



Université
de Toulouse

THÈSE

En vue de l'obtention du

DOCTORAT DE L'UNIVERSITÉ DE TOULOUSE

Délivré par :

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

Présentée et soutenue par :
Frédéric Guérin

Le mercredi 28 mars 2012

Titre :

Caractérisation structurale et fonctionnelle d'amylosaccharases

ED SEVAB : Ingénieries microbienne et enzymatique

Unité de recherche :

LISBP UMR 5504 CNRS/INSA UMR 792 INRA/INSA - IPBS UMR 5089

Directeur(s) de Thèse :

REMAUD-SIMEON Magali, Professeur, INSA Toulouse

TRANIER Samuel, Maître de conférence, Univ. Paul Sabatier Toulouse III

Rapporteurs :

IMBERTY Anne, Directrice de recherche, CNRS

CZJZEK Mirjam, Directrice de recherche, CNRS

Autre(s) membre(s) du jury :

Président : GOUET Patrice, Professeur, Univ. Lyon

Examineur : ANDRE Isabelle, Directrice de recherche, CNRS

Invité : MOUREY Lionel, Directeur de recherche, CNRS

THESE

Présentée devant

L'INSTITUT NATIONAL DES SCIENCES APPLIQUEES DE TOULOUSE

En vue de l'obtention du

DOCTORAT

Spécialité : Sciences Ecologiques, Vétérinaires, Agronomiques et Bioingénieries

Filière : Ingénieries microbienne et enzymatique

Par **Frédéric Guérin**

**Caractérisation structurale et fonctionnelle
d'amylosaccharases**

Soutenue le 28 Mars 2012 devant la commission d'examen :

Directeurs de thèse :

Mme. Magali Remaud-Siméon Professeur, INSA Toulouse

M. Samuel Tranier Maître de Conférences, Univ. Paul Sabatier, Toulouse

Rapporteurs :

Mme. Anne Imberty Directrice de Recherche, CNRS

Mme. Mirjam Czjzek Directrice de Recherche, CNRS

Membres du Jury :

M. Patrice Gouet Professeur, Univ. Lyon

Mme. Isabelle André Directrice de Recherche, CNRS

M. Lionel Mourey Directeur de Recherche, CNRS

Nom : GUERIN

Prénom : Frédéric

Titre : Caractérisation structurale et fonctionnelle d'amylosaccharases

Spécialité : Science Ecologiques, Vétérinaire, Agronomique et Bioingénieries. Filière : Ingénieries microbienne et enzymatique.

Date de soutenance : 28 Mars 2012, **n° d'ordre** 1125, **nombre de page** : 247.

Jury : P. GOUET, Président – M. REMAUD-SIMEON et S. TRANIER, Directeurs de thèse – A. IMBERTY et M. CZJZEK, Rapporteurs – I. ANDRE et L. MOUREY, Examineurs.

Cette thèse a été préparée au Laboratoire d'Ingénierie des Systèmes Biologiques et des Procédés de l'INSA de Toulouse et à l'Institut de Pharmacologie et de Biologie Structurale

Résumé :

Les amylosaccharases sont des α -transglucosylases catalysant naturellement la synthèse exclusive d' α -1,4-glucanes à partir du saccharose. Ces enzymes produisent également des composés secondaires et, en particulier, des isomères du saccharose tels que le turanose et le tréhalulose.

L'objectif de cette thèse a consisté à utiliser un panel de techniques biophysiques et biochimiques afin d'étudier les amylosaccharases de *Deinococcus geothermalis* (ASDg) et *Neisseria polysaccharea* (ASNp) afin de comprendre les relations unissant la structure, la flexibilité et la fonction de ces enzymes.

La première étude rapporte la caractérisation structurale et biophysique de l'amylosaccharase la plus thermostable connue à ce jour, l'amylosaccharase de *Deinococcus geothermalis*. La structure tridimensionnelle révèle une organisation dimérique en solution, jamais rapportée pour une amylosaccharase. Grâce à l'analyse de l'interface dimérique et à des travaux d'analyse de séquences, une séquence signature de dimérisation a été identifiée. En rigidifiant la structure de l'ASDg, la structure quaternaire contribue à l'augmentation de la stabilité thermique de la protéine. La spécificité de production des isomères du saccharose par les amylosaccharases a été étudiée. Les résultats décrivent, pour la première fois, les structures de l'ASDg et de l'ASNp en complexe avec le turanose. Dans l'ASNp, les résidus clefs forcent le résidu fructosyle à adopter une conformation linéaire positionnant idéalement le O3' pour sa glucosylation expliquant la formation préférentielle de turanose par l'enzyme. Ces résidus sont absents ou placés différemment dans l'ASDg. En conséquence, l'ASDg lie principalement les formes furanoses du fructose avec un faible réseau d'interactions. La topologie du sous-site +1 permet donc différents modes de liaison du fructose en accord avec la capacité de l'ASDg à produire une plus grande quantité de tréhalulose par rapport à l'ASNp.

Dans la seconde étude, des techniques de mutagenèse à saturation et combinatoire ciblées sur les acides aminés voisins du site actif ont été utilisées pour modifier la spécificité d'accepteur de l'ASNp. Le criblage de trois bibliothèques semi-rationnelles représentant un total de 20 000 variants a permis d'isoler trois doubles mutants montrant une amélioration spectaculaire de spécificité à la fois vis-à-vis du saccharose, le substrat donneur et de l'accepteur α -allyl-*N*-acetyl-2-désoxy- α -D-glucopyranoside par rapport au type sauvage de l'ASNp. De tels niveaux d'amélioration d'activité n'ont jamais été signalés auparavant pour cette classe d'enzymes actives sur les sucres. L'analyse par cristallographie des rayons X de la structure des meilleures enzymes mutantes suivie par des simulations de dynamique moléculaire ont montré une rigidité locale du sous-site -1 couplée à une flexibilité des boucles impliquées dans la topologie du site actif. Ces faits pourraient être à l'origine des performances catalytiques accrues de ces enzymes mutantes. L'étude démontre l'importance, lors de la conception des bibliothèques de variants, de tenir compte de la conformation locale des résidus catalytiques ainsi que de la dynamique des protéines au cours du processus catalytique.

Mots-clefs : glucane-saccharase, amylosaccharase, transglucosylation, isomérisation du saccharose, turanose, tréhalulose, thermostabilité, cristallographie des rayons X, dynamique moléculaire, ingénierie semi-rationnelle, interaction protéine-ligand

Structural and functional characterization of amylosucrases

Summary:

Amylosucrases are sucrose-utilizing α -transglucosylases that naturally catalyze the synthesis of α -glucans, exclusively linked through α -1,4 linkages. Side-products and in particular sucrose isomers such as turanose and trehalulose are also produced by these enzymes.

The objective of this thesis concerned the application of biophysical and biochemical techniques to study amylosucrases from *Deinococcus geothermalis* (DgAS) and *Neisseria polysaccharea* (NpAS) in order to investigate relationships between structure, flexibility and function of these enzymes.

In the first study, we report the first structural and biophysical characterization of the most thermostable amylosucrase identified so far, the amylosucrase from *Deinococcus geothermalis*. The 3D-structure revealed a homodimeric quaternary organization, never reported before for other amylosucrases. A sequence signature of dimerization was identified from the analysis of the dimer interface and sequence alignments. By rigidifying DgAS structure, the quaternary organization is likely to participate in the enhanced thermal stability

of the protein. Amylosucrase specificity with respect to sucrose isomer formation (turanose or trehalulose) was also investigated. We report the first structures of the DgAS and NpAS in complex with turanose. In NpAS, key residues were found to force the fructosyl moiety to bind in an open state with the O3' ideally positioned to explain the preferential formation of turanose by NpAS. Such residues are either not present or not similarly placed in DgAS. As a consequence, DgAS binds the furanose tautomers of fructose through a weak network of interactions to enable turanose formation. Such topology at subsite +1 is likely favoring other possible fructose binding modes in agreement with the higher amount of trehalulose formed by DgAS.

In the second study, iterative saturation mutagenesis and combinatorial active site saturation focused on vicinal amino acids were used to alter the acceptor specificity of NpAS and sort out improved variants. From the screening of three semi-rational sub-libraries accounting in total for 20,000 variants, we report here the isolation of three double-mutants displaying a spectacular specificity enhancement towards both sucrose, the donor substrate, and the α -allyl-*N*-acetyl-2-deoxy- α -D-glucopyranoside acceptor compared to wild-type *N. polysaccharea* amylosucrase. Such levels of activity improvement have never been reported before for this class of carbohydrate-active enzymes. X-ray structural analysis of the best performing enzymes followed by Molecular Dynamics simulations showed both local rigidity of the -1 subsite and flexibility of loops involved in active site topology which both account for the enhanced catalytic performances of the mutants. The study well illustrates the importance when designing enzyme libraries of taking into account the local conformation of catalytic residues as well as protein dynamics during the catalytic process.

Keywords: glucansucrase, amylosucrase, transglucosylation, sucrose isomerization, turanose, trehalulose, thermostability, X-ray crystallography, molecular dynamic, semi-rational engineering, protein-ligand interaction

A tous ceux qui m'ont permis d'arriver jusqu'ici

Remerciements

Voilà, cette thèse est définitivement terminée et c'est avec mon inoubliable cadeau de thèse en tête que j'écris ces quelques lignes, une thèse c'est comme un saut en parachute, au début on se demande un peu ce qu'on fait là, pourquoi a-t-on eu une idée aussi saugrenue ? Mais très vite on rencontre des personnes qui vous disent à quel point c'est une aventure merveilleuse, une passion qui vous prend pour ne plus vous lâcher et on se met à y croire, à le sentir... Puis vient le moment de la formation, on nous explique comment fonctionne le matériel, où trouver ce dont on a besoin, un moment où les émotions naissent, où l'on est sûr d'une chose, c'est qu'on le fera jusqu'au bout quoiqu'il puisse se passer... Alors on prend le train en marche (ou plutôt l'avion pour rester dans la métaphore) pour un temps qui nous semble au début infini, on passe à travers tous nos doutes, nos craintes, on apprend des choses, sur nous-même, sur ce qu'on a entrepris, sur les gens qui nous entourent tout en ayant une seule certitude : une fois arrivé au bout du chemin ce sera le grand saut, ce moment où l'erreur ne sera pas permise, mais nous avons encore le temps, non ?... Le temps s'accélère, la porte s'ouvre, dernier regard, grande inspiration, on se lance, toutes les incertitudes s'envolent pour laisser place à un plaisir court mais intense où tout défile à une vitesse incroyable, on est seul et on en profite, c'est l'aboutissement de ce que l'on a entrepris qui se réalise... Ralentissement brutal, on sort de notre petit nuage on prend conscience de ce qui nous entoure, on nous pose des questions et on se rend compte qu'on a les rennes en mains, tout peut encore basculer mais on virevolte après tout on a été bien préparé et on atterrit en douceur, voilà, c'est fait et c'est plein d'émotion qu'il faut songer à remercier tout le monde alors je remercie mes deux directeurs de thèse pour ce saut en parachute, le Professeur Magali Remaud-Siméon pour sa sagesse, sa gentillesse, sa bonne humeur et de ses conseils avisés ainsi que le Docteur Samuel Tranier, qui m'a supporté durant presque cinq ans dans son bureau, en première ligne de toute mes plaintes et qui m'a toujours soutenu moralement et prodigué un enseignement de cristallographie « à la carte ». Merci à mon tandem, le Docteur Isabelle André, qui m'a fait atterrir sur une pièce de 2 €, je tiens à souligner ta patience infinie et ta disponibilité de tous les instants à mon égard même dans mes moments les plus « hermétiques ». Je remercie le Docteur Lionel Mourey et le Professeur Pierre Monsan de m'avoir accueilli dans leurs clubs de parachutisme respectifs.

