polymers. As the affinity towards glucan was higher for DSR-OK than for DSR-M domain V, one can imagine that DSR-OK pockets might have better binding efficiency than the ones of DSR-M. We can envisage that the

unconserved residues composing these pockets may also participate to the affinity and that subtle differences may induce significant binding strength differences. Another hypothesis is that their positioning relative to the active site is different, thus highlighting the fact that a co-evolution of both the catalytic site, the domain V and also the domain IV (shown to play a role in the ability of domain V to move toward the catalytic domain in GTF-180ΔN) influence their efficiency by allowing a productive positioning of the dextran chain for elongation.

While the sequence of pockets D, E, F and G of DSR-OK all comprise what we have (at least) identified as a pre-requisite for binding (presence of a stacking aromatic residue together with the QxK motif), mutations of the aromatic residues in pockets F and G did not significantly impact the polymer size. However, in comparison with the parental enzyme, “elimination” of these pockets induced the apparition of a LMM dextran population as well as a decrease of the molar mass of the HMM dextran one. Moreover, additional investigations on the QxK motif and on the other residues composing these pockets are still needed. Thus, the link between the number, the sequence, the binding strength, the localization and the eventual synergetic action of these pockets still remains to be elucidated and would deserve further investigations.

Analysis of the direct interaction between glucans and specific sugar binding pockets may bring some answers. In that goal, the construction of domain V isolated from the rest of the protein has been initiated but as we have been confronted to solubility issues, improvement of the expression conditions will be required. We have then pursued protein-glucan interaction characterization directly with the entire enzyme forms. Affinity gel electrophoresis allowed us to obtain qualitative information on the enzyme binding strength. Unfortunately quantitative information could not be obtained, although different methods were tested: thermophoresis (Nanotemper), surface plasmon resonance (Biacore), isothermal calorimetry … Several difficulties were encountered with each method: requirement of too high concentrations of pure proteins and ligands, that are often commercially non available (ITC), problem with ligand response (Nanotemper), effect of the immobilization on the protein activity and presence of dextran on the commercially available chips (SPR). Thus, efforts need to be pursued to either i) improve or find more adapted methods for protein-ligand interactions analyses or ii) obtain soluble constructions of isolated domain V and more particularly of isolated sugar binding pockets.

By monitoring dextran synthesis of DSR-MΔ2, DSR-OKΔ1 and the double mutant DSR-OKΔ1-Y1162A-F1228A, we showed that the distributive or semi-processive mode of elongation of the dextransucrases was conferred by some of the sugar binding pocket of their domain V. Moreover, the final molar mass of the dextran produced would be directly linked to the mode of dextran

synthesis. However, the dynamic aspects of the elongation process of these enzymes are still questioning and would require analyses at the molecular level. Notably, we showed that the SAXS envelops of both candidates were totally different (linear for DSR-OKΔ1 versus a "horseshoe" shape for DSR-MΔ2). Whether this is of functional relevance remains to be elucidated.

In that purpose, we have tried to set up a SAXS monitoring of the full polymerization mechanism. The objective was to follow protein structural rearrangements and polymer elongation during the reaction. To do so, we designed a set of experiments to: i) determine and compare the molecular envelops of the enzymes in the rest state and during the synthesis, ii) obtain structural parameters on both protein and polymer during the synthesis, in particular on the initial stages of the elongation. Data acquisition on the protein and the polymer during the synthesis needed much more experimental set up than for the observation in rest state. Such experiment was previously done for the synthesis of amylose by amylosucrases, and the obtained results demonstrated that it was possible to simultaneously follow enzymes and polymers during the synthesis (Roblin et al., 2013). In collaboration with Pierre Roblin on the Swing beamline of the synchrotron SOLEIL, we set up a full process to allow SAXS analysis of the reaction *in situ*. A reactor and a fraction collector were built for the experiment (Figure 57). A large number of data were acquired, but we rapidly realized that the observation of the protein would require too high enzymatic concentrations incompatible with the acquisition time of the data (the reaction was too fast!). We then decided to focus on the polymer formation, but here again, limitations appeared. Probably because of the polymer size, with DSR-OK (even with very low enzyme (<0.005 mg/mL) and substrate concentrations and low temperature), the detector was saturated since the beginning of the reaction preventing data acquisition. Concerning DSR-MΔ2, we acquired data on the kinetic of the reactions with different sucrose concentrations (5, 100 and 300 g/L). By fitting the data with a Gauss chain model, we obtained some information on the polymer characteristics along its synthesis (see example of the 100 g/L kinetic on the graphics of Figure 57). However, we still need to work on these data to validate the method and refine interpretation. It would also be interesting to combine these investigations with other technics such as real time experiments either by AFM or through single molecule experiments that could be very informative to understand the dynamics of these amazing catalysts. NMR analyses of the protein dynamics may also bring complementary information.

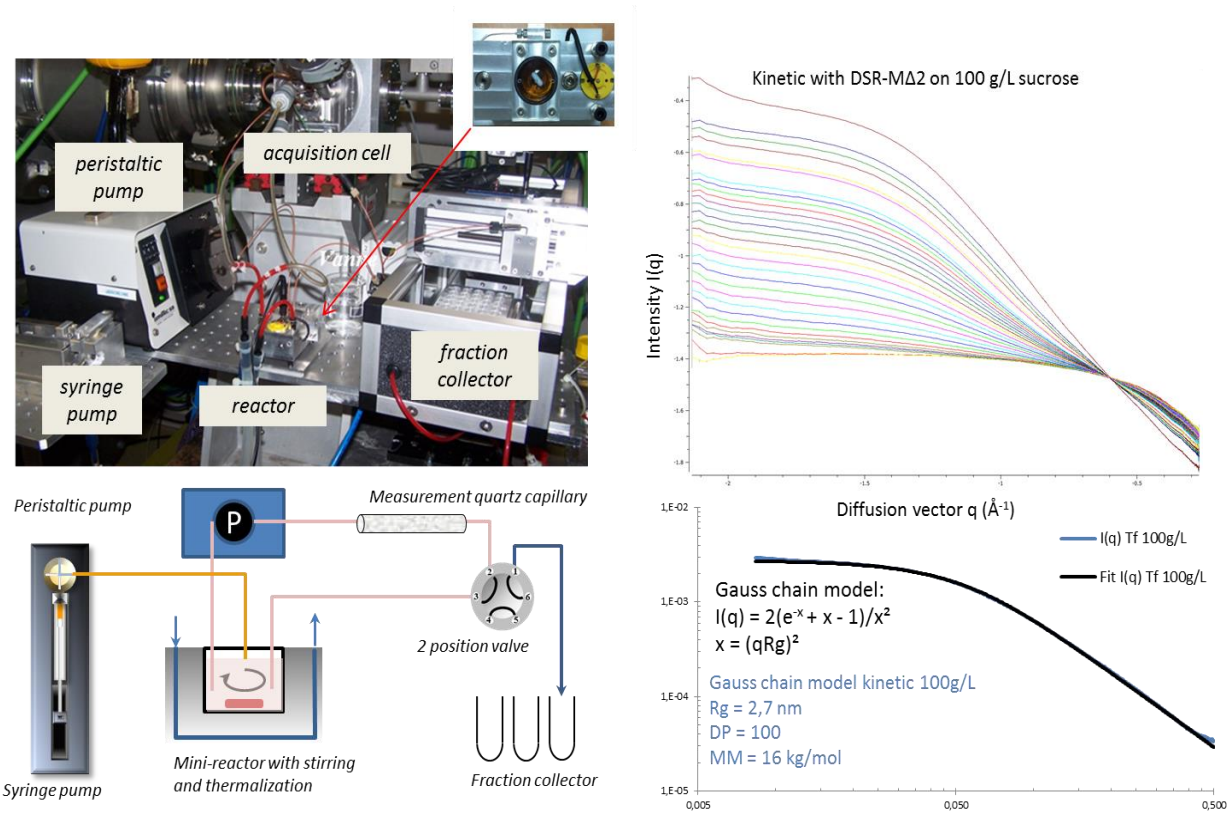


Figure 57: **SAXS experimental set up and preliminary results.**

In the present work, we also report the first series of ^{1}H ^{15}N NMR experiments performed on GSs and particularly on DSR-MΔV, a 100 kDa truncated form of DSR-M devoid of its entire domain V. NMR on such a large protein is difficult and we already have been surprised to be able to make observations and to start assigning the various signals. Of course, the results still need further validation. In that goal, protein expression using marked indole precursor (Schörghuber et al., 2017) will be undertaken and should allow an accurate assignment of the signals. However, it can already be concluded that the loop B2 surrounding the catalytic cleft and particularly the Trp624 of this loop play a role on the polymerization ability of the enzyme. We can think that further NMR experiments may bring additional information for the understanding of the elongation mechanism dynamics. These findings added to previous works, notably on the GTF180-ΔN dextransucrase of *Lactobacillus reuteri* 180 (Meng et al., 2014, 2015a), highlight the involvement of residues around the catalytic area in the control of the polymer size. For that reason, in the absence of crystallographic structural data, a 3D-model of DSR-OK catalytic core (domains A, B and C) was proposed. Although no major distinct features were observable, first analyses of the model highlighted the presence of two aromatic residues in helix α6 of the $(\beta/\alpha)_8$ barrel. Even if they are at longer distance from the active site than residues identified in DSR-MΔ2, we suspect that they have a potential role in the elongation mechanism, and we will rapidly investigate it by mutagenesis.