Je remercie également les Docteurs Anne Imberty et Mirjam Czjzek d'avoir accepté d'être rapporteur de ce travail ainsi que le Professeur Patrice Gouet d'avoir suivi durant trois ans et demi le déroulement de cette thèse.

Je remercie le PRES de Toulouse et la région Midi-Pyrénées pour avoir financé ce travail qui, j'espère, a porté ses fruits.

Mais ce travail n'aurait pas pu aboutir sans l'aide de nombreuses personnes, tout particulièrement, Hoai Thu Chan, Sophie Barbe, Sandra Pizzut-Serin, Claire Moulis, Stéphane Emond et Elise Champion avec qui j'ai pu collaborer "à la paillasse" directement.

On aurait besoin d'un annuaire pour lister et remercier l'ensemble des membres de l'équipe de biophysique structurale de l'IPBS et de catalyse et ingénierie microbienne et enzymatique du LISBP, je vous remercie pour avoir su faire de ces laboratoires un lieu de travail agréable où il fait bon « manipuler ».

Je remercie particulièrement Sylvianne pour tous ses *conseils de culture*, Valérie pour m'avoir fait découvrir un *autre monde*, Juliette et Frisouille ma Fripouille mes *colocs* de thèse, les *scientilivresque* Zouzou et Virginie ainsi que Fabien pour son *sens de l'observation* unique.

Il me vient en tête une bonne petite soirée qui m'a bien remonté le moral au moment où j'en avais besoin, je fais une spéciale dédicace aux membres de la *Glouglou Team* Marlène, Yannick, Manue et Karine, merci pour votre amitié, votre soutien et tous ces bons moments passés aussi bien à l'intérieur qu'à l'extérieur du labo, quelques lettres vous sont spécialement adressées sur cette page.

Tout les mercis du monde ne suffiraient pas à exprimer ma gratitude envers ma famille, mes amis et Bayonetta pour m'avoir soutenu et encouragé tout au long de mes études.

Outre ma thèse, je remercie Julie Arneau et mon tandem Vincent Noré pour ce saut en parachute (sans métaphore cette fois) qui vient achever cette thèse et là je me dis que la boucle des remerciements est bouclée.

Sommaire

Liste des Abréviations	19
Introduction	23
Chapitre 1 : Etude Bibliographique	27
Partie I : Généralités sur les Amylosaccharases	29
I-1 Historique	31
I-2 Une activité catalytique originale dans la famille 13 des Glycoside-Hydrolases	32
I-3 Traits structuraux communs de la famille GH13	35
I-4 Mécanisme d'action catalytique des enzymes de la famille GH13	38
Partie II : Relations Structure-Fonction des Amylosaccharases.....	43
II-1 La sous-famille des glycoside-hydrolases GH13.4.....	45
II-2 Production et purification des amylosaccharases recombinantes	45
II-3 Propriétés biochimiques	47
II-3.1 Profils de produits	47
II-3.2 Effet du glycogène	49
II-3.3 Un large panel d'applications	50
II-4 Structure	51
II-4.1 Structure de l'ASNp sous forme <i>apo</i>	51
II-4.2 Structure du complexe ASNp E328Q-saccharose.....	54
II-4.3 Structure de l'intermédiaire covalent β -D-glucosyl-enzyme	56
II-4.4 Structure du complexe ASNp E328Q-maltoheptaose.....	57
II-5 Relations structure-activité	60
II-6 Réaction de disproportionnation	61
Partie III : Ingénierie des Amylosaccharases.....	65
III-1 Amélioration de l'activité et de la stabilité de l'ASNp	67
III-2 Modification de la spécificité de l'ASNp.....	68
III-2.1 Stratégie de construction des bibliothèques de variants.....	69
III-2.2 Caractérisation des propriétés biochimiques des mutants de spécificité modifiée.....	72

Partie IV : Structure, Stabilité des Protéines et Interactions Protéine-Ligand	75
<i>IV.1</i> Le dichroïsme circulaire.....	77
<i>IV.2</i> La fluorimétrie différentielle à balayage.....	80
<i>IV-3</i> Cristallographie des protéines par diffraction des rayons X.....	84
<i>IV-3.1</i> Introduction à la résolution de structure de protéines par cristallographie des rayons X	84
<i>IV-3.2</i> Cristallisation des protéines	85
<i>IV-3.3</i> Diffraction des rayons X.....	89
<i>IV-3.4</i> Traitement des données de diffraction.....	91
<i>IV-3.5</i> Phasage.....	93
<i>IV-4</i> Modélisation moléculaire	97
<i>IV-4.1</i> La mécanique moléculaire	98
<i>IV-4.2</i> La dynamique moléculaire.....	99
<i>IV.5</i> La calorimétrie de titrage isotherme	101
Objectifs de la Thèse.....	109
Références Bibliographiques	113
Chapitre 2 : Etude de la Thermostabilité et de la Spécificité de Produits de l'Amylosaccharase de <i>Deinococcus geothermalis</i>	123
Summary	127
Introduction	128
Experimental procedures.....	129
Results and discussion.....	135
Overall structure and dynamics	135
Oligomerization and thermostability	139
Specificity toward sucrose isomer formation	143
Supplementary informations	150
References.....	158

Chapitre 3 : Mutants de l'amylosaccharase de <i>Neisseria polysaccharea</i> adaptés à la glycosylation d'accepteurs non naturels.....	163
Introduction	167
Experimental procedures.....	170
Results and discussion	177
Biochemical characterization of improved amylosucrase variants.....	179
Impact of mutations and ligand on protein stability	182
Structural insight on improved acceptor specificity	184
Role of mutations on amylosucrase flexibility	187
Conclusions.....	190
Supplementary informations	192
References.....	198
Conclusions Générales et Perspectives.....	201
Documentations Annexes I Résultats Complémentaires.....	209
1 Effet du tampon sur la stabilité de l'ASNp	211
2 Structure de l'ASNp cristallisée en tampon HEPES.....	212
3 Structure basse résolution de l'ASDg	213
4 Détermination d'un nouveau conditionnement pour la cristallisation du mutant I228Y de l'ASNp par DSF.....	214
5 Analyse des interactions protéine-ligand par ITC.....	215
Documentations Annexes II.....	217
Table des illustrations	241

Liste des Abréviations

A. macleodii : *Alteromonas macleodii*

ANR : Agence Nationale de la Recherche

AS : Amylosaccharase (amylosucrase en anglais)

ASAm : Amylosaccharase d'*Alteromonas macleodii*

ASDg : Amylosaccharase de *Deinococcus geothermalis*

ASDr : Amylosaccharase de *Deinococcus radiodurans*

ASNp : Amylosaccharase de *Neisseria polysacharea*

ASNm : Amylosaccharase de *Neisseria meningitidis*

ATCC : American Type Collection Culture

CAZy : Carbohydrate-Active enZYmes

CCD : Charged Coupled Device

CCP4 : Collaborative Computational Project No. 4

CDase : CyclomaltoDextrinase

CGTase : Cyclodextrine GlycosylTransferase

D : *N*-acetyl-glucosamine

D' : α -allyl-*N*-acetyl-glucosamine

D. geothermalis : *Deinococcus geothermalis*

D. radiodurans : *Deinococcus radiodurans*

DM : Dynamique moléculaire

DP : Degré de Polymérisation

DSF : Differential Scanning Fluorimetry

DSM : Deutsche Sammlung von Mikroorganismen

DTT : DiThioThréitol

ED : α -D-glucopyranosyl-(1,4)-*N*-acetyl-glucosamine

ED' : α -D-glucopyranosyl- α -(1,4)-allyl-*N*-acetyl-glucosamine

EED' : α -D-glucopyranosyl- α -(1,4)-D-glucopyranosyl- α -(1,4)-allyl-*N*-acetyl-glucosamine

E.C. : Enzyme Commission number

EDTA : acide Ethylène Diamine Tétra Acétique

ESRF : European Synchrotron Radiation Facility

fru : fructose

HPAEC-PAD : High Performance Anion Exchange Chromatography with Pulsed Amperometric Detection

GH : Glycoside-Hydrolase

glc : glucose

GST : Glutathion S Transférase

HEPES : acide hydroxyéthyle-1-piperazineéthanesulfonique

His₆ : étiquette à 6 histidines
HPLC : Chromatographie Liquide Haute Performance
IPBS : Institut de Pharmacologie et de Biologie Structurale
IPTG : Isopropyl β-D-1-thiogalactopyranoside
ITC : Isothermal titration calorimetry
K_a : constante d'association
K_d : constante de dissociation
KCTC : Korean Collection for Type Cultures
LB : milieu Luria Bertani (ou Lysogeny Broth en anglais)
LISBP : Laboratoire d'Ingénierie des Systèmes Biologiques et Procédés
LPS : LipoPolySaccharide
MES : acide 2-(N-morpholino)ethanesulfonique
MD : Molecular Dynamic
N. polysaccharea : *Neisseria polysaccharea*
nd : Non Déterminé
O-Ag : O-Antigen
OB : Oligosaccharide Binding site
PCR : Polymerase Chain Reaction
PDB : Protein Data Bank
PEG : PolyEthylène Glycol
R : facteur d'accord cristallographique
RFU : Relative Fluorescence Unit
RM : Remplacement Moléculaire
S. flexneri : *Shigella flexneri*
SAXS : Small Angle X-ray Scattering
SEC-MALLS : Size Exclusion Chromatography with Multi-Angle Laser Light Scattering
SB : Sucrose Binding site
SDS-PAGE : Sodium Dodecyl Sulfate PolyAcrylamide Gel Electrophoresis
T : Température
T_m : Température de dénaturation
Tris : 2-amino-2-hydroxyméthyl-1,3-propanediol
U : une unité U correspond à 1 μmol de fructose libéré par minute
UV : Ultra Violet

Introduction

Les amylosaccharases (E.C. 2.4.1.4) sont des transglucosidases d'origine bactérienne appartenant à la famille 13 des glycoside-hydrolases selon la classification CAZy (1). Ces enzymes catalysent la synthèse d'amylose à partir de saccharose comme unique substrat donneur d'unités glucosyle. Il existe différentes amylosaccharases possédant chacune des propriétés variées en fonction de leur origine. Ces propriétés peuvent être exploitées pour la synthèse de divers types de molécules : amylose, maltooligosaccharides ou glucodérivés (2). De plus, de récents travaux d'ingénierie ont abouti à la création d'amylosaccharases artificielles dotées de nouvelles propriétés catalytiques, les rendant ainsi capables de produire de nouvelles molécules glucidiques (3). L'étude approfondie de ces amylosaccharases naturelles ou artificielles devrait permettre, à terme, de mieux contrôler la synthèse de mono- ou polysaccharides et d'étendre le potentiel de ces enzymes à la synthèse de glucodérivés non naturels.