Altogether, these findings suggest an interplay of the domain V with the catalytic domain in the polymerization mechanism. This was supported by the construction of chimeric variants. It consisted in interchanging the domain V of the two candidates and revealed that the domain swapping is not sufficient to enable a complete size modification of the resulting polymers. Hence, the interplay between domain V and the enzyme core seems to play a crucial role as already mentioned. Structural data with long glucan products (DP>10) that would be bound in both these domains would be of great help to characterize their possible cooperation. Considering that GH70 enzymes are quite difficult to crystallize and that such pure ligand are not commercially available, this would require important efforts. However, as crystals are not needed, emerging technics such as Electron Cryo-Microscopy (cryo-EM) might offer new solutions to circumvent this bottleneck.

Overall, the work accomplished during this thesis allowed the identification of some of the structural determinants controlling the GH70 product size specificity in two different dextransucrases. It seems clear that from a polymerase to another, the domain V plays different role, the strength of its affinity towards polymers likely being the principal explanation. The loops surrounding the active site are also involved in the size determination of the produced polymers. Altogether, these findings open promising strategies for GH70 enzyme engineering aiming at finely tune the size of the synthesized dextrans to respond to precise applications (clinical fractions of 1, 10, 40 or 70 kg/mol for example). Some of the constructions realized in this thesis are already interesting. For example, DSR-MΔV-Y180A is efficient to produce dextrans of 11 kg/mol and the mutant W624A presents high potential for the synthesis of oligosaccharides from 1 to 2.5 kg/mol. For higher fractions between 10 and 70 kg/mol, a full strategy can be imagine and would consist in the assembly of different building blocks. The idea would be to use a common catalytic core and a combination of sugar binding pockets taking in consideration their strength and their localization. To illustrate this new route, and using the results that we obtained, multiple combination of building blocks composed of functional and/or non-functional sugar binding pockets could be imagined (some of them are presented in Figure 58). The objective would be to keep the global folding of the domain V and modulate the distance of dextran anchoring points from the catalytic core. As it is active on its own, the catalytic core of DSR-M (domains A, B, C and IV, corresponding to DSR-MΔV) could be used as the main block, then the binding pockets D or E of DSR-OK could be used as the functional binding blocks while pocket C of DSR-M could be used as the non-functional one, for instance.

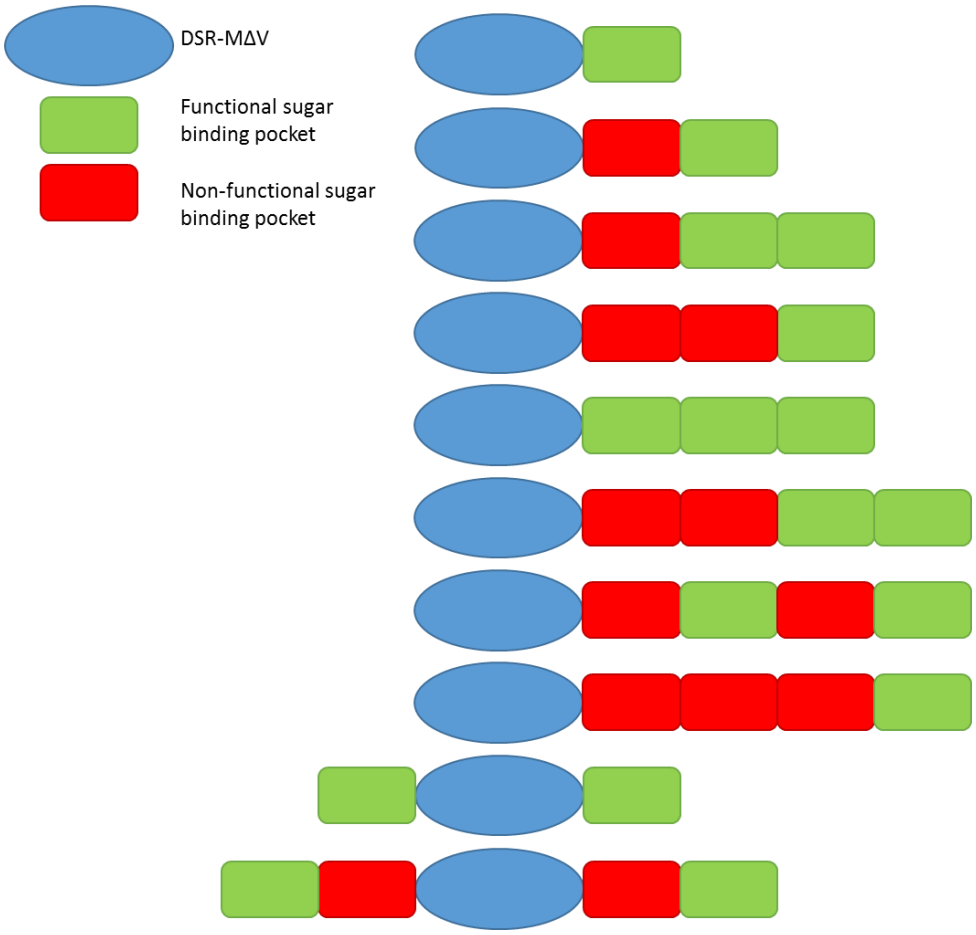


Figure 58: **Schematic representation of some of the potential combinations to engineer dextransucrases in order to produce a catalogue of dextran fraction with controlled molar masses.**

However, in order to envisage such “tailor-made” dextransucrases, several prerequisites need to be encompassed. First, the full understanding of all determinants governing pocket functionality as well as their interplay with the catalytic core appears necessary. Additional structural data on other dextransucrases and other GH70 enzymes in general, if possible in complex with dextrans, would be of great help. Also, the production of soluble isolated pockets (some suspected to be functional & others not functional) followed by a deep characterization of their interaction with dextran is needed. Different technics such as the use of chaperones or cell free system may improve the expression and we hope that the analyses of the interactions would be facilitated by working with much smaller enzymatic forms. These perspectives are quite challenging but would represent major advances from both fundamental and applicative point of views.

References

Abo, H., Matsumura, T., Kodama, T., Ohta, H., Fukui, K., Kato, K., and Kagawa, H. (1991). Peptide sequences for sucrose splitting and glucan binding within Streptococcus sobrinus glucosyltransferase (water-insoluble glucan synthetase). J. Bacteriol. *173*, 989–996.

Amari, M., Arango, L.F.G., Gabriel, V., Robert, H., Morel, S., Moulis, C., Gabriel, B., Remaud-Siméon, M., and Fontagné-Faucher, C. (2013). Characterization of a novel dextransucrase from Weissella confusa isolated from sourdough. Appl. Microbiol. Biotechnol. *97*, 5413–5422.

Amari, M., Valérie, G., Robert, H., Morel, S., Moulis, C., Gabriel, B., Remaud-Siméon, M., and 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.

André, I., Potocki-Véronèse, G., Morel, S., Monsan, P., and Remaud-Siméon, M. (2010). Sucrose-utilizing transglucosidases for biocatalysis. Top. Curr. Chem. *294*, 25–48.

Argüello Morales, M.A., Remaud-Simeon, M., Willemot, R.-M., Vignon, M.R., and Monsan, P. (2001). Novel oligosaccharides synthesized from sucrose donor and cellobiose acceptor by alternansucrase. Carbohydr. Res. *331*, 403–411.

Asensio, J.L., Ardá, A., Cañada, F.J., and Jiménez-Barbero, J. (2013). Carbohydrate–Aromatic Interactions. Acc. Chem. Res. *46*, 946–954.

Auerbach, M., Witt, D., Toler, W., Fierstein, M., Lerner, R.G., and Ballard, H. (1988). Clinical use of the total dose intravenous infusion of iron dextran. J. Lab. Clin. Med. *111*, 566–570.

Awad, S., Hassan, A.N., and Halaweish, F. (2005). Application of Exopolysaccharide-Producing Cultures in Reduced-Fat Cheddar Cheese: Composition and Proteolysis. J. Dairy Sci. *88*, 4195–4203.

Badel, S., Bernardi, T., and Michaud, P. (2011). New perspectives for Lactobacilli exopolysaccharides. Biotechnol. Adv. *29*, 54–66.

Baheti, A., Kumar, L., and Bansal, A.K. (2010). Excipients used in lyophilization of small molecules. J. Excip. Food Chem. *1*.

Bai, Y., Gangoiti, J., Dijkstra, B.W., Dijkhuizen, L., and Pijning, T. (2017). Crystal Structure of 4,6-α-Glucanotransferase Supports Diet-Driven Evolution of GH70 Enzymes from α-Amylases in Oral Bacteria. Structure.

Bar-Even, A., Milo, R., Noor, E., and Tawfik, D.S. (2015). The Moderately Efficient Enzyme: Futile Encounters and Enzyme Floppiness. Biochemistry (Mosc.) *54*, 4969–4977.

Barham, S., and Tippett, J.M. (2004). Dextran starch and flocculant combination for improving red mud clarification.

Barker, P.E., Ganetsos, G., and Ajongwen, N.J. (2007). A novel approach to the production of clinical-grade dextran. J. Chem. Technol. Biotechnol. *57*, 21–26.

Basedow, A.M., and Ebert, K.H. (1977). Ultrasonic degradation of polymers in solution. In Physical Chemistry, (Berlin, Heidelberg: Springer Berlin Heidelberg), pp. 83–148.

Bertrand, A., Morel, S., Lefoulon, F., Rolland, Y., Monsan, P., and Remaud-Simeon, M. (2006). Leuconostoc mesenteroides glucansucrase synthesis of flavonoid glucosides by acceptor reactions in aqueous-organic solvents. Carbohydr. Res. *341*, 855–863.

Bindels, L.B., Delzenne, N.M., Cani, P.D., and Walter, J. (2015). Towards a more comprehensive concept for prebiotics. Nat. Rev. Gastroenterol. Hepatol. *12*, 303–310.

Binder, T.P., Côté, G.L., and Robyt, J.F. (1983). Disproportionation reactions catalyzed by Leuconostoc and Streptococcus glucansucrases. Carbohydr. Res. *124*, 275–286.

Bixler, G.H., Hines, G.E., McGhee, R.M., and Shurter, R.A. (1953). DEXTRAN. Ind. Eng. Chem. *45*, 692–705.

Bizien, T., Durand, D., Roblin, P., Thureau, A., Vachette, P., and Pérez, J. (2016). A Brief Survey of State-of-the-Art BioSAXS. Protein Pept. Lett. *23*, 217–231.

Borneman, A.R., McCarthy, J.M., Chambers, P.J., and Bartowsky, E.J. (2012). Functional Divergence in the Genus Oenococcus as Predicted by Genome Sequencing of the Newly-Described Species, Oenococcus kitaharae. PLoS ONE *7*, e29626.

Bounaix, M.-S., Robert, H., Gabriel, V., Morel, S., Remaud-Siméon, M., Gabriel, B., and Fontagné-Faucher, C. (2010). Characterization of dextran-producing Weissella strains isolated from sourdoughs and evidence of constitutive dextransucrase expression: Characterization of dextran-producing Weissella strains. FEMS Microbiol. Lett. *311*, 18–26.

Bourne, E.J., Sidebotham, R.L., and Weigel, H. (1972). Studies on dextrans and dextranases. Carbohydr. Res. *22*, 13–22.

Bozonnet, S., Dols-Laffargue, M., Fabre, E., Pizzut, S., Remaud-Simeon, M., Monsan, P., and Willemot, R.-M. (2002). Molecular Characterization of DSR-E, an a-1,2 Linkage-Synthesizing Dextransucrase with Two Catalytic Domains. J. Bacteriol. *184*, 5753–5761.

Braswell, E., and Stern, K.G. (1959). Studies on the enzymatic synthesis of dextran. J. Polym. Sci. *41*, 467–473.

Braswell, E., Goodman, A., and Stern, K.G. (1962). Studies on the enzymatic synthesis of dextran. Part II. J. Polym. Sci. *61*, 143–154.

Brison, Y., Fabre, E., Moulis, C., Portais, J.-C., Monsan, P., and 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.

Brison, Y., Pijning, T., Malbert, Y., Fabre, É., Mourey, L., Morel, S., Potocki-Véronèse, G., Monsan, P., Tranier, S., Remaud-Siméon, M., et al. (2012). Functional and structural characterization of α-(1->2) branching sucrase derived from DSR-E glucansucrase. J. Biol. Chem. *287*, 7915–7924.

Brison, Y., Malbert, Y., Czaplicki, G., Mourey, L., Remaud-Simeon, M., and 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.

Burokas, A., Moloney, R.D., Dinan, T.G., and Cryan, J.F. (2015). Microbiota Regulation of the Mammalian Gut–Brain Axis. In Advances in Applied Microbiology, (Elsevier), pp. 1–62.

Burokas, A., Arboleya, S., Moloney, R.D., Peterson, V.L., Murphy, K., Clarke, G., Stanton, C., Dinan, T.G., and Cryan, J.F. (2017). Targeting the Microbiota-Gut-Brain Axis: Prebiotics Have Anxiolytic and Antidepressant-like Effects and Reverse the Impact of Chronic Stress in Mice. Biol. Psychiatry.

Cantarel, B.L., Coutinho, P.M., Rancurel, C., Bernard, T., Lombard, V., and Henrissat, B. (2009). The Carbohydrate-Active EnZymes database (CAZy): an expert resource for Glycogenomics. Nucleic Acids Res. *37*, D233–D238.

Carlson, T.L., Woo, A., and Carlson, T.L. (2006). Use of low-glycemic sweeteners in food and beverage compositions.

CASTAING, J.C., GELO, P.M., LABEAU, M.P., MEYRANT, P., REMAUD, S.M., MOULIS, C., and GRIMAUD, F. (2017). Procede De Bioproduction De Dextrane En Milieu Tensioactif.

Chaouat, M., Le Visage, C., Autissier, A., Chaubet, F., and Letourneur, D. (2006). The evaluation of a small-diameter polysaccharide-based arterial graft in rats. Biomaterials *27*, 5546–5553.

Cheetham, N.W.H., Slodki, M.E., and 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.

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

Cho, Y.W., Allen, R.P., and Earley, C.J. (2013). Lower molecular weight intravenous iron dextran for restless legs syndrome. Sleep Med. *14*, 274–277.

Chourasia, M., Sastry, G.M., and Sastry, G.N. (2011). Aromatic-Aromatic Interactions Database, A(2)ID: an analysis of aromatic π-networks in proteins. Int. J. Biol. Macromol. *48*, 540–552.

Chung, C.-H., and Day, D.F. (2002). Glucooligosaccharides from Leuconostoc mesenteroides B-742 (ATCC 13146): A potential prebiotic. J. Ind. Microbiol. Biotechnol. *29*, 196–199.

Chung, C.H., and Day, D.F. (2004). Efficacy of Leuconostoc mesenteroides (ATCC 13146) isomaltooligosaccharides as a poultry prebiotic. Poult. Sci. *83*, 1302–1306.

Claverie, M., Cioci, G., Vuillemin, M., Monties, N., Roblin, P., Lippens, G., Remaud-Siméon, M., and Moulis, C. (2017). Investigations on the determinants responsible for low molar mass dextran formation by DSR-M dextransucrase. ACS Catal.

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

Côté, G.L., and 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.

Côté, G.L., and 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.

Crittenden, R.G., and Playne, M.J. (1996). Production, properties and applications of food-grade oligosaccharides. Trends Food Sci. Technol. *7*, 353–361.

Dai, T., Zhou, S., Yin, C., Li, S., Cao, W., Liu, W., Sun, K., Dou, H., Cao, Y., and Zhou, G. (2014). Dextran-based fluorescent nanoprobes for sentinel lymph node mapping. Biomaterials *35*, 8227–8235.

Das, D., and Goyal, A. (2014). Characterization and biocompatibility of glucan: a safe food additive from probiotic *Lactobacillus plantarum* DM5: Characterization and biocompatibility of glucan. J. Sci. Food Agric. *94*, 683–690.

Das, D., Baruah, R., and Goyal, A. (2014). A food additive with prebiotic properties of an α-d-glucan from Lactobacillus plantarum DM5. Int. J. Biol. Macromol. *69*, 20–26.

David, G., and Pérez, J. (2009). Combined sampler robot and high-performance liquid chromatography: a fully automated system for biological small-angle X-ray scattering experiments at the Synchrotron SOLEIL SWING beamline. J. Appl. Crystallogr. *42*, 892–900.

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

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

Day, D.F., and Kim, D. (1992). A Novel Process for the Production of Dextrans of Selected Molecular Sizes. Ann. N. Y. Acad. Sci. *672*, 573–576.

De Belder, A. (2003). Dextran (Uppsala: Amersham Biociences AB).

Deitering, S., Anselmo, L., Burnett, A., and Jakeman, B. (2016). Safety and effectiveness of extended versus shortened iron dextran infusion time for the treatment of iron deficiency anemia. Am. J. Hematol. *91*, E498–E499.

Desbonnet, L., Clarke, G., Traplin, A., O'Sullivan, O., Crispie, F., Moloney, R.D., Cotter, P.D., Dinan, T.G., and Cryan, J.F. (2015). Gut microbiota depletion from early adolescence in mice: Implications for brain and behaviour. Brain. Behav. Immun. *48*, 165–173.

Djouzi, Z., and Andrieux, C. (1997). Compared effects of three oligosaccharides on metabolism of intestinal microflora in rats inoculated with a human faecal flora. Br. J. Nutr. *78*, 313–324.

Djouzi, Z., Andrieux, C., Pelenc, V., Somarriba, S., Popot, F., Paul, F., Monsan, P., and Szylit, O. (1995). Degradation and fermentation of alpha-gluco-oligosaccharides by bacterial strains from human colon: in vitro and in vivo studies in gnotobiotic rats. J. Appl. Bacteriol. *79*, 117–127.

Dobruchowska, J.M., Meng, X., Leemhuis, H., Gerwig, G.J., Dijkhuizen, L., and 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.

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

Emsley, P., Lohkamp, B., Scott, W.G., and Cowtan, K. (2010). Features and development of Coot. Acta Crystallogr. D Biol. Crystallogr. *66*, 486–501.