Dans ce cadre, cette thèse vise à élucider les facteurs moléculaires impliqués, d'une part, dans les différences de propriétés observées chez les amylosaccharases d'origine naturelle, et, d'autre part, dans les changements de spécificité des enzymes artificielles. A cette fin, nous avons utilisé différentes techniques expérimentales et théoriques (biochimie, biologie moléculaire, cristallographie des rayons X, dichroïsme circulaire, microcalorimétrie, fluorimétrie, bioinformatique, modélisation et dynamique moléculaire, ...). Conduits en cotutelle entre deux laboratoires toulousains, le Laboratoire d'Ingénierie des Systèmes Biologiques et des Procédés (LISBP) et l'Institut de Pharmacologie et de Biologie Structurale (IPBS), nos travaux ont permis d'aborder la problématique sous des angles différents mais néanmoins complémentaires afin de mieux comprendre les relations intimes entre la structure, la dynamique et l'activité des enzymes de cette famille et, à terme, permettre de mieux guider le modelage "à façon" de néo-amylosaccharases dotées de nouvelles propriétés.

La thèse débute par une étude bibliographique scindée en quatre parties. La première partie vise à décrire les enzymes de la famille 13 de la classification CAZy, et plus particulièrement, les amylosaccharases naturelles caractérisées à ce jour. La deuxième partie est consacrée à la description structurale des amylosaccharases et aux études des relations entre leur structure et leur fonction. La troisième partie dresse un bilan des travaux d'ingénierie effectués sur l'amylosaccharase de *Neisseria polysaccharea*, l'amylosaccharase la plus étudiée à ce jour. La dernière partie présente les méthodes et techniques d'étude biophysique des protéines utilisées dans le cadre de nos travaux.

Les résultats de nos investigations sont exposés dans les chapitres 2 et 3 sous forme de publications. Ainsi, le deuxième chapitre présente l'étude de la thermostabilité et spécificité de produits de l'amylosaccharase de *Deinococcus geothermalis* par des approches biochimiques et biophysiques. Le troisième chapitre est dédié aux travaux d'ingénierie réalisés sur l'amylosaccharase de *Neisseria polysaccharea* et à la caractérisation biochimique et structurale d'amylosaccharases construites "à façon" afin d'améliorer la spécificité vis-à-vis d'accepteurs non-naturels. Ce travail s'est inscrit dans le cadre d'une collaboration avec l'Institut Pasteur. Il visait à générer des enzymes de glycosylation adaptées à la production d'intermédiaires glycosidiques intervenant dans la synthèse chimio-enzymatique de motifs antigéniques. La caractérisation approfondie des meilleurs mutants obtenus a été effectuée afin de mieux comprendre les gains de spécificité. Le quatrième chapitre présente des résultats complémentaires non inclus dans les précédents chapitres. Enfin, le manuscrit se termine par la conclusion générale et les perspectives à nos travaux qui peuvent être envisagées à court ou moyen termes.

En annexe est présenté un troisième article auquel j'ai été associé et qui traite de la conformation en solution des polymères d'amyloses synthétisés par l'amylosaccharase de *Neisseria polysaccharea* et de *Deinococcus geothermalis* par Diffusion de rayons X aux petits angles (DXPA, SAXS en anglais pour Small Angle X-ray Scattering). L'étude inclut également une étude en solution de ces deux enzymes.

Chapitre 1 : Etude Bibliographique

Partie I : Généralités sur les Amylosaccharases

I-1 Historique

En 1946, Hehre et Hamilton sont les premiers à isoler une souche appartenant à l'espèce *Neisseria perflava* capable de synthétiser un α -glucane en l'absence de sucres activés de type nucléotide-sucres et en utilisant comme seul substrat le saccharose (4, 5). Puis il a été montré que l' α -glucane produit présentait une majorité de liaisons glucosidiques de type α -1,4, liaison la plus représentée dans l'amidon. L'enzyme fut donc nommée "amylosaccharase". L'activité "amylosaccharase" a été ensuite identifiée dans d'autres espèces du genre *Neisseria* isolées du microbiote de la cavité buccale (6). En particulier, Riou et coll. ont isolé en 1983, dans le rhinopharynx d'enfants sains d'Europe et d'Asie, une autre espèce qui sécrétait une amylosaccharase, *Neisseria polysaccharea* (6). Cette espèce fut utilisée une quinzaine d'années plus tard comme source d'amylosaccharase recombinante. Le gène codant l'amylosaccharase de *N. polysaccharea* a ainsi été cloné dans *Escherichia coli* (2). L'analyse de la structure primaire de la protéine a permis de révéler que l'enzyme, jusqu'alors classée par l'Enzyme Commission sous le n° E.C. 2.4.1.4, appartient à la famille 13 des glycoside-hydrolases de la classification CAZy (7). Ces travaux ont été à l'origine de la résolution de la structure tridimensionnelle de l'amylosaccharase de *N. polysaccharea*, seule structure tridimensionnelle répertoriée dans la Protein Data Bank (PDB) au début de nos travaux (8). La recherche de séquences homologues à celle de l'amylosaccharase de *N. polysaccharea* a permis de repérer de nombreux autres organismes potentiellement producteurs de cette enzyme comme nous le verrons dans le paragraphe II-1.

Sur le plan fonctionnel, seules les amylosaccharases de *N. perflava*, *N. polysaccharea*, *Deinococcus radiodurans*, *Deinococcus geothermalis* et *Alteromonas macleodii* étaient caractérisées au début de nos travaux. A partir de la seule molécule de saccharose et de la formation d'un intermédiaire glucosyl-enzyme, ces enzymes catalysent trois réactions principales, selon un mécanisme de rétention de la configuration présenté dans le paragraphe I-4 (Figure 1) :

- ✓ l'hydrolyse du saccharose,
- ✓ l'isomérisation du saccharose,
- ✓ la formation d' α -1,4 glucane par réaction de polymérisation.

La formation de polymère de glucose est la principale réaction catalysée.

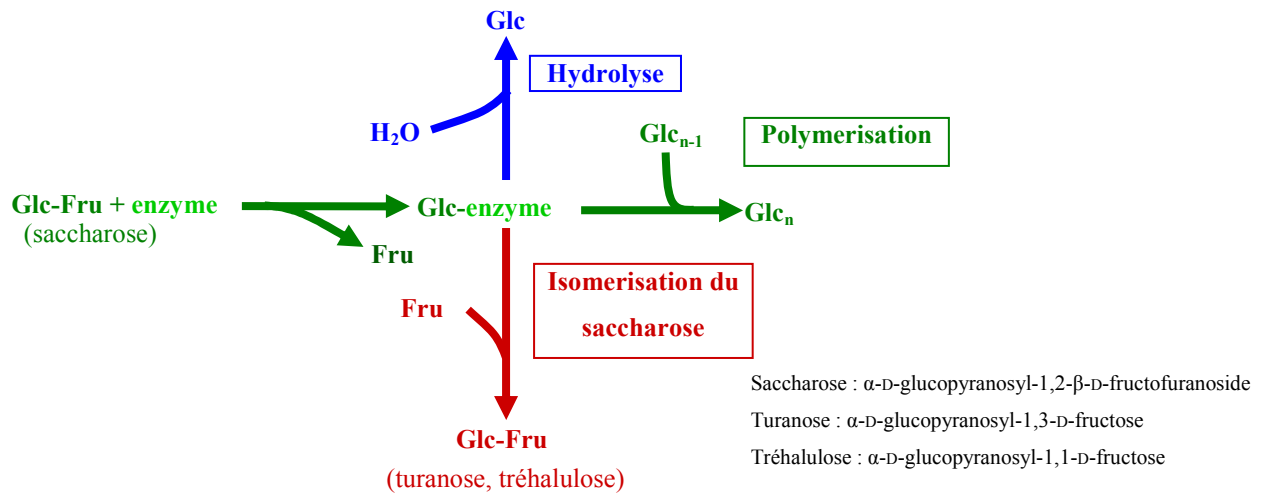


Figure 1 : Représentation schématique des différentes réactions catalysées par les amylosaccharases

En vert est représentée la réaction de polymérisation, en bleu la réaction d'hydrolyse et en rouge la réaction d'isomérisation du saccharose
Avec : Glc : glucose, Fru : fructose

A titre d'exemple, les produits synthétisés par l'amylosaccharase de *N. polysaccharaea* se répartissent de la manière suivante :

- ✓ glucose, 4 %,
- ✓ isomères du saccharose, 19 %,
- ✓ maltooligosaccharides solubles, 20 %,
- ✓ maltooligosaccharides insolubles, 57 %.

Ces réactions réalisées en présence de saccharose 106 mM ; les pourcentages correspondent à la fraction du glucose provenant du saccharose incorporée dans les différents produits.

I-2 Une activité catalytique originale dans la famille 13 des Glycoside-Hydrolases

Les amylosaccharases sont classées dans la famille 13 des glycoside-hydrolases (GH) de la classification CAZy. Cette classification regroupe les enzymes sur la base de leurs homologues de séquences reflétant des traits structuraux et mécanistiques communs (1, 9–12). La famille 13 des glycoside-hydrolases est communément appelée famille des " α -amylases". C'est la plus large des 128 familles des glycoside-hydrolases. Elle comprend plus de 10 000 enzymes qui catalysent le clivage de liaisons α -glucosidiques par un même mécanisme de

rétenion de la configuration anomérique et partagent un même repliement du type tonneau (β/α)₈. L'analyse des séquences des enzymes de cette famille par Stam et coll. (13) a permis de proposer une division de la famille 13 en 35 sous-familles dotées d'activités distinctes. Cette sous-classification comprend aussi des transporteurs d'acides aminés répertoriés dans les sous-familles GH13.34 et GH13.35. En effet, Gabrisko et coll. (14) ont mis en évidence des similitudes entre les chaînes lourdes de transporteurs d'acides aminés possédant un tonneau (β/α)₈ et certaines enzymes comme les pullulanases justifiant leur classement dans la famille GH13.

Si la famille 13 réunit principalement des enzymes impliquées dans la biotransformation de l'amidon, classées par l'Enzyme Commission dans la catégorie des hydrolases (E.C. 3) ou des transférases (E.C. 2), s'y trouvent aussi représentées des isomérases (E.C. 5).

De plus, on trouve aussi au sein de cette famille plusieurs enzymes utilisant le saccharose comme substrat. Parmi ces dernières, se distinguent les sucrose-phosphorylases, sucrose-isomérases, sucrose-hydrolases et enfin les amylosaccharases, seules glucane-saccharases de la famille capable de polymériser les unités glucosyle provenant du saccharose pour produire un polymère de type amylose (Tableau 1). A titre d'information, la famille 70 des glycoside-hydrolases qui est apparentée à la famille 13 et appartient au même clan GH-H réunit d'autres glucane-saccharases de beaucoup plus haute masse moléculaire que les amylosaccharases. Ces enzymes synthétisent des α -glucanes présentant des liaisons glucosidiques de type α -1,6, α -1,3, ou α -1,2 et une structure tridimensionnelle beaucoup plus complexe (15–17). Pour revue, il est possible de se référer aux articles de Monsan et coll. (18) et André et coll. (19). Nous avons fait le choix ici de limiter notre étude bibliographique à la famille 13.