Fabre, E., Bozonnet, S., Arcache, A., Willemot, R.-M., Vignon, M., Monsan, P., and Remaud-Simeon, M. (2005). Role of the Two Catalytic Domains of DSR-E Dextransucrase and Their Involvement in the Formation of Highly -1,2 Branched Dextran. J. Bacteriol. *187*, 296–303.

Falconer, D.J., Mukerjea, R., and Robyt, J.F. (2011). Biosynthesis of dextrans with different molecular weights by selecting the concentration of Leuconostoc mesenteroides B-512FMC dextransucrase, the sucrose concentration, and the temperature. Carbohydr. Res. *346*, 280–284.

Fernández-Arrojo, L., Marín, D., Gómez De Segura, A., Linde, D., Alcalde, M., Gutiérrez-Alonso, P., Ghazi, I., Plou, F.J., Fernández-Lobato, M., and Ballesteros, A. (2007). Transformation of maltose into prebiotic isomaltooligosaccharides by a novel α-glucosidase from Xantophyllomyces dendrorhous. Process Biochem. *42*, 1530–1536.

Fernández-Tornero, C., López, R., García, E., Giménez-Gallego, G., and Romero, A. (2001). A novel solenoid fold in the cell wall anchoring domain of the pneumococcal virulence factor LytA. Nat. Struct. Biol. *8*, 1020–1024.

Finkenstadt, V.L., Côté, G.L., and Willett, J.L. (2011). Corrosion protection of low-carbon steel using exopolysaccharide coatings from Leuconostoc mesenteroides. Biotechnol. Lett. *33*, 1093–1100.

Flodin, P.G.M., and Porath, J.O. (1961). Process of separating materials having different molecular weights and dimensions.

Foster, J., Killen, D., Jolly, P., and Kirtley, J. (1966). Low molecular weight dextran in vascular surgery: prevention of early thrombosis following arterial reconstruction in 85 cases. Ann. Surg. 764–770.

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

Gan, W., Zhang, H., Zhang, Y., and Hu, X. (2014). Biosynthesis of oligodextrans with different Mw by synergistic catalysis of dextransucrase and dextranase. Carbohydr. Polym. *112*, 387–395.

Gangoiti, J., Pijning, T., and Dijkhuizen, L. (2016a). 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.

Gangoiti, J., van Leeuwen, S.S., Vafiadi, C., and Dijkhuizen, L. (2016b). 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 BBA - Gen. Subj. *1860*, 1224–1236.

Gangoiti, J., van Leeuwen, S.S., Gerwig, G.J., Duboux, S., Vafiadi, C., Pijning, T., and Dijkhuizen, L. (2017). 4,3-α-Glucanotransferase, a novel reaction specificity in glycoside hydrolase family 70 and clan GH-H. Sci. Rep. *7*, 39761.

GEEL-SCHUTTEN, V. (2003). Novel Glucans and Novel Glucansucrases Derived from Lactic Acid Bacteria.

Ghosh, T., Chattopadhyay, K., Marschall, M., Karmakar, P., Mandal, P., and Ray, B. (2009). Focus on antivirally active sulfated polysaccharides: From structure–activity analysis to clinical evaluation. Glycobiology *19*, 2–15.

Gibson, D.G. (2009). Synthesis of DNA fragments in yeast by one-step assembly of overlapping oligonucleotides. Nucleic Acids Res. *37*, 6984–6990.

Giffard, P.M., and 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.

Goffin, D., Delzenne, N., Blecker, C., Hanon, E., Deroanne, C., and Paquot, M. (2011). Will Isomalto-Oligosaccharides, a Well-Established Functional Food in Asia, Break through the European and American Market? The Status of Knowledge on these Prebiotics. Crit. Rev. Food Sci. Nutr. *51*, 394–409.

Goulas, A.K., Fisher, D.A., Grimble, G.K., Grandison, A.S., and Rastall, R.A. (2004a). Synthesis of isomaltooligosaccharides and oligodextrans by the combined use of dextransucrase and dextranase. Enzyme Microb. Technol. *35*, 327–338.

Goulas, A.K., Cooper, J.M., Grandison, A.S., and Rastall, R.A. (2004b). Synthesis of isomaltooligosaccharides and oligodextrans in a recycle membrane bioreactor by the combined use of dextransucrase and dextranase. Biotechnol. Bioeng. *88*, 778–787.

Greco, A., Ho, J.G.S., Lin, S.-J., Palcic, M.M., Rupnik, M., and Ng, K.K.-S. (2006). Carbohydrate recognition by Clostridium difficile toxin A. Nat. Struct. Mol. Biol. *13*, 460–461.

Grimoud, J., Durand, H., Courtin, C., Monsan, P., Ouarné, F., Theodorou, V., and Roques, C. (2010). In vitro screening of probiotic lactic acid bacteria and prebiotic glucooligosaccharides to select effective synbiotics. Anaerobe *16*, 493–500.

Gutiérrez-Alonso, P., Gimeno-Pérez, M., Ramírez-Escudero, M., Plou, F.J., Sanz-Aparicio, J., and Fernández-Lobato, M. (2016). Molecular characterization and heterologous expression of a Xanthophyllomyces dendrorhous α-glucosidase with potential for prebiotics production. Appl. Microbiol. Biotechnol. *100*, 3125–3135.

Hamada, S., and Slade, H.D. (1980). Biology, immunology, and cariogenicity of Streptococcus mutans. Microbiol. Rev. *44*, 331–384.

Hedin, H., and Richter, W. (1982). Pathomechanisms of dextran-induced anaphylactoid/anaphylactic reactions in man. Int. Arch. Allergy Appl. Immunol. *68*, 122–126.

Hellman, N.N., Tsuchiya, H.M., Rogovin, S.P., Lamberts, B.L., Tobin, R., Glass, C.A., Stringer, C.S., Jackson, R.W., and Senti, F.R. (1955). Controlled Enzymatic Synthesis of Dextran. Ind. Eng. Chem. *47*, 1593–1598.

Hellmuth, H., Hillringhaus, L., Höbbel, S., Kralj, S., Dijkhuizen, L., and Seibel, J. (2007). Highly Efficient Chemoenzymatic Synthesis of Novel Branched Thiooligosaccharides by Substrate Direction with Glucansucrases. ChemBioChem *8*, 273–276.

Hellmuth, H., Wittrock, S., Kralj, S., Dijkhuizen, L., Hofer, B., and Seibel, J. (2008). Engineering the Glucansucrase GTFR Enzyme Reaction and Glycosidic Bond Specificity: Toward Tailor-Made Polymer and Oligosaccharide Products †. Biochemistry (Mosc.) *47*, 6678–6684.

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

Hershline, R.K. (2004). Antiviral composition.

Hu, Y., Ketabi, A., Buchko, A., and Gänzle, M.G. (2013). Metabolism of isomalto-oligosaccharides by *Lactobacillus reuteri* and bifidobacteria. Lett. Appl. Microbiol. *57*, 108–114.

Hung, V.S., Hatada, Y., Goda, S., Lu, J., Hidaka, Y., Li, Z., Akita, M., Ohta, Y., Watanabe, K., Matsui, H., et al. (2005). α-Glucosidase from a strain of deep-sea Geobacillus: a potential enzyme for the biosynthesis of complex carbohydrates. Appl. Microbiol. Biotechnol. *68*, 757–765.

Ignatova-Ivanova, T., and Ivanov, R. (2014). Exopolysaccharides from lactic acid bacteria as corrosion inhibitors. Acta Sci. Nat. *3*.

Irague, R., Rolland-Sabaté, A., Tarquis, L., Doublier, J.L., Moulis, C., Monsan, P., Remaud-Siméon, M., Potocki-Véronèse, G., and Buléon, A. (2012). Structure and Property Engineering of α-D-Glucans Synthesized by Dextransucrase Mutants. Biomacromolecules *13*, 187–195.

Irague, R., Tarquis, L., André, I., Moulis, C., Morel, S., Monsan, P., Potocki-Véronèse, G., and Remaud-Siméon, M. (2013). Combinatorial Engineering of Dextransucrase Specificity. PLoS ONE *8*, e77837.

Ito, K., Ito, S., Shimamura, T., Weyand, S., Kawarasaki, Y., Misaka, T., Abe, K., Kobayashi, T., Cameron, A.D., and Iwata, S. (2011). Crystal structure of glucansucrase from the dental caries pathogen Streptococcus mutans. J. Mol. Biol. *408*, 177–186.

Jeanes, A., Wilham, C.A., and Miers, J.C. (1948). Preparation and characterization of dextran from Leuconostoc mesenteroides. J. Biol. Chem. *176*, 603–615.

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., and Rist, C.E. (1954). Characterization and Classification of Dextrans from Ninety-six Strains of Bacteria [1b]. J. Am. Chem. Soc. *76*, 5041–5052.

Joucla, G., Pizzut, S., Monsan, P., and Remaud-Simeon, M. (2006). Construction of a fully active truncated alternansucrase partially deleted of its carboxy-terminal domain. FEBS Lett. *580*, 763–768.

Kabsch, W. (2010). XDS. Acta Crystallogr. D Biol. Crystallogr. *66*, 125–132.

Kaneko, T., Kohmoto, T., Kikuchi, H., Shiota, M., Iino, H., and Mitsuoka, T. (1994). Effects of Isomaltooligosaccharides with Different Degrees of Polymerization on Human Fecal Bifidobactcria. Biosci. Biotechnol. Biochem. *58*, 2288–2290.

Katina, K., Maina, N.H., Juvonen, R., Flander, L., Johansson, L., Virkki, L., Tenkanen, M., and Laitila, A. (2009). In situ production and analysis of Weissella confusa dextran in wheat sourdough. Food Microbiol. *26*, 734–743.

Ketabi, A., Dieleman, L.A., and Gänzle, M.G. (2011). Influence of isomalto-oligosaccharides on intestinal microbiota in rats: Intestinal IMO fermentation in rats. J. Appl. Microbiol. *110*, 1297–1306.

Khalikova, E., Susi, P., and Korpela, T. (2005). Microbial Dextran-Hydrolyzing Enzymes: Fundamentals and Applications. Microbiol. Mol. Biol. Rev. *69*, 306–325.

Kim, D., Robyt, J.F., Lee, S.-Y., Lee, J.-H., and Kim, Y.-M. (2003). Dextran molecular size and degree of branching as a function of sucrose concentration, pH, and temperature of reaction of Leuconostoc mesenteroides B-512FMCM dextransucrase. Carbohydr. Res. *338*, 1183–1189.

Kingston, K.B., Allen, D.M., and Jacques, N.A. (2002). Role of the C-terminal YG repeats of the primer-dependent streptococcal glucosyltransferase, GtfJ, in binding to dextran and mutan. Microbiology *148*, 549–558.

Kobayashi, M., and Matsuda, K. (1977). Structural Characteristics of Dextrans Synthesized by Dextransucrases from Leuconostoc mesenteroides NRRL B-1299. Agric. Biol. Chem. *41*, 1931–1937.

Koepsell, H.J., Tsuchiya, H.M., Hellman, N.N., Kazenko, A., Hoffman, C.A., Sharpe, E.S., and Jackson, R.W. (1953). Enzymatic synthesis of dextran; acceptor specificity and chain initiation. J. Biol. Chem. *200*, 793–801.

Koga, T., Oho, T., Shimazaki, Y., and Nakano, Y. (2002). Immunization against dental caries. Vaccine *20*, 2027–2044.

Komatsu, H., Katayama, M., Sawada, M., Hirata, Y., Mori, M., Inoue, T., Fukui, K., Fukada, H., and 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.

Koshland, D.E. (1953). STEREOCHEMISTRY AND THE MECHANISM OF ENZYMATIC REACTIONS. Biol. Rev. *28*, 416–436.

Kothari, D., Baruah, R., and Goyal, A. (2012). Immobilization of glucansucrase for the production of gluco-oligosaccharides from Leuconostoc mesenteroides. Biotechnol. Lett. *34*, 2101–2106.

Kothari, D., Das, D., Patel, S., and Goyal, A. (2014). Dextran and Food Application. In Polysaccharides, K.G. Ramawat, and J.-M. Mérillon, eds. (Cham: Springer International Publishing), pp. 1–16.

Kralj, S. (2004a). Glucan synthesis in the genus Lactobacillus: isolation and characterization of glucansucrase genes, enzymes and glucan products from six different strains. Microbiology *150*, 3681–3690.

Kralj, S. (2004b). Biochemical and molecular characterization of Lactobacillus reuteri 121 reuteransucrase. Microbiology *150*, 2099–2112.

Kralj, S., van Geel-Schutten, G.H., Rahaoui, H., Leer, R.J., Faber, E.J., van der Maarel, M.J.E.C., and Dijkhuizen, L. (2002). Molecular Characterization of a Novel Glucosyltransferase from Lactobacillus reuteri Strain 121 Synthesizing a Unique, Highly Branched Glucan with -(1->4) and -(1->6) Glucosidic Bonds. Appl. Environ. Microbiol. *68*, 4283–4291.

Kralj, S., Stripling, E., Sanders, P., van Geel-Schutten, G.H., and Dijkhuizen, L. (2005a). Highly Hydrolytic Reuteransucrase from Probiotic Lactobacillus reuteri Strain ATCC 55730. Appl. Environ. Microbiol. *71*, 3942–3950.

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

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

Kralj, S., Grijpstra, P., Leeuwen, S.S. van, Leemhuis, H., Dobruchowska, J.M., Kaaij, R.M. van der, Malik, A., Oetari, A., Kamerling, J.P., and 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.

Kubik, C., Sikora, B., and Bielecki, S. (2004). Immobilization of dextransucrase and its use with soluble dextranase for glucooligosaccharides synthesis. Enzyme Microb. Technol. *34*, 555–560.

Kuge, T., Kobayashi, K., Kitamura, S., and Tanahashi, H. (1987). Degrees of long-chain branching in dextrans11Part II of the series G.p.c. Analyses of Polysaccharides. Carbohydr. Res. *160*, 205–214.

LABEAU, M.P., REMAUD, S.M., MOULIS, C., and GRIMAUD, F. (2017a). Dextrane Carboxyle.

LABEAU, M.P., GELO, P.M., MEYRANT, P., REMAUD, S.M., MOULIS, C., and GRIMAUD, F. (2017b). Procede De Bioproduction De Dextrane En Milieu Salin.

Lambie, J.M., Barber, D.C., Dhall, D.P., and Matheson, N.A. (1970). Dextran 70 in Prophylaxis of Postoperative Venous Thrombosis. A Controlled Trial. BMJ *2*, 144–145.

Larm, O., Lindberg, B., and Svensson, S. (1971). Studies on the length of the side chains of the dextran elaborated by Leuconostoc mesenteroides NRRL B-512. Carbohydr. Res. *20*, 39–48.

Le Visage, C., Gournay, O., Benguirat, N., Hamidi, S., Chaussumier, L., Mougenot, N., Flanders, J.A., Isnard, R., Michel, J.-B., Hatem, S., et al. (2012). Mesenchymal stem cell delivery into rat infarcted myocardium using a porous polysaccharide-based scaffold: a quantitative comparison with endocardial injection. Tissue Eng. Part A *18*, 35–44.

Leathers, T.D., Nunnally, M.S., and Côté, G.L. (2009). Modification of alternan by dextranase. Biotechnol. Lett. *31*, 289–293.

Lee, D., Barsky, D., Hughes, R., Driggers, J., Poupard, N., and Mascarenhas, M. (2017). Evaluation of the safety of iron dextran with parenteral nutrition in the paediatric inpatient setting: Safety of iron in parenteral nutrition. Nutr. Diet.

Lee, R.-S., Li, Y.-C., and Wang, S.-W. (2015). Synthesis and characterization of amphiphilic photocleavable polymers based on dextran and substituted-ε-caprolactone. Carbohydr. Polym. *117*, 201–210.

Leemhuis, H., Pijning, T., Dobruchowska, J.M., Dijkstra, B.W., and Dijkhuizen, L. (2012). Glycosidic bond specificity of glucansucrases: on the role of acceptor substrate binding residues. Biocatal. Biotransformation *30*, 366–376.

Leemhuis, H., Pijning, T., Dobruchowska, J.M., van Leeuwen, S.S., Kralj, S., Dijkstra, B.W., and Dijkhuizen, L. (2013a). Glucansucrases: Three-dimensional structures, reactions, mechanism, α-glucan analysis and their implications in biotechnology and food applications. J. Biotechnol. *163*, 250–272.

Leemhuis, H., Dijkman, W.P., Dobruchowska, J.M., Pijning, T., Grijpstra, P., Kralj, S., Kamerling, J.P., and Dijkhuizen, L. (2013b). 4,6-α-Glucanotransferase activity occurs more widespread in Lactobacillus strains and constitutes a separate GH70 subfamily. Appl. Microbiol. Biotechnol. *97*, 181–193.

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

van Leeuwen, S.S., Kralj, S., Gerwig, G.J., Dijkhuizen, L., and Kamerling, J.P. (2008b). Structural Analysis of Bioengineered α-Glucan Produced by a Triple Mutant of the Glucansucrase GTF180 Enzyme from Lactobacillus reuteri Strain 180: Generation of (α1→4) Linkages in a Native (1→3)(1→6)-α- D -Glucan. Biomacromolecules *9*, 2251–2258.

Leme, A.F.P., Koo, H., Bellato, C.M., Bedi, G., and Cury, J.A. (2006). The Role of Sucrose in Cariogenic Dental Biofilm Formation—New Insight. J. Dent. Res. *85*, 878–887.

Li, B., Wang, Q., Wang, X., Wang, C., and Jiang, X. (2013). Preparation, drug release and cellular uptake of doxorubicin-loaded dextran-b-poly(ε-caprolactone) nanoparticles. Carbohydr. Polym. *93*, 430–437.

Likotrafiti, E., Tuohy, K.M., Gibson, G.R., and Rastall, R.A. (2014). An in vitro study of the effect of probiotics, prebiotics and synbiotics on the elderly faecal microbiota. Anaerobe *27*, 50–55.

Lindberg, B., and Svensson, S. (1968). Structural studies on dextran from Leuconostoc mesenteroides NRRL B-512F. Acta Chem. Scand. *22*, 1907–1912.

Lis, M., Shiroza, T., and Kuramitsu, H.K. (1995). Role of C-terminal direct repeating units of the Streptococcus mutans glucosyltransferase-S in glucan binding. Appl. Environ. Microbiol. *61*, 2040–2042.

Liu, M., Mao, X., Ye, C., Huang, H., Nicholson, J.K., and Lindon, J.C. (1998). Improved WATERGATE Pulse Sequences for Solvent Suppression in NMR Spectroscopy. J. Magn. Reson. *132*, 125–129.