Tableau 1 : Les différentes enzymes constituant la famille 13 des Glycoside-Hydrolases

Enzyme	Classification E.C.	Activité, d'après ExplorEnz http://www.enzyme-database.org/	Exemple de structure
α-amylase	3.2.1.1	Endohydrolyse de liaisons (1→4)- α -D-glucosidiques de polysaccharides contenant 3 ou plus d'unités glucosyles	3REN (20)
pullulanase	3.2.1.41	Hydrolyse de liaisons (1→6)- α -D-glucosidiques du pullulane, de l'amylopectine et du glycogène	2WAN (21) 3FAW (22)
cyclomaltodextrine glucanotransferase	2.4.1.19	Cyclisation de partie de chaîne (1→4)- α -D-glucane par formation d'une liaison (1→4)- α -D-glucosidique	1A47 (23) 1PAM (24)
cyclomaltodextrinase	3.2.1.54	cyclomaltodextrine + H ₂ O = maltodextrine linéaire	1H3G (25)
trehalose-6-phosphate hydrolase	3.2.1.93	trehalose 6-phosphate + H ₂ O = D-glucose + D-glucose-6-phosphate	
oligo-α-glucosidase	3.2.1.10	Hydrolyse de liaisons (1→6)- α -D-glucosidiques d'oligosaccharides produits à partir du clivage d'amidon et de glycogène par une α -amylase	3AJ7 (26)
maltogenic amylase	3.2.1.133	Hydrolyse de liaisons (1→4)- α -D-glucosidiques depuis l'extrémité non réductrice en libérant du maltose	4AEE (27)
neopullulanase	3.2.1.135	Hydrolyse du pullulane en panose (6- α -D-glucosylmaltose)	1J0H (28)
α-glucosidase	3.2.1.20	Hydrolyse de liaisons (1→4)- α -D-glucosidiques depuis l'extrémité non réductrice en libérant du D-glucose	2ZE0 (29)
maltotetraose-forming α-amylase	3.2.1.60	Hydrolyse de liaisons (1→4)- α -D-glucosidiques depuis l'extrémité non réductrice en libérant du maltotetraose	2AMG (30)
isoamylase	3.2.1.68	Hydrolyse des branchements (1→6)- α -D-glucosidiques de l'amylopectine et du glycogène	2WSK (31) 1BF2 (32)
glucodextranase	3.2.1.70	Hydrolyse de liaisons (1→6)- α -D-glucosidiques de dextrans	2ZIC (33)
maltohexaose-forming α-amylase	3.2.1.98	Hydrolyse de liaisons (1→4)- α -D-glucosidiques depuis l'extrémité non réductrice en libérant du maltohexaose	1WP6 (34)
maltotriose-forming α- amylase	3.2.1.116	Hydrolyse de liaisons (1→4)- α -D-glucosidiques depuis l'extrémité non réductrice en libérant du maltotriose	
branching enzyme	2.4.1.18	Transfert d'un segment d'une chaîne (1→4)- α -D-glucane sur un groupement alcool primaire d'une chaîne similaire	3AML (35) 3K1D (36)
trehalose synthase	5.4.99.16	maltose → α,α -trehalose	1IV8 (37)
4-α-glucanotransferase	2.4.1.25	Transfert d'un segment d'une chaîne (1→4)- α -D-glucane sur une nouvelle position d'un accepteur qui peut être un glucose ou une chaîne (1→4)- α -D-glucane	1LWH (38)
maltopentaose-forming α-amylase	3.2.1.-	Endohydrolyse de liaisons (1→4)- α -D-glucosidiques dans des polysaccharides contenant 3 liaisons (1→4)- α -D-glucosidiques ou plus	
amylosaccharase	2.4.1.4	saccharose + (α -D-glucopyranosyl-(1→4)- α -D-glucopyranoside) _n = D-fructose + (α -D-glucopyranosyl-(1→4)- α -D-glucopyranoside) _{n+1}	1G5A (39)
sucrose phosphorylase	2.4.1.7	α -D-glucopyranosyl-(1→2)- β -D-fructofuranoside + phosphate = D-fructose + α -D-glucose-1-phosphate	1R7A (40)
malto-oligosyltrehalose trehalohydrolase	3.2.1.141	Hydrolyse de liaisons (1→4)- α -D-glucosidiques du 4- α -D-[(1→4)- α -D-glucanoyl] _n -trehalose pour libérer du trehalose et un (1→4)- α -D-glucane	2BHU (41)
isomaltulose synthase	5.4.99.11	α -D-glucopyranosyl-(1→2)- β -D-fructofuranoside → α -D-glucopyranosyl-(1→6)- β -D-fructofuranose	1M53 (42)
saccharose hydrolase	3.2.1.-	Saccharose + H ₂ O → glucose + fructose	3CZE (43) 2WPG (44)
Transporteur d'acides aminés			

I-3 Traits structuraux communs de la famille GH13

L'architecture des enzymes de la famille, révélée par cristallographie et diffraction des rayons X, fait apparaître une organisation commune en trois domaines distincts : les domaines A, B et C.

Le domaine A organisé en tonneau (β/α)₈ renferme le site catalytique en forme de sillon ou de poche selon les enzymes. Le site actif possède différents sous-sites adaptés à l'arrimage des différentes unités glycosyle des substrats et/ou produits dont le nombre peut varier (45) en fonction de l'enzyme. Par convention, ces sous-sites sont numérotés de -n (à l'extrémité non-réductrice) à +n (extrémité réductrice) (46).

La rupture de la liaison osidique s'opère entre les sous-sites -1 et +1 (47) par un mécanisme de rétention de la configuration faisant intervenir deux acides carboxyliques catalytiques dont le rôle sera décrit plus précisément dans le paragraphe I-4 (Figure 2). D'autres résidus impliqués dans la stabilisation du ligand au niveau du sous-site -1 sont également très conservés. Il s'agit d'un deuxième acide aspartique, de deux histidines et d'une arginine participant à la stabilisation de l'état de transition et d'une tyrosine participant à la stabilisation des complexes enzyme-substrat par des interactions d'empilement avec les unités osidiques (48).

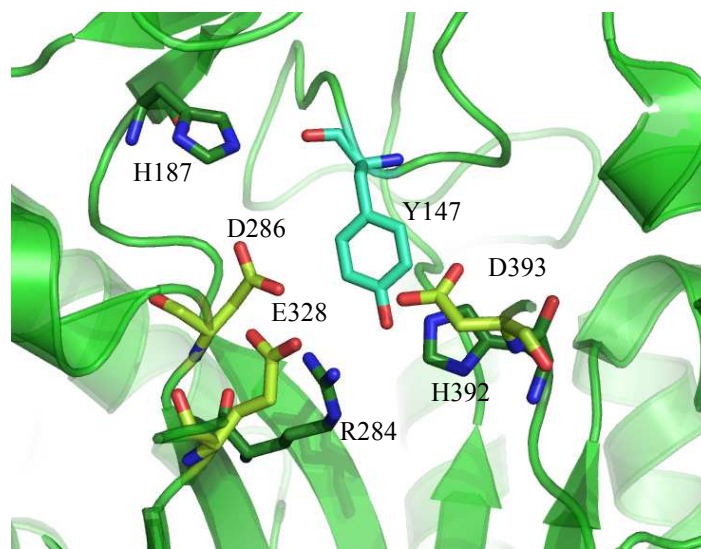


Figure 2 : Représentation tridimensionnelle des résidus d'acides aminés conservés dans les sous-sites -1 et +1 des enzymes de la famille GH13

*La numérotation des résidus est en accord avec la séquence de l'amylosaccharase de *Neisseria polysaccharea* (code PDB 1G5A) (39, 49). Avec, en vert clair : la triade catalytique, en turquoise : la tyrosine impliquée dans les interactions d'empilement, en vert foncé : les histidines et l'arginine impliquées dans la stabilisation du complexe intermédiaire. Image générée avec Pymol (84)*

Des travaux d'alignements de séquences réalisés par Jespersen et coll. (50) et par Sarçabal et coll. (51) ont mis à jour plusieurs zones très conservées du tonneau catalytique (β/α)₈ et qui constituent la signature de la famille GH13 (Figure 3).

	I β 2	II β 3	III L3	IV L3	V β 4	VI β 5	VII β 7	VIII β 8
AS	134-GLTYLHLM-P-142	190-DFIFNH-195		262-QWDLN-266	290-ILRMDAVAF-298		396-YVR--SHD-401	488-GLPLIYLGD-496
AMY	36-GFSGVQVS-P-44	96-DAVINH-101	150-SYND-153	165-LDLA-169	193-GFRIDASKH-201	233-EVID-236	295-FVD--NHD-300	334-GFTRVMSSY-342
OGL	44-GIDVWLS-P-52	98-DLVVNH-103	150-QYDE-153	167-QPDLN-171	195-GFRMDVINF-203	255-EMPG-258	324-YNN--NHD-329	360-GTPYIYQGE-368
AGL	52-GVDAIWC-P-60	106-DLVVNH-111	164-TFDE-167	181-QVDLN-185	210-GFRIDTAGL-218	276-EVAH-279	344-YIE--NHD-349	381-GTLYVYQGG-389
PUL	210-GVTHVELL-P-218	281-DVVVNH-286	308-AYGN-311	319-GNDIA-323	348-GFRFDLMGI-356	381-EGWD-384	464-YVE--SHD-469	505-GIPFLHSGQ-513
APU	435-GISVIYLN-P-443	488-DGVFNH-493	527-PYGD-530	565-WADFI-569	593-GWRDLVANE-601	626-ELWG-629	698-LLG--SHD-703	745-GMPSIYYGD-753
CMD	187-GVNALYFN-P-195	240-DAVFNH-245	284-TYDT-287	294-MPKLN-298	323-GWRDLVANE-331	356-EIMH-359	418-LLG--SHD-423	450-GTPCIYYGD-458
MTH	50-GFSAIWMPVP-59	112-DVVVNH-117	145-NYPN-148	160-ESDLN-164	189-GFRFDVVRG-197	219-ELWK-222	289-FVD--NHD-294	327-GTPVYVWSH-335
ISA	217-GVTAVFEL-P-225	291-DVVVNH-296	328-TSGN-331	341-GANFN-345	370-GFRFDLASV-378	416-EFTV-419	502-FID--VHD-507	570-GTPLMQGGD-578
DGL	44-GVMAIWS-P-52	98-DLVVNH-103	145-QYDD-148	162-QPDLN-166	190-GFRMDVIDM-198	236-ETWG-239	308-FWN--NHD-313	344-GTPYIYQGE-352
MHH	38-GITAVWLP-P-46	102-DVVVNH-107	166-DWDQ-169	203-YADID-207	232-GFRIDAVKH-240	266-EFWK-269	328-FVD--NHD-333	362-GYPSVYFGD-370
NPU	188-GINGIYLT-P-196	241-DAVFNH-246	284-NYDT-287	294-MPKLN-298	323-GWRDLVANE-331	356-EVWH-359	418-LLG--SHD-423	450-GTPCIYYGD-458
BRE	280-GPTHLELL-P-288	335-DWVPGH-340	356-LYEH-359	367-HQDWN-371	401-ALRVDVAVS-409	458-EEST-461	519-FVLPLSHD-526	555-GWMWAFPGK-563
CGT	70-GVTALWISQP-79	135-DFAPNH-140	185-SLEN-188	197-LADFN-201	225-GIRVDVAVK-233	257-EWFL-260	323-FID--NHD-328	354-GVPAIYYGT-362
GDE	137-GYNMIHFT-P-145	198-DVVVNH-203	248-KYKE-251	451-LRNFA-455	505-GVRLDNCHS-513	538-ELFT-541	604-FMD--NHD-610	642-GYDELVPHQ-650
TAA	56-GFTAIWIT-P-64	117-DVVVNH-122	155-YEDQ-158	173-LPDLN-177	202-GLRIDTVKH-210	230-EVLD-233	292-FVE--NHD-297	323-GIPIIYAGQ-331
		★			★	★	★★	

Figure 3 : Résidus conservés du tonneau catalytique des GH13

d'après Sarçabal et coll. (51)

AS : amylosaccharase de *Neisseria polysaccharea*, AMY : α -amylase pancréatique de porc, OGL : oligo-1,6-glucosidase de *Bacillus cereus*, AGL : α -glucosidase de *Saccharomyces cerevisiae*, PUL : pullulanase de *Bacillus stearothermophilus*, APU : amylopullulanase de *Clostridium thermohydrosulfuricum*, CMD : cyclomaltodextrinase de *Bacillus sphaericus*, MTH : maltotetraohydrolase de *Pseudomonas saccharophila*, ISA : isoamylase de *Pseudomonas amyloideramosa*, DGL : dextran-glucosidase de *Streptococcus mutans*, MHH : maltohexaohydrolase de *Bacillus sp. Strain 707*, NPU : neopullulanase de *Bacillus stearothermophilus*, BRE : enzyme de branchement d'*Escherichia coli*, CGT : cyclodextrin-glycosyltransferase de *Bacillus circulans*, GDE : enzyme de débranchement du glycogène humaine, TAA : α -amylase d'*Aspergillus oryzae*.