Lombard, V., Golaconda Ramulu, H., Drula, E., Coutinho, P.M., and Henrissat, B. (2014). The carbohydrate-active enzymes database (CAZy) in 2013. Nucleic Acids Res. *42*, D490-495.

MacGregor, E.A., Jespersen, H.M., and Svensson, B. (1996). A circularly permuted α-amylase-type α/β-barrel structure in glucan-synthesizing glucosyltransferases. FEBS Lett. *378*, 263–266.

Madsen, L.R., Stanley, S., Swann, P., and Oswald, J. (2017). A Survey of Commercially Available Isomaltooligosaccharide-Based Food Ingredients: A survey of IMO-based food ingredients…. J. Food Sci. *82*, 401–408.

Maina, N.H., Virkki, L., Pynnönen, H., Maaheimo, H., and Tenkanen, M. (2011). Structural Analysis of Enzyme-Resistant Isomaltooligosaccharides Reveals the Elongation of α-(1→3)-Linked Branches in Weissella confusa Dextran. Biomacromolecules *12*, 409–418.

Malbert, Y., Pizzut-Serin, S., Massou, S., Cambon, E., Laguerre, S., Monsan, P., Lefoulon, F., Morel, S., André, I., and Remaud-Simeon, M. (2014). Extending the Structural Diversity of α-Flavonoid Glycosides with Engineered Glucansucrases. ChemCatChem *6*, 2282–2291.

Matthews, B.W. (1968). Solvent content of protein crystals. J. Mol. Biol. *33*, 491–497.

Mayer, R.M. (1987). [57] Dextransucrase: A glucosyltransferase from Streptococcus sanguis. In Methods in Enzymology, (Elsevier), pp. 649–661.

McCoy, A.J., Grosse-Kunstleve, R.W., Adams, P.D., Winn, M.D., Storoni, L.C., and Read, R.J. (2007). Phaser crystallographic software. J. Appl. Crystallogr. *40*, 658–674.

MEISSNER, A., SCHULTE-HERBRÜGGEN, T., BRIAND, J., and SØRENSEN, O.W. (1998). Double spin-state-selective coherence transfer. Application for two-dimensional selection of multiplet components with long transverse relaxation times. Mol. Phys. *95*, 1137–1142.

Mende, S., Peter, M., Bartels, K., Dong, T., Rohm, H., and Jaros, D. (2013). Concentration dependent effects of dextran on the physical properties of acid milk gels. Carbohydr. Polym. *98*, 1389–1396.

Meng, X., Dobruchowska, J.M., Pijning, T., Lopez, C.A., Kamerling, J.P., and 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.

Meng, X., Dobruchowska, J.M., Pijning, T., Gerwig, G.J., Kamerling, J.P., and Dijkhuizen, L. (2015a). Truncation of domain V of the multidomain glucansucrase GTF180 of Lactobacillus reuteri 180 heavily impairs its polysaccharide-synthesizing ability. Appl. Microbiol. Biotechnol. *99*, 5885–5894.

Meng, X., Pijning, T., Dobruchowska, J.M., Gerwig, G.J., and 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 from *Lactobacillus reuteri* 180. J. Biol. Chem. *290*, 30131–30141.

Meng, X., Dobruchowska, J.M., Gerwig, G.J., Kamerling, J.P., and Dijkhuizen, L. (2015c). Synthesis of oligo- and polysaccharides by Lactobacillus reuteri 121 reuteransucrase at high concentrations of sucrose. Carbohydr. Res. *414*, 85–92.

Meng, X., Gangoiti, J., Bai, Y., Pijning, T., Van Leeuwen, S.S., and Dijkhuizen, L. (2016a). 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. *73*, 2681–2706.

Meng, X., Pijning, T., Dobruchowska, J.M., Yin, H., Gerwig, G.J., and Dijkhuizen, L. (2016b). Structural determinants of alternating (α1 → 4) and (α1 → 6) linkage specificity in reuteransucrase of Lactobacillus reuteri. Sci. Rep. *6*.

Meng, X., Dobruchowska, J.M., Pijning, T., Gerwig, G.J., and Dijkhuizen, L. (2016c). Synthesis of New Hyperbranched α-Glucans from Sucrose by *Lactobacillus reuteri* 180 Glucansucrase Mutants. J. Agric. Food Chem. *64*, 433–442.

Meng, X., Pijning, T., Tietema, M., Dobruchowska, J.M., Yin, H., Gerwig, G.J., Kralj, S., and Dijkhuizen, L. (2017). Characterization of the glucansucrase GTF180 W1065 mutant enzymes producing polysaccharides and oligosaccharides with altered linkage composition. Food Chem. *217*, 81–90.

Messaoudi, M., Lalonde, R., Violle, N., Javelot, H., Desor, D., Nejdi, A., Bisson, J.-F., Rougeot, C., Pichelin, M., Cazaubiel, M., et al. (2011). Assessment of psychotropic-like properties of a probiotic formulation (Lactobacillus helveticus R0052 and Bifidobacterium longum R0175) in rats and human subjects. Br. J. Nutr. *105*, 755–764.

Meulenbeld, G.H., Zuilhof, H., van Veldhuizen, A., van den Heuvel, R.H., and Hartmans, S. (1999). Enhanced (+)-catechin transglucosylating activity of Streptococcus mutans GS-5 glucosyltransferase-D due to fructose removal. Appl. Environ. Microbiol. *65*, 4141–4147.

Miller, G.L. (1959). Use of Dinitrosalicylic Acid Reagent for Determination of Reducing Sugar. Anal. Chem. *31*, 426–428.

Monchois, V., Willemot, R.M., Remaud-Simeon, M., Croux, C., and Monsan, P. (1996). Cloning and sequencing of a gene coding for a novel dextransucrase from Leuconostoc mesenteroides NRRL B-1299 synthesizing only alpha (1-6) and alpha (1-3) linkages. Gene *182*, 23–32.

Monchois, V., Reverte, A., Remaud-Simeon, M., Monsan, P., and Willemot, R.M. (1998a). Effect of Leuconostoc mesenteroides NRRL B-512F dextransucrase carboxy-terminal deletions on dextran and oligosaccharide synthesis. Appl. Environ. Microbiol. *64*, 1644–1649.

Monchois, V., Remaud-Simeon, M., Monsan, P., and Willemot, R.M. (1998b). Cloning and sequencing of a gene coding for an extracellular dextransucrase (DSRB) from Leuconostoc mesenteroides NRRL B-1299 synthesizing only a alpha (1-6) glucan. FEMS Microbiol. Lett. *159*, 307–315.

Monchois, V., Arguello-Morales, M., and Russell, R.R. (1999). Isolation of an active catalytic core of Streptococcus downei MFe28 GTF-I glucosyltransferase. J. Bacteriol. *181*, 2290–2292.

Monsan, P., Bozonnet, S., Albenne, C., Joucla, G., Willemot, R.-M., and Remaud-Siméon, M. (2001). Homopolysaccharides from lactic acid bacteria. Int. Dairy J. *11*, 675–685.

Monsan, P., Remaud-Siméon, M., and André, I. (2010). Transglucosidases as efficient tools for oligosaccharide and glucoconjugate synthesis. Curr. Opin. Microbiol. *13*, 293–300.

Moody, G.M., and Rushforth, C.A. (1991). Recovery of alumina from bauxite.

Mooser, G., and Wong, C. (1988). Isolation of a glucan-binding domain of glucosyltransferase (1,6-alpha-glucan synthase) from Streptococcus sobrinus. Infect. Immun. *56*, 880–884.

Mooser, G., Hefta, S.A., Paxton, R.J., Shively, J.E., and 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.

Mori, T., Asakura, M., and Okahata, Y. (2011). Single-molecule force spectroscopy for studying kinetics of enzymatic dextran elongations. J. Am. Chem. Soc. *133*, 5701–5703.

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

Moulis, C., Arcache, A., Escalier, P.-C., Rinaudo, M., Monsan, P., Remaud-Simeon, M., and Potocki-Veronese, G. (2006a). High-level production and purification of a fully active recombinant dextransucrase from Leuconostoc mesenteroides NRRL B-512F. FEMS Microbiol. Lett. *261*, 203–210.

Moulis, C., Joucla, G., Harrison, D., Fabre, E., Potocki-Veronese, G., Monsan, P., and Remaud-Simeon, M. (2006b). Understanding the Polymerization Mechanism of Glycoside-Hydrolase Family 70 Glucansucrases. J. Biol. Chem. *281*, 31254–31267.

Moulis, C., Vaca Medina, G., Suwannarangsee, S., Monsan, P., Remaud-Simeon, M., and Potocki-Veronese, G. (2008). One-step synthesis of isomalto-oligosaccharide syrups and dextrans of controlled size using engineered dextransucrase. Biocatal. Biotransformation *26*, 141–151.

Moulis, C., André, I., and Remaud-Simeon, M. (2016). GH13 amylosucrases and GH70 branching sucrases, atypical enzymes in their respective families. Cell. Mol. Life Sci. *73*, 2661–2679.

Mountzouris, K.C., Gilmour, S.G., and Rastall, R.A. (2002). Continuous Production of Oligodextrans via Controlled Hydrolysis of Dextran in an Enzyme Membrane Reactor. J. Food Sci. *67*, 1767–1771.