Une étude par mutagenèse dirigée guidée par ces travaux d'alignement a permis à Sarçabal et coll. d'identifier les résidus catalytiques (51). Précédemment, Jespersen et coll. (50) ont proposé que les boucles du tonneau catalytique, interviennent dans la spécificité de reconnaissance de ces enzymes.

Par ailleurs, des séquences consensus ont été récemment proposées par Kumar (Figure 4) pour les α -amylases, les cyclodextrine glycosyltransferases (CGTases), les enzymes de branchement et les cyclomaltodextrinases (CDases) (52).

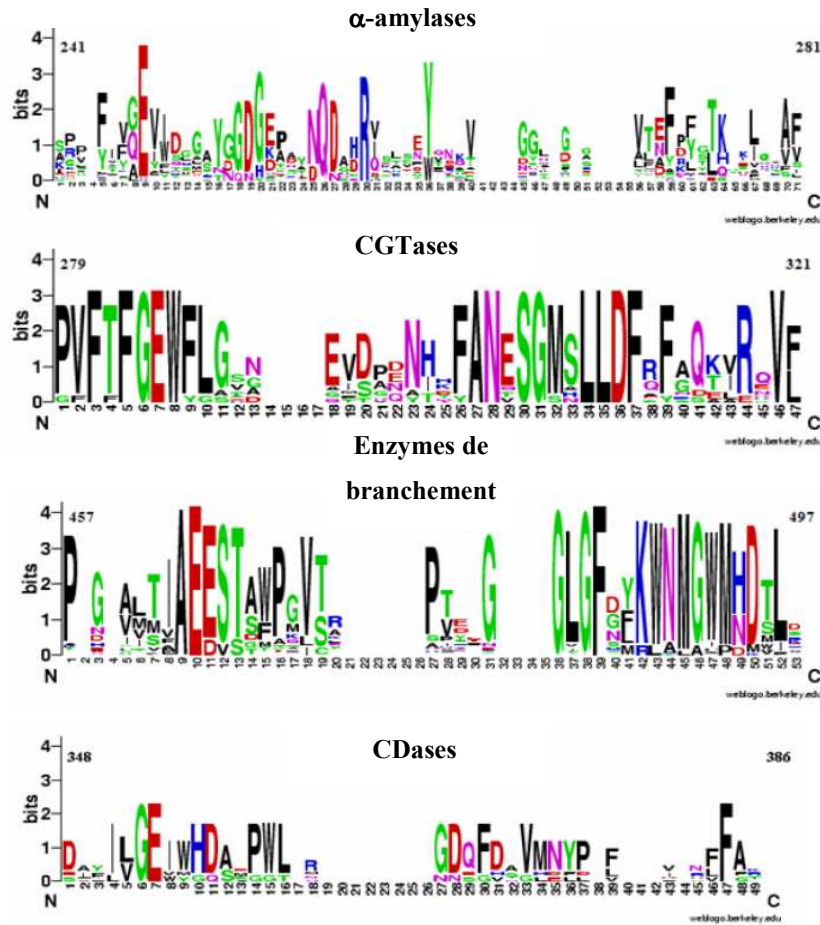


Figure 4 : Séquences consensus de GH13 aux activités différentes

d'après Kumar (52) générées avec Weblogo (53)

Le domaine B se situe entre le troisième brin β et la troisième hélice α du tonneau catalytique. Sa structure primaire et sa structure tertiaire varient grandement d'une enzyme à l'autre et de nombreux travaux indiquent qu'il joue un rôle important dans le contrôle de la spécificité de substrats ou de produits des enzymes de la famille 13 (54).

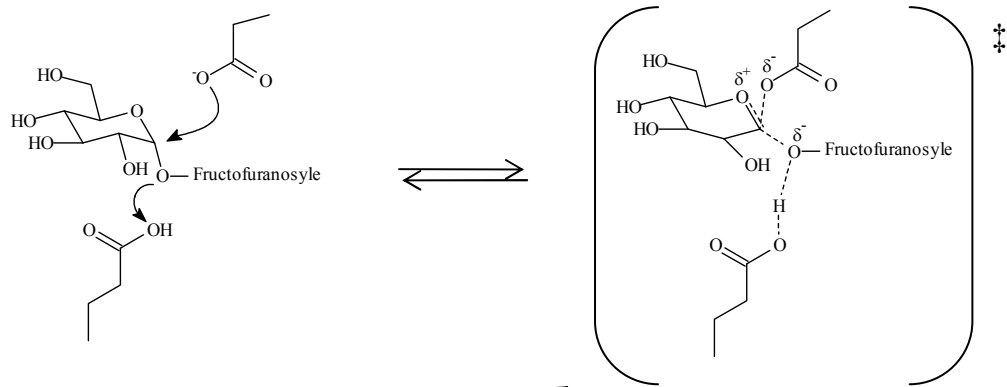
Le domaine C suit le domaine A et possède une structure tertiaire en forme de clef grecque. Sa fonction a été très peu explorée. Dans l' α -amylase du grain d'orge AMY1, ce domaine abrite un site particulier appelé la "pince à sucre" qui serait impliqué dans la fixation de chaînes maltooligosaccharidiques (55). La construction de mutants supprimant la "pince à sucre" par Bozonnet et coll. (56) et montrant une perte drastique d'affinité pour les maltooligosaccharides a permis de confirmer le rôle clef de cette "pince à sucre" dans la

fixation des chaînes maltooligosaccharidiques. Ce domaine est aussi connu pour accueillir un site de surface d'oligosaccharide (57).

I-4 Mécanisme d'action catalytique des enzymes de la famille GH13

Le clivage de la liaison osidique s'opère entre les sous-sites -1 et +1. La catalyse procède selon un double-déplacement impliquant un acide glutamique, jouant alternativement le rôle de catalyseur acide-base et un aspartate jouant le rôle de nucléophile. L'acide glutamique va céder son proton à l'oxygène de la liaison osidique tandis que le nucléophile attaque le carbone anomérique (en configuration α) pour former un intermédiaire oxocarbone qui génère l'intermédiaire covalent β -glucosyl-enzyme (d'anomérisation inversée) et la libération de l'extrémité réductrice de la molécule (58) (Figure 5). Dans une deuxième étape, l'acide glutamique déprotoné (glutamate) va activer l'hydroxyle d'une molécule acceptrice (la molécule d'eau pour la réaction d'hydrolyse ou un sucre pour la transglucosylation) qui attaquera à son tour l'intermédiaire β -glucosyl-enzyme pour former un α -glucosyl-accepteur (seconde inversion de la configuration responsable de la rétention de configuration). Un troisième résidu catalytique, (un acide aspartique) intervient conjointement avec deux histidines conservées au sein de la famille GH13, afin de stabiliser le complexe intermédiaire covalent.

Etape 1 : Formation de l'intermédiaire covalant β -D-glucosyl-enzyme



Etape 2 : Transfert du résidu β -D-glucosyl-enzyme

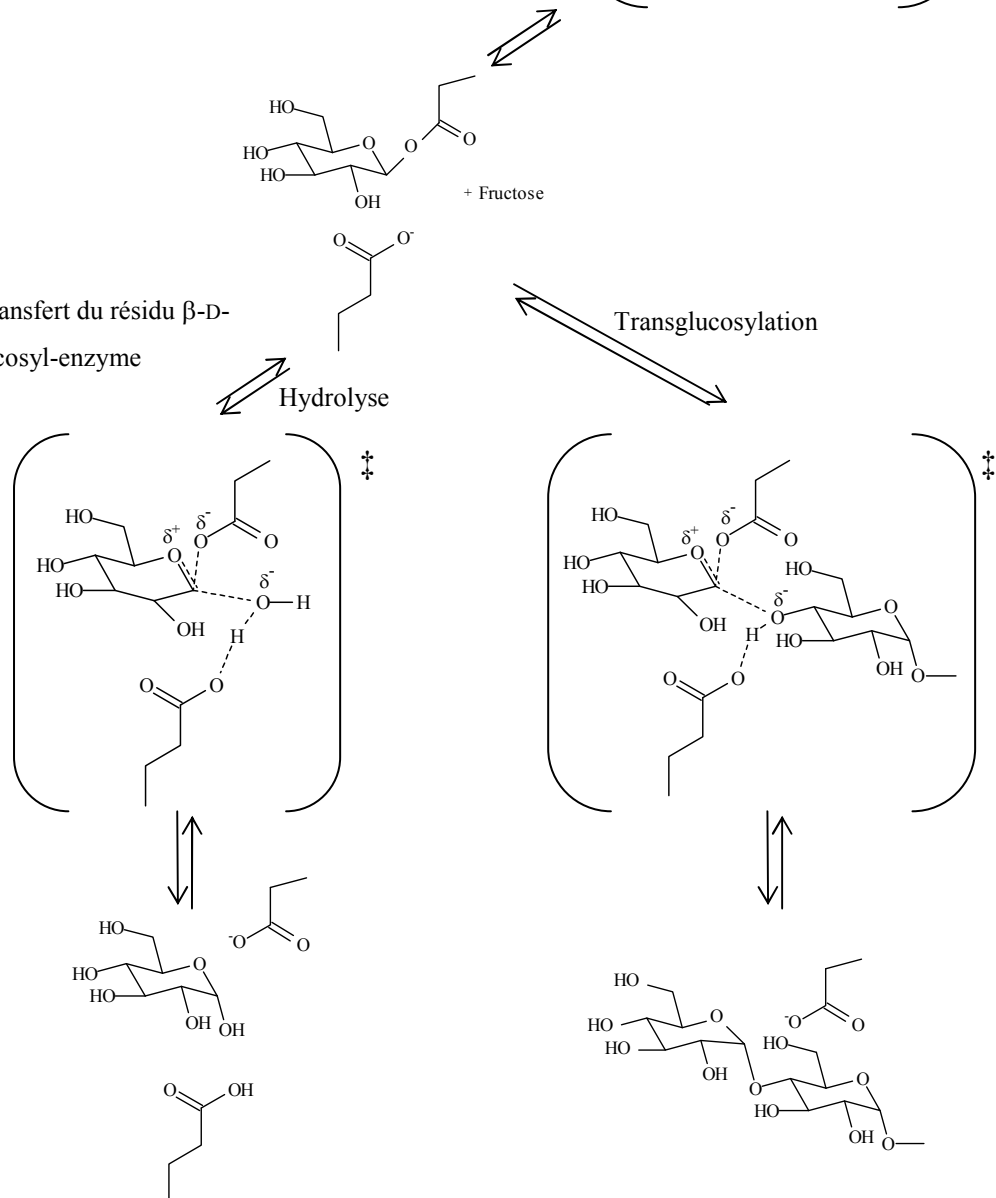


Figure 5 : Mécanisme réactionnel des glycoside-hydrolases et transglycosidases à rétention de configuration

D'après Withers et coll. (59)

Conclusion

La famille GH13 de la classification CAZy (famille des α -amylases) est la famille la plus vaste avec plus de 10 000 représentants. Cependant, ces enzymes présentent des caractéristiques communes comme le mécanisme d'action catalytique opérant par rétention de la configuration anomérique du résidu glucosyle transféré, une forte homologie structurale avec trois domaines A, B et C souvent observés, ou la conservation de résidus impliqués dans la catalyse enzymatique. Au sein de cette famille, les amylosaccharases occupent une position originale puisqu'elles sont les seules enzymes à être douées d'une activité de polymérisation du glucose. Par la suite, les amylosaccharases ont montré un véritable potentiel d'utilisation biotechnologique comme outils de transglucosylation étant donné que ces enzymes utilisent le saccharose, une agro-ressource abondante et peu onéreuse, comme substrat de base. Afin de l'exploiter au mieux, différentes études abordant les relations entre la structure et la fonction de ces enzymes ont été menées et sont exposées dans la partie suivante.

*Partie II : Relations Structure-Fonction
des Amylosaccharases*

II-1 La sous-famille des glycoside-hydrolases GH13.4

Les amylosaccharases sont regroupées dans la sous-famille GH13.4 qui réunit deux types d'enzymes :

- ✓ les amylosaccharases (E.C. 2.4.1.4), transglucosidases synthétisant un polymère de type amylose,
- ✓ les sucrose-hydrolases (E.C. 3.2.1.-) qui hydrolysent le saccharose en glucose et fructose.

Ces deux enzymes utilisent le saccharose comme substrat naturel et opèrent selon le mécanisme d'action décrit en paragraphe I-4.

La plupart des travaux portant sur les amylosaccharases mettent en œuvre des amylosaccharases recombinantes en raison principalement de la facilité de production et de purification. L'amylosaccharase de *N. polysaccharea* (ASNp) (Potocki de Montalk et coll. en 2000 (2)) reste l'amylosaccharase la plus étudiée. Au début de nos travaux, c'était d'ailleurs la seule amylosaccharase de structure connue et disponible dans la Protein Data Bank (PDB) (39, 49). Les amylosaccharases de *D. radiodurans* (ASDr) et *D. geothermalis* (ASDg) ont été étudiées plus tardivement en 2005 et 2008, respectivement (Pizzut-Serin et coll. (60), Emond et coll. (61)). Enfin l'amylosaccharase d'*A. macleodii* (ASAm) est la dernière à avoir été caractérisée en 2009 (Ha et coll. (62)). Les identités de séquences entre ces enzymes varient de 40 à 75 % (tableau 2).

Tableau 2 : Conservation de séquences entre les différentes amylosaccharases caractérisées à ce jour

	ASDg	ASDr	ASAm
ASNp	Identité : 42 % Similarité : 60 %	Identité : 43 % Similarité : 60 %	Identité : 49 % Similarité : 65 %
ASDg		Identité : 75 % Similarité : 83 %	Identité : 41 % Similarité : 59 %
ASDr			Identité : 42 % Similarité : 58 %

II-2 Production et purification des amylosaccharases recombinantes

Pour pallier les problèmes de production par les organismes natifs et faciliter la purification des enzymes, le gène codant les amylosaccharases a généralement été cloné en fusion avec

une étiquette codant pour la glutathion S-transférase (GST) ou pour une étiquette à six histidines (His₆) tableau 3.

Tableau 3 : *Bilan des amylosaccharases recombinantes caractérisées à ce jour*

Amylosaccharases de	Souche d'origine	Souche recombinante	Plasmide	Etiquette	Référence
<i>Neisseria polysaccharea</i>	ATCC* 43768	<i>Escherichia coli</i> BL21	pGEX-6P-3	GST	(63)
<i>Neisseria polysaccharea</i>	ATCC* 43768	<i>Escherichia coli</i> BL21 (DE3)	pGEX-4T-1	His ₆	(64)
<i>Neisseria polysaccharea</i>	ATCC* 43768	<i>Escherichia coli</i> BL21 (DE3)	pJEA1000	His ₆	(65)
<i>Deinococcus geothermalis</i>	DSM* 11300	<i>Escherichia coli</i> JM109	pGEX-6P-3	GST	(61)
<i>Deinococcus geothermalis</i>	DSM* 11300	<i>Escherichia coli</i> MC1061	pHCXHD	His ₆	(66)
<i>Deinococcus radiodurans</i>	ATCC* 13939	<i>Escherichia coli</i> BL21	pGEX-6P-3	GST	(60)
<i>Alteromonas macleodii</i>	KCTC* 2957	<i>Escherichia coli</i> BL21 (DE3)	pGEX-4T-1	GST	(62)

*ATCC : American Type Collection Culture / DSM : Deutsche Sammlung von Mikroorganismen / KCTC : Korean Collection for Type Cultures

Les méthodes de production et de purification des amylosaccharases recombinantes possédant l'étiquette GST (clivable par la PreScission protéase) varient peu d'une enzyme à l'autre. La souche d'*Escherichia coli* transformée avec le plasmide recombinant est cultivée en milieu LB supplémenté d'ampicilline (100 µg.mL⁻¹). Le gène codant l'amylosaccharase se trouve généralement sous la dépendance d'un promoteur inductible à l'isopropylthiogalactoside (IPTG). La concentration d'IPTG joue un rôle crucial dans le niveau d'expression des protéines et varie pour chaque enzyme (2 mM, 1 mM, 10 µM et 5 µM pour l'ASNp, l'ASAm, l'ASDg et l'ASDr, respectivement). La culture est centrifugée et le culot bactérien est resuspendu dans un tampon Tris 50 mM pH 7 NaCl 150 mM et lysé par sonication. La purification de l'enzyme est réalisée par chromatographie d'affinité en mode batch. L'échantillon est déposé sur une résine Glutathione Sepharose. L'enzyme est éluée par addition de glutathion libre ou directement par coupure de l'étiquette GST grâce à l'ajout de 250 UE de PreScission Protease (GE Healthcare) sur la résine. Il est ainsi possible de produire

facilement des quantités importantes d'enzymes purifiées (10 à 40 mg de protéines purifiées par litre de culture). D'autres protocoles de purification existent, ainsi Scheider et coll. (65) ont développé un protocole permettant d'obtenir jusqu'à 170 mg d'ASNp par litre de culture. D'autre part, Seo et coll. (66) produisent de l'ASDg, fusionnée à une étiquette à 6 histidines, de manière constitutive et dont l'activité augmente d'un facteur 145.

II-3 Propriétés biochimiques

Ces quatre amylosaccharases recombinantes (ASNp, ASDg, ASDr et ASAm), bien que relativement proches en terme de similarité de séquences, ont révélé, suite à leur caractérisation, un certain nombre de différences. Une différence majeure réside dans la thermostabilité à 50 °C de l'ASDg produite par la bactérie thermophile *D. geothermalis*, possédant de ce fait des enzymes pouvant fonctionner à des températures modérément élevées comprises entre 45 °C et 50 °C (67).

Le tableau 4 rassemble certaines caractéristiques biochimiques de ces enzymes.

Tableau 4 : Caractéristiques biochimiques des amylosaccharases caractérisées à ce jour

	Température	Activité spécifique (U.mg ⁻¹)			Temps de demi-vie			Ref
	optimale	30 °C	45 °C	50 °C	30 °C	45 °C	50 °C	
ASNp	30 °C	9,6	<i>nd</i>	<i>nd</i>	21 h	<i>nd</i>	3 min	(2)
ASDg	50 °C	19	<i>nd</i>	44	65 h	<i>nd</i>	22 h	(61)
ASDr	<i>nd</i>	4	<i>nd</i>	<i>nd</i>	15 h	<i>nd</i>	<i>nd</i>	(60)
ASAm	45 °C	~150	432	~400	<i>nd</i>	33,3 min	14,3 min	(62)

nd : Non déterminé

D'après les résultats obtenus, l'ASDg serait la plus stable à 30 °C et à 50 °C, avec des temps de demi-vie de 65 h et 22 h respectivement, alors que l'enzyme possédant la meilleure activité spécifique est l'ASAm.

II-3.1 Profils de produits

Les seuls profils de produits décrits dans la littérature à ce jour sont ceux de l'ASNp et de l'ASDg (61), représentés dans la figure 6. On distingue, parmi les produits formés par les deux enzymes, du glucose, du maltose, du maltotriose et des maltooligosaccharides de degrés de

polymérisation (DP) supérieurs à 3 et pouvant atteindre un DP voisin de 80 ainsi que des isomères du saccharose : turanose et tréhalulose, pouvant exister sous différentes formes en solution (68). Les formules développées de ces molécules sont présentées sur la figure 7. En présence de 100 mM de saccharose et à 30 °C, 62 % du glucose provenant du saccharose est incorporé dans les maltooligosaccharides de DP>3 par l'ASNp. Avec l'ASDg, l'incorporation du glucose provenant du saccharose dans les maltooligosaccharides de DP>3 est supérieure et atteint 75 %. En effet, l'ASDg produit beaucoup moins de maltotriose (~2 % contre ~20 % pour l'ASNp). En présence de 600 mM de saccharose, les différences observées sont atténuées avec des taux de production de maltooligosaccharides (DP>3) sensiblement du même ordre. En revanche, à cette concentration de 600 mM, l'ASDg produit toujours moins de maltotriose (2 % contre 8 % pour l'ASNp) mais semble produire légèrement plus de turanose (20 % contre 15 % pour l'ASNp). La production du second isomère du saccharose est anecdotique pour les deux enzymes avec moins de 5 % de tréhalulose produit en fin de réaction.