Murshudov, G.N., Vagin, A.A., and Dodson, E.J. (1997). Refinement of macromolecular structures by the maximum-likelihood method. Acta Crystallogr. D Biol. Crystallogr. *53*, 240–255.

Naessens, M., Cerdobbel, A., Soetaert, W., and Vandamme, E.J. (2005). Leuconostoc dextransucrase and dextran: production, properties and applications. J. Chem. Technol. Biotechnol. *80*, 845–860.

Nam, S.H., Ko, E.A., Jang, S.S., Kim, D.W., Kim, S.Y., Hwang, D.S., and Kim, D. (2007). Maximization of dextransucrase activity expressed in E. coli by mutation and its functional characterization. Biotechnol. Lett. *30*, 135–143.

Naoi, T., Fujioka, T., Okamura, H., and Satake, M. (1987). Silver halide photographic material.

Ni, M., Xu, K., and Lin, J. (2013). Application of FITC-dextran in manufacturing fundus angiographic agent.

Novak, L.J., and Stoycos, G.S. (1958). Method for producing clinical dextran.

Olano-Martin, E., Mountzouris, K.C., Gibson, G.R., and Rastall, R.A. (2000). In vitro fermentability of dextran, oligodextran and maltodextrin by human gut bacteria. Br. J. Nutr. *83*, 247–255.

Ölçer, Z., and Tanriseven, A. (2010). Co-immobilization of dextransucrase and dextranase in alginate. Process Biochem. *45*, 1645–1651.

Olivares-Illana, V., Lopez-Munguia, A., and Olvera, C. (2003). Molecular Characterization of Inulosucrase from Leuconostoc citreum: a Fructosyltransferase within a Glucosyltransferase. J. Bacteriol. *185*, 3606–3612.

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

Paini, M., Aliakbarian, B., Casazza, A.A., Perego, P., Ruggiero, C., and Pastorino, L. (2015). Chitosan/dextran multilayer microcapsules for polyphenol co-delivery. Mater. Sci. Eng. C *46*, 374–380.

Passerini, D., Vuillemin, M., Ufarté, L., Morel, S., Loux, V., Fontagné-Faucher, C., Monsan, P., Remaud-Siméon, M., and 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.

Paul, F., Oriol, E., Auriol, D., and Monsan, P. (1986). Acceptor reaction of a highly purified dextransucrase with maltose and oligosaccharides. Application to the synthesis of controlled-molecular-weight dextrans. Carbohydr. Res. *149*, 433–441.

Paul, F.B., Lopez, M.C., Remaud, M.M., Pelenc, V.P., and Monsan, P.F. (1992). Method for the Production of (alpha)(1->2) Oligodextrans Using Leuconostoc Mesenteroides B-1299.

Pervushin, K., Riek, R., Wider, G., and Wüthrich, K. (1997). Attenuated T2 relaxation by mutual cancellation of dipole-dipole coupling and chemical shift anisotropy indicates an avenue to NMR structures of very large biological macromolecules in solution. Proc. Natl. Acad. Sci. U. S. A. *94*, 12366–12371.

Pervushin, K.V., Wider, G., and Wüthrich, K. (1998). Single Transition-to-single Transition Polarization Transfer (ST2-PT) in [15N,1H]-TROSY. J. Biomol. NMR *12*, 345–348.

Petoukhov, M.V., Franke, D., Shkumatov, A.V., Tria, G., Kikhney, A.G., Gajda, M., Gorba, C., Mertens, H.D.T., Konarev, P.V., and Svergun, D.I. (2012). New developments in the ATSAS program package for small-angle scattering data analysis. J. Appl. Crystallogr. *45*, 342–350.

Pijning, T., Vujičić-Žagar, A., Kralj, S., Dijkhuizen, L., and 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.

Pijning, T., Vujičić-Žagar, A., Kralj, S., Dijkhuizen, L., and Dijkstra, B.W. (2014). Flexibility of truncated and full-length glucansucrase GTF180 enzymes from Lactobacillus reuteri 180. FEBS J. *281*, 2159–2171.

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

Pleszczyńska, M., Wiater, A., and Szczodrak, J. (2010). Mutanase from Paenibacillus sp. MP-1 produced inductively by fungal α-1,3-glucan and its potential for the degradation of mutan and Streptococcus mutans biofilm. Biotechnol. Lett. *32*, 1699–1704.

Pollock, R.F., and Muduma, G. (2017). A budget impact analysis of parenteral iron treatments for iron deficiency anemia in the UK: reduced resource utilization with iron isomaltoside 1000. Clin. Outcomes Res. *Volume 9*, 475–483.

Porath, J., and Flodin, P. (1959). Gel Filtration: A Method for Desalting and Group Separation. Nature *183*, 1657–1659.

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

Remaud, M., Paul, F., Monsan, P., Heyraud, A., and Rinaudo, M. (1991). Molecular Weight Characterization and Structural Properties of Controlled Molecular Weight Dextrans Synthesized by Acceptor Reaction using Highly Purified Dextransucrase. J. Carbohydr. Chem. *10*, 861–876.

Remaud, M., Paul, F., Monsan, P., Lopez-Munguia, A., and Vignon, M. (1992). Characterization of α-(1→3) Branched Oligosaccharides Synthesized by Acceptor Reaction with the Extracellular Glucosyltransferases from *L. Mesenteroides* NRRL B-742. J. Carbohydr. Chem. *11*, 359–378.

Roberfroid, M. (2002). Functional food concept and its application to prebiotics. Dig. Liver Dis. *34*, S105–S110.

Roblin, P., Potocki-Véronèse, G., Guieysse, D., Guerin, F., Axelos, M.A.V., Perez, J., and Buleon, A. (2013). SAXS Conformational Tracking of Amylose Synthesized by Amylosucrases. Biomacromolecules *14*, 232–239.

Robyt, J.F. (1985). Dextran (New-York).

Robyt, J.F., Kimble, B.K., and Walseth, T.F. (1974). The mechanism of dextransucrase action. Direction of dextran biosynthesis. Arch. Biochem. Biophys. *165*, 634–640.

Röper, H., and Koch, H. (1988). New Carbohydrate Derivatives from Biotechnical and Chemical Processes. Starch - Stärke *40*, 453–464.

Saburi, W., Okuyama, M., Kumagai, Y., Kimura, A., and Mori, H. (2015). Biochemical properties and substrate recognition mechanism of GH31 α-glucosidase from Bacillus sp. AHU 2001 with broad substrate specificity. Biochimie *108*, 140–148.

Saldías, C., Velásquez, L., Quezada, C., and Leiva, A. (2015). Physicochemical assessment of Dextran-g-Poly (ε-caprolactone) micellar nanoaggregates as drug nanocarriers. Carbohydr. Polym. *117*, 458–467.

Sarbini, S.R., Kolida, S., Naeye, T., Einerhand, A., Brison, Y., Remaud-Simeon, M., Monsan, P., Gibson, G.R., and Rastall, R.A. (2011). In Vitro Fermentation of Linear and -1,2-Branched Dextrans by the Human Fecal Microbiota. Appl. Environ. Microbiol. *77*, 5307–5315.

Sarbini, S.R., Kolida, S., Naeye, T., Einerhand, A.W., Gibson, G.R., and 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.

Savignac, H.M., Corona, G., Mills, H., Chen, L., Spencer, J.P.E., Tzortzis, G., and Burnet, P.W.J. (2013). Prebiotic feeding elevates central brain derived neurotrophic factor, N-methyl-d-aspartate receptor subunits and d-serine. Neurochem. Int. *63*, 756–764.

Schmidt, K., Cowen, P.J., Harmer, C.J., Tzortzis, G., Errington, S., and Burnet, P.W.J. (2015). Prebiotic intake reduces the waking cortisol response and alters emotional bias in healthy volunteers. Psychopharmacology (Berl.) *232*, 1793–1801.

Schörghuber, J., Geist, L., Bisaccia, M., Weber, F., Konrat, R., and Lichtenecker, R.J. (2017). Anthranilic acid, the new player in the ensemble of aromatic residue labeling precursor compounds. J. Biomol. NMR *69*, 13–22.

Serino, M., Luche, E., Gres, S., Baylac, A., Bergé, M., Cenac, C., Waget, A., Klopp, P., Iacovoni, J., Klopp, C., et al. (2012). Metabolic adaptation to a high-fat diet is associated with a change in the gut microbiota. Gut *61*, 543–553.

Seymour, F.R., and Knapp, R.D. (1980). Structural analysis of α-d-glucans by 13C-nuclear magnetic resonance, spin-lattice relaxation studies. Carbohydr. Res. *81*, 67–103.

Seymour, F.K., Knapp, R.D., Chen, E.C.M., Bishop, S.H., and Jeanes, A. (1979a). 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.

Seymour, F.R., Slodki, M.E., Plattner, R.D., and 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.

Seymour, F.R., Chen, E.C.M., and Bishop, S.H. (1979b). Methylation structural analysis of unusual dextrans by combined gas-liquid chromatography-mass spectrometry. Carbohydr. Res. *68*, 113–121.

Shah, D.S.H., Joucla, G., Remaud-Simeon, M., and Russell, R.R.B. (2004). Conserved repeat motifs and glucan binding by glucansucrases of oral streptococci and Leuconostoc mesenteroides. J. Bacteriol. *186*, 8301–8308.