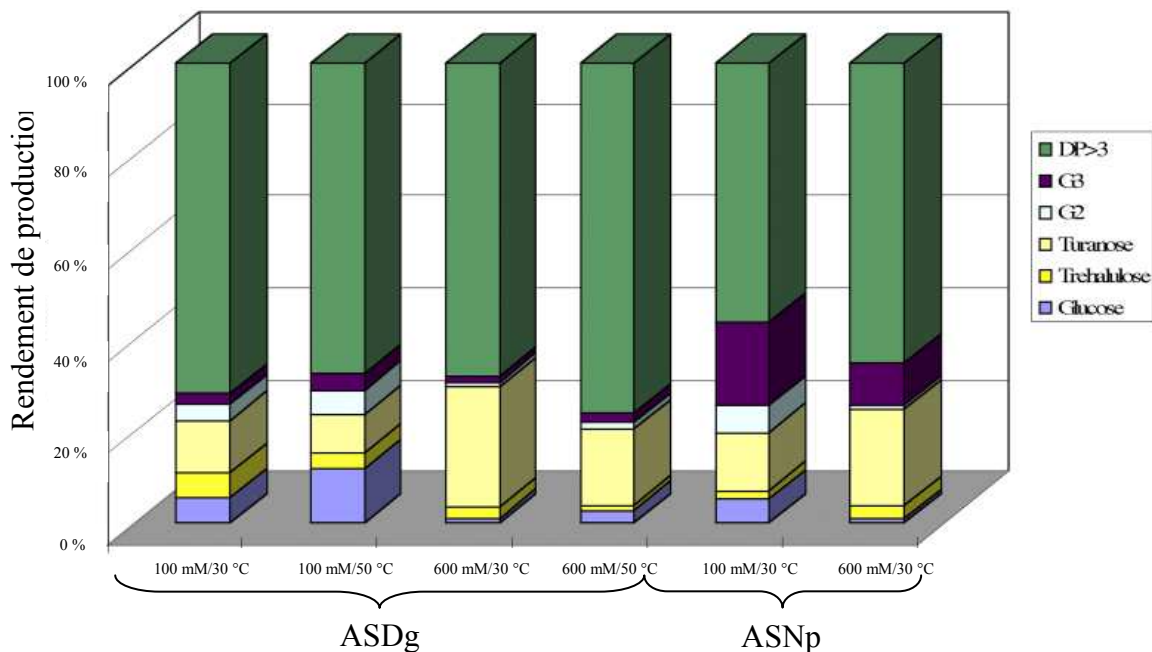


Figure 6 : Comparaison des produits formés par l'ASNp et l'ASDg en fonction des conditions réactionnelles

d'après Emond et coll. (61)

Rendement de production : pourcentage d'unités glucopyranosyle issues du saccharose transférées sur un accepteur

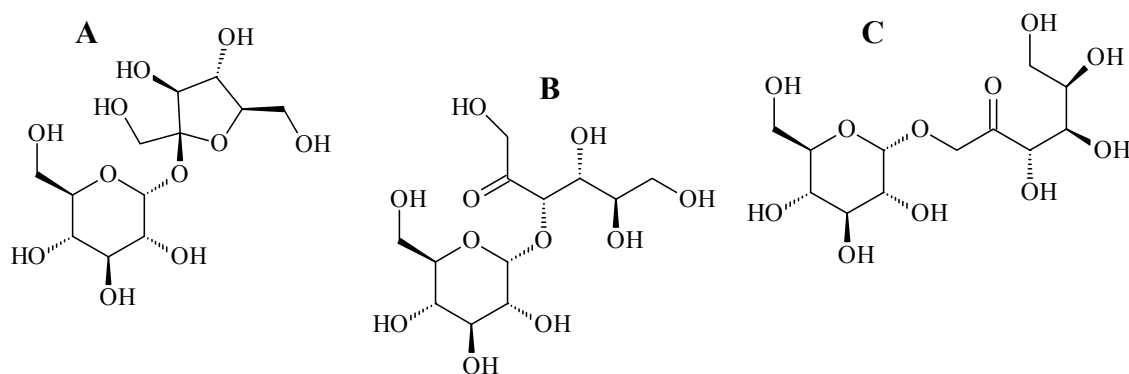


Figure 7 : Formules semi-développées des isomères du saccharose produits par les amylosaccharases

(A) : saccharose : α -D-glucopyranosyl-1,2- β -D-fructofuranose

(B) : turanose* : α -D-glucopyranosyl-1,3-D-fructose

(C) : tréhalulose* : α -D-glucopyranosyl-1,1-D-fructose

*Il peut également exister des conformations où le fructose est sous forme furanose ou pyranose (68)

II-3.2 Effet du glycogène

La vitesse de consommation du saccharose par l'ASNp en présence de glycogène exogène est fortement augmentée par rapport à celle observée en absence de glycogène (69). De plus, Potocki de Montalk et coll. ont mis en évidence un effet dose-dépendant du polymère. Ainsi en présence de 10 g.L^{-1} de glycogène (valeur à laquelle il y a saturation de l'effet activateur), la vitesse de consommation du saccharose est accélérée d'un facteur 98. Dans ces conditions, le glycogène joue le rôle d'accepteur des résidus glucosyle provenant du saccharose et l'enzyme catalyse l'élongation des ramifications du polymère (69). Cet effet activateur diminue avec la concentration en glycogène. A la concentration de $0,5 \text{ g.L}^{-1}$ de glycogène, la vitesse de consommation du saccharose est peu augmentée et les produits synthétisés correspondent à ceux formés à partir de saccharose seul. Notons que cet effet activateur du glycogène n'est pas observé pour l'ASDg.

II-3.3 Un large panel d'applications

L'amylosaccharase est une enzyme très attractive pour des applications biotechnologiques. On dénombre à ce jour près de 103 brevets sur le site de l'organisation mondiale de la propriété industrielle (OMPI) pour une recherche effectuée à partir du mot-clef "amylosucrase". Grâce à ses propriétés, l'amylosaccharase peut être utilisée pour produire des polymères de type amylose à partir de saccharose seul. En outre, le degré de polymérisation moyen, la morphologie et la taille des grains d'amylose peuvent être modulés par la simple concentration en saccharose initiale (70). De plus, l'amylosaccharase peut également modifier la structure de polymères en présence de saccharose (amylose, amylopectine, amidon, glycogène) par élongation des chaînes d' α -(1,4)-glucane (71), améliorant ainsi leur résistance aux agents de dégradation enzymatique (72). Ceci ouvre de nombreuses perspectives de modification et d'utilisation d'amidons modifiés et résistants pour le secteur agroalimentaire (73). Les amylosaccharases peuvent également être employées pour la synthèse d'isomères du saccharose, molécules édulcorantes acariogéniques (74). L'utilisation combinée de l'ASNp et d'une 4- α -glucanotransférase par Kim et coll. (75) a permis le développement d'une nouvelle voie de synthèse de cycloamylose (molécule présentant une cavité hydrophobe pouvant transporter des molécules d'intérêt alimentaire, cosmétique ou pharmaceutique).

Ces enzymes ont aussi la capacité originale de glucosyler naturellement des accepteurs hydroxylés exogènes (de nature saccharidiques ou non) introduits dans le milieu réactionnel. Ainsi les amylosaccharases recombinantes ont été employées dans le cadre de la glucosylation de la (+)-catechine (76), de l'arbutine (77) et de l'hydroquinone (66) par l'ASDg ou de la salicine (64) par l'ASNp et l'ASDg. La glucosylation pourrait conduire à une amélioration de la stabilité et de la solubilité de ces molécules tout en conservant leurs propriétés biologiques. Ces propriétés couplées à de l'ingénierie enzymatique ayant permis d'atténuer la formation de longues chaînes au profit de réactions de mono- et di-glucosylations (78) rendent les amylosaccharases fort intéressantes pour la synthèse par voie enzymatique de glucoconjugués d'intérêt (19).

Plus récemment, l'ASNp a été immobilisée de manière non covalente pour produire un polymère d'amylose insoluble (79). Wang et coll. ont constaté que l'immobilisation de l'ASNp permet d'améliorer sa thermostabilité. En outre, ce procédé d'immobilisation permet un recyclage du biocatalyseur en fin de manipulation qui peut s'avérer économiquement très intéressant pour des applications biotechnologiques et industrielles.

II-4 Structure

Le clonage de l'ASNp sous forme recombinante couplé au développement et à l'optimisation de sa purification ont permis de mener à bien la détermination de sa structure par cristallographie des rayons X. Jusqu'à très récemment, la structure tridimensionnelle de l'ASNp était la seule structure d'amylosaccharase disponible (39, 49). Le tableau 5 recense les différentes structures de l'ASNp accessibles dans la PDB, sous forme *apo* ou en complexe avec des ligands. L'obtention de complexes avec le saccharose, le maltoheptaose ou encore l'intermédiaire covalent β -D-glucosyl-enzyme n'a pu être possible que grâce à l'introduction d'une mutation sur un résidu catalytique (E328Q) inactivant l'enzyme.

II-4.1 Structure de l'ASNp sous forme *apo*

La structure primaire de l'amylosaccharase de *Neisseria polysaccharea* comporte 636 résidus dont 6 cystéines, cependant aucun pont disulfure n'est observé au sein de cette protéine. La détermination de la structure de l'ASNp sous sa forme *apo* (code PDB : 1G5A (39, 49)) a permis de mettre en lumière le repliement de la chaîne polypeptidique en 5 domaines nommés N, A, B, B' et C (Figure 8A). Les domaines A, B et C sont communs aux α -amylases, enzymes composant majoritairement la famille GH13 (cf paragraphe I-3).

Le domaine N (1-89) se compose de six hélices α , sa fonction n'a pas été élucidée. **Le domaine A**, domaine catalytique caractéristique de la famille GH13, adopte un repliement en tonneau (β/α)₈, et comprend les résidus 90-184, 261-395 et 461-553. Les différents éléments de structures secondaires du tonneau catalytique sont reliés par des boucles (Figure 8B) dont la boucle 3 (reliant le brin β 3 et l'hélice α 3 du tonneau (β/α)₈) est relativement longue (185-260) et forme un domaine à part entière : **le domaine B**. Ce domaine comporte deux brins β anti-parallèles et deux hélices α . La boucle 7 (reliant le brin β 7 à l'hélice α 7 du tonneau (β/α)₈) présente aussi une extension formant **le domaine B'** (396-460) et est exclusivement composée d'hélices α . Enfin, **le domaine C** (554-628) est essentiellement composé de brins β organisés en clef grecque.

Tableau 5 : Tableau récapitulatif des structures d'ASNp déposées dans la PDB (Octobre 2011) et de leurs conditions de cristallisation

Code PDB	Mutations Introduites	Ligand	Conditionnement	Cristallisation	Résolution (Å)	Référence
1G5A	G537D*		Tris 50 mM pH 7 NaCl 150 mM EDTA 1 mM DTT 1 mM 4-6 g.L ⁻¹	PEG6000	1,4	(39, 49)
1JG9	G537D*	D-glucose		HEPES 0,1 M pH 7	1,7	(80)
1JGI	E328Q** H131Y* G537D*	Saccharose		PEG6000 HEPES 0,1 M pH 7 saccharose 14 mM	2,0	(80)
1MW1	G537D*	Saccharose		PEG6000	2,1	(81)
1MW2	G537D*	Saccharose		HEPES 0,1 M	2,1	(81)
1MW3	G537D*	Saccharose		pH 7	2,0	(81)
1MW0	E328Q** H131Y*	Maltoheptaose		PEG6000 HEPES 0,1 M pH 7	2,0	(81)
1MVY	E328Q** H131Y*	Maltoheptaose		maltoheptaose 14 mM	2,0	(81)
1ZS2	E328Q** H131Y*	Saccharose Maltoheptaose		PEG6000 HEPES 0,1 M pH 7 saccharose 14 mM maltoheptaose 14 mM	2,2	(82)
1S46	E328Q** H131Y*	α-D-glucosyl lié covalamment		PEG6000 HEPES 0,1 M pH 7	2,2	(83)

* Les structures déposées sont susceptibles de contenir deux mutations fortuites : H131Y et G537D. Ces mutations sont dues à des erreurs de PCR mais n'ont, *a priori*, aucune influence sur l'activité de l'enzyme.

** Mutation responsable de la perte d'activité de l'enzyme.

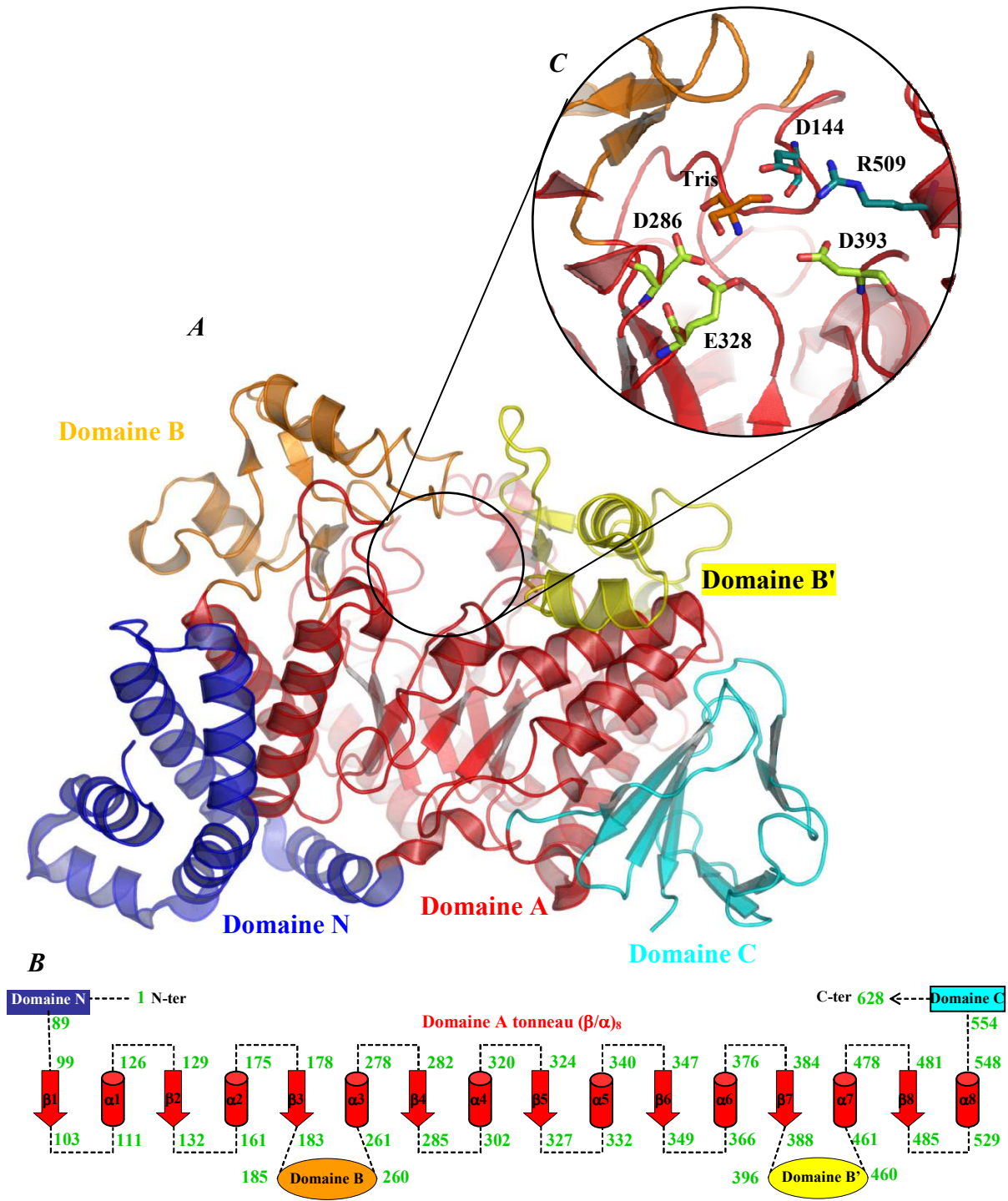


Figure 8 : Organisation tridimensionnelle de l'ASNp (code PDB : 1G5A)

- (A) Structure tertiaire de l'ASNp divisée en 5 domaines (bleu : domaine N, rouge : domaine A, orange : domaine B, jaune : domaine B' et cyan : domaine C). Image générée avec Pymol (84)
- (B) Schéma topologique de l'ASNp avec en vert les bornes des différents domaines et éléments de structures secondaires du tonneau catalytique
- (C) Site actif de l'ASNp, avec, en vert la triade catalytique, en bleu le pont salin et en orange une molécule de Tris. D'après Skov et coll. (39, 49). Image générée avec Pymol (84)

Au sein du domaine A se trouve le site actif de l'enzyme. Il est recouvert par les domaines B et B' lui conférant une topologie en forme de poche. On y retrouve la triade catalytique composée des résidus D286, E328 et D393 (Figure 8C) ainsi que le pont salin impliquant les résidus D144 et R509 qui délimite le fond de la poche catalytique. Au sein du site actif se trouve une molécule de Tris complexée, issue du tampon de conditionnement de la protéine utilisé lors de la cristallisation de l'ASNp.

II-4.2 Structure du complexe ASNp E328Q-saccharose

L'étude du complexe entre le mutant inactif de l'ASNp (E328Q) et le saccharose (code PDB : 1JGI (80)) met en évidence deux sites de fixation du saccharose sur la protéine (Figure 9A). Le premier site, nommé SB1 (pour Sucrose Binding site 1) se situe au niveau du site catalytique. Le résidu glucopyranosyle est présent dans une conformation chaise 4C_1 et se trouve en interactions avec deux des résidus de la triade catalytique (D286, D393), les résidus du pont salin (D144, R509) ainsi que d'autres résidus conservés au sein de la famille GH13 : H187, H392 et R284. Une interaction d'empilement est aussi observée entre le résidu glucopyranosyle et le résidu Y147. Le résidu fructofuranosyle, en conformation 4T_3 est en interaction avec les résidus D393, Q328, D394 et R446 (Figure 9B). Mirza et coll. (80) ont souligné l'importance de l'obtention de ce complexe dans la compréhension de la première étape du mécanisme réactionnel (cf paragraphe I-4). En s'appuyant sur cette structure, ils ont montré que la molécule de saccharose est parfaitement alignée pour un transfert optimal du proton du résidu acide E328, entraînant la formation de l'intermédiaire covalent β -D-glucosyl-enzyme et une libération du fructose. Mirza et coll. rapportent aussi la structure du complexe ASNp-glucose (code PDB : 1JG9 (80)). Les interactions observées sont similaires à celles du résidu glucopyranosyle du complexe ASNp-saccharose, cependant, le glucose au sein du site actif est majoritairement retrouvé sous forme d'anomère β en conformation chaise 4C_1 permettant à Mirza et coll. de conclure que la poche catalytique de l'ASNp serait optimisée pour la sélection de l'anomère β .

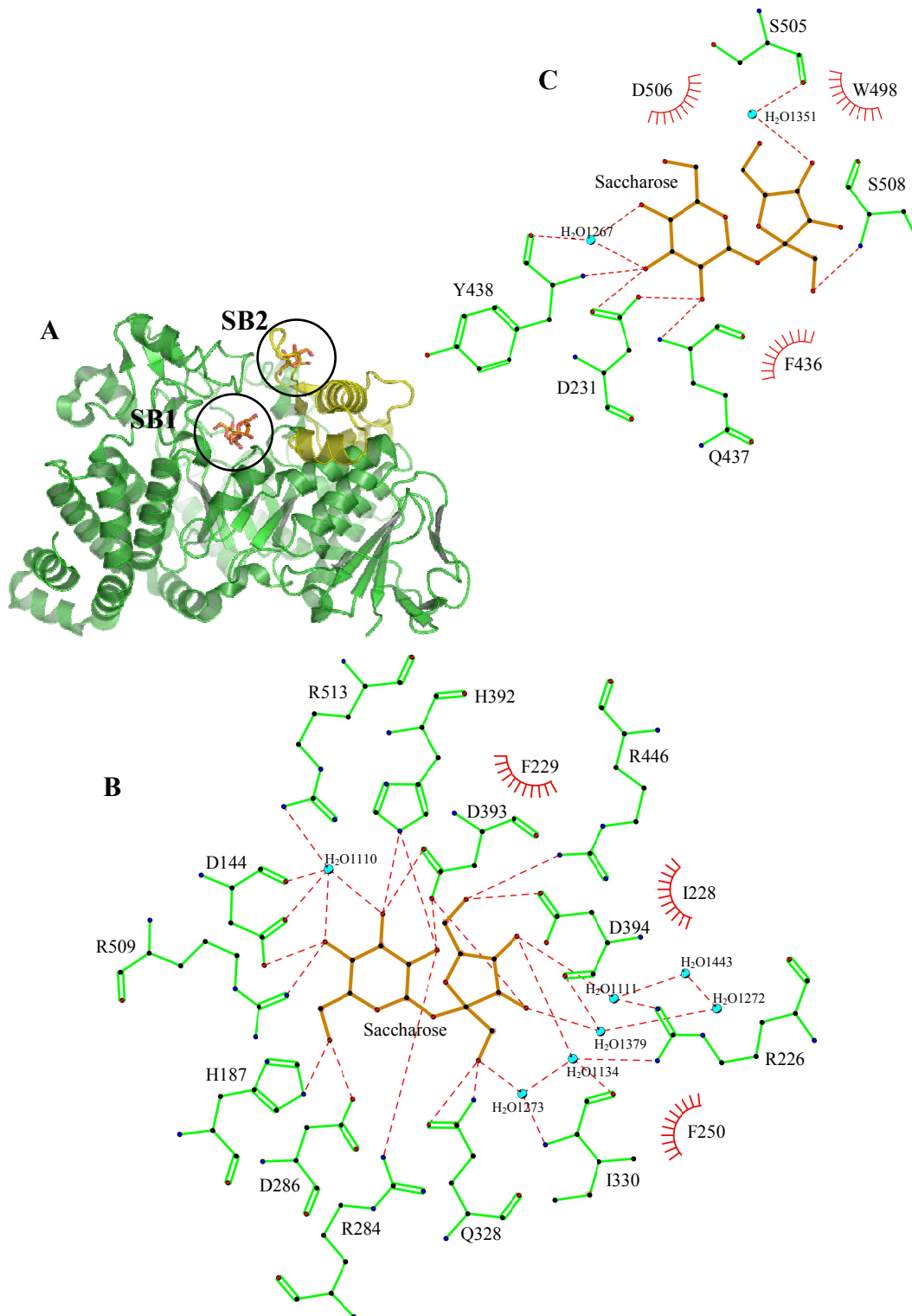


Figure 9 : Sites de fixation du saccharose dans l'ASNp

(A) Vue générale de l'ASNp (code PDB : 1JGI (80)), en orange, les sites de fixation du saccharose (SB1 et 2).
Image générée avec Pymol (84)

(B) Détail du site SB1 au sein du site catalytique. Image générée avec Ligplot (85)

(C) Détail du site de surface SB2 sur le domaine B', en jaune. Image générée avec Ligplot (85)

Skov et coll. (81) ont observé un second site de fixation du saccharose SB2 en surface du domaine B' (code PDB : 1MW1 (81)). Ce site nommé SB2 implique 5 résidus en surface de la protéine (Figure 9C) et se trouve très proche du site catalytique. Skov et coll. émettent l'hypothèse qu'une molécule de saccharose fixée à cet endroit pourrait pénétrer dans le site actif suite à un mouvement de large amplitude du domaine B' suggérant ainsi une régulation allostérique par le saccharose lui-même.

II-4.3 Structure de l'intermédiaire covalent β -D-glucosyl-enzyme

Cette structure