Shen, M., and Sali, A. (2006). Statistical potential for assessment and prediction of protein structures. Protein Sci. *15*, 2507–2524.

Shukla, R., and Goyal, A. (2013). Novel dextran from Pediococcus pentosaceus CRAG3 isolated from fermented cucumber with anti-cancer properties. Int. J. Biol. Macromol. *62*, 352–357.

Shukla, S., Verma, A.K., Kajala, I., Nyyssolä, A., Baruah, R., Katina, K., Juvonen, R., Tenkanen, M., and Goyal, A. (2016). Structure modeling and functional analysis of recombinant dextransucrase from *Weissella confusa* Cab3 expressed in *Lactococcus lactis*. Prep. Biochem. Biotechnol. *46*, 822–832.

Silva, A.K.A., Juenet, M., Meddahi-Pellé, A., and Letourneur, D. (2015). Polysaccharide-based strategies for heart tissue engineering. Carbohydr. Polym. *116*, 267–277.

Skov, L.K., Mirza, O., Henriksen, A., De Montalk, G.P., Remaud-Simeon, M., Sarçabal, P., Willemot, R.-M., Monsan, P., and Gajhede, M. (2001). Amylosucrase, a Glucan-synthesizing Enzyme from the α-Amylase Family. J. Biol. Chem. *276*, 25273–25278.

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

Sun, G., and Chu, C.-C. (2011). Biodegradable nanospheres self-assembled from complementary hydrophilic dextran macromers. Carbohydr. Polym. *86*, 910–916.

Sun, G., and Mao, J.J. (2012). Engineering dextran-based scaffolds for drug delivery and tissue repair. Nanomed. *7*, 1771–1784.

Suresh Kumar, A., Mody, K., and Jha, B. (2007). Bacterial exopolysaccharides – a perception. J. Basic Microbiol. *47*, 103–117.

Suwannarangsee, S., Moulis, C., Potocki-Veronese, G., Monsan, P., Remaud-Simeon, M., and Chulalaksananukul, W. (2007). Search for a dextransucrase minimal motif involved in dextran binding. FEBS Lett. *581*, 4675–4680.

Taguchi, M. (1999). Processing method of silver halide light sensitive photographic material.

Takemoto, K.K., and Liebhaber, H. (1962). Virus-polysaccharide interactions. Virology *17*, 499–501.

Tingirikari, J.M.R., Kothari, D., and Goyal, A. (2014). Superior prebiotic and physicochemical properties of novel dextran from Weissella cibaria JAG8 for potential food applications. Food Funct. *5*, 2324.

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

Van Hijum, S.A.F.T., Kralj, S., Ozimek, L.K., Dijkhuizen, L., and van Geel-Schutten, I.G.H. (2006). Structure-Function Relationships of Glucansucrase and Fructansucrase Enzymes from Lactic Acid Bacteria. Microbiol. Mol. Biol. Rev. *70*, 157–176.

Van Leeuwen, S.S., Kralj, S., Eeuwema, W., Gerwig, G.J., Dijkhuizen, L., and 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.

Vandamme, E.J., Renard, C.E.F.G., Arnaut, F.R.J., Vekemans, N.M.F., and Tossut, P.P.A. (2002). Process for obtaining improved structure build-up of baked products.

Vettori, M.H.P.B., Blanco, K.C., Cortezi, M., Lima, C.J.B., and Contiero, J. (2012). Dextran: effect of process parameters on production, purification and molecular weight and recent applications. Diálogos Ciênc. *2012*, 171–186.

Villemson, A., Couvreur, P., Gillet, B., Larionova, N., and Gref, R. (2006). Dextran-poly-ε-caprolactone micro- and nanoparticles: preparation, characterization and tamoxifen solubilization. J. Drug Deliv. Sci. Technol. *16*, 307–313.

Vuillemin, M., CLAVERIE, M., MOULIS, C., REMAUD-SIMEON, M., MONSAN, P., SEVERAC, E., and FONTAGNE-FAUCHER, C. (2015a). Protein with Dextran-Saccharase Activity, and Uses.

Vuillemin, M., Claverie, M., Moulis, C., Remaud-Simeon, M., Grimaud, F., Monsan, P., Sabate, A., Garnier, C., Dols-Lafargue, M., and Lucas, P. (2015b). Very High Molar Mass Dextrans.

Vuillemin, M., Claverie, M., Brison, Y., Séverac, E., Bondy, P., Morel, S., Monsan, P., Moulis, C., and Remaud-Siméon, M. (2016). Characterization of the First α-(1→3) Branching Sucrases of the GH70 Family. J. Biol. Chem. *291*, 7687–7702.

Vuillemin, M., Grimaud, F., Claverie, M., Rolland-Sabaté, A., Garnier, C., Lucas, P., Monsan, P., Dols-Lafargue, M., Remaud-Siméon, M., and Moulis, C. (2017). A dextran with unique rheological properties produced by the dextransucrase from Oenococcus kitaharae DSM 17330. Carbohydr. Polym.

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

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

Wang, W., and Malcolm, B.A. (1999). Two-stage PCR protocol allowing introduction of multiple mutations, deletions and insertions using QuikChange Site-Directed Mutagenesis. BioTechniques *26*, 680–682.

Wang, Y.-H., Jiang, Y., Duan, Z.-Y., Shao, W.-L., and Li, H.-Z. (2009). Expression and characterization of an α-glucosidase from Thermoanaerobacter ethanolicus JW200 with potential for industrial application. Biologia (Bratisl.) *64*.

Wasiak, I., Kulikowska, A., Janczewska, M., Michalak, M., Cymerman, I.A., Nagalski, A., Kallinger, P., Szymanski, W.W., and Ciach, T. (2016). Dextran Nanoparticle Synthesis and Properties. PLOS ONE *11*, e0146237.

Webb, B., and Sali, A. (2014). Comparative Protein Structure Modeling Using MODELLER: Comparative Protein Structure Modeling Using MODELLER. In Current Protocols in Bioinformatics, A. Bateman, W.R. Pearson, L.D. Stein, G.D. Stormo, and J.R. Yates, eds. (Hoboken, NJ, USA: John Wiley & Sons, Inc.), p. 5.6.1-5.6.32.

Wever, D.A.Z., Picchioni, F., and Broekhuis, A.A. (2011). Polymers for enhanced oil recovery: A paradigm for structure–property relationship in aqueous solution. Prog. Polym. Sci. *36*, 1558–1628.

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., et al. (2011). Overview of the CCP4 suite and current developments. Acta Crystallogr. D Biol. Crystallogr. *67*, 235–242.

Withers, S.G. (2001). Mechanisms of glycosyl transferases and hydrolases. Carbohydr. Polym. *44*, 325–337.

Wolff, I.A., Mellies, R.L., Lohmar, R.L., Hellman, N.N., Rogovin, S.P., Watson, P.R., Sloan, J.W., Hofreiter, B.T., Fisher, B.E., and Rist, C.R. (1954). Production of Clinical-Type Dextran. Ind. Eng. Chem. *46*, 2605–2610.

Xu, Y., and Matthews, S. (2013). TROSY NMR spectroscopy of large soluble proteins. Top. Curr. Chem. *335*, 97–119.

Zhang, H.-B., Gan, W.-W., Zhang, Y.-Q., and Hu, X.-Q. (2013). Synthesis of isomalto-oligosaccharides by using recombinant dextransucrase and Hypocrea lixii dextranase. J. Chem. Pharm. Res. *5*.

Zhou, S., Dou, H., Zhang, Z., Jin, Y., Shen, Z., and Sun, K. (2013). Facile preparation and drug delivery behavior of novel dextran-based nanogels conjugated with doxorubicin via a pH-labile bond. J. Controlled Release *172*, e67–e68.

Artworks & Table Contents

Figure content

NMR figures

Table content

Supplementary information content

Supplementary figures

Supplementary tables

Abbreviations

AFM: Atomic Force Microscopy

CAZy: Carbohydrate-Active enZymes

CD: Catalytic Domain

BRS: Branching sucrase

DNS: Dinitrosalicylic acid

DP: Degree of Polymerization

DSR: Dextransucrase

GBD: Glucan Binding Domain

GBD-CD2: Glucan Binding Domain – Catalytic Domain 2 (α-(1→2) branching sucrase)

GH: Glycoside Hydrolase

GS: Glucansucrase

His: six histidine tag

HMM: High molar mass

HPAEC-PAD: High performance anion exchange chromatography with pulsed amperometric detection

HPLC: High Performance Liquid Chromatography

HPSEC: High Pressure Size Exclusion Chromatography

HSQC: Heteronuclear Single-Quantum Correlation

IMOS: Isomaltooligosaccharides (Ix, isomaltooligosaccharide of DPx)

ITC: Isothermal Calorimetry

kDa: kiloDalton

LB: Lysogeny Broth

LMM: Low molar mass

MES: 2-(N-morpholino)ethanesulfonic acid

MM: Molar Mass

NMR: Nuclear Magnetic resonance

PCR: Polymerase Chain Reaction

PDB: Protein Data Bank

SDS-PAGE: Sodium Dodecyl Sulfate PolyAcrylamide Gel Electrophoresis

SPR: Surface Plasmon Resonance

Strep: Streptavidine tag

U: Unit of enzyme activity

WT: Wild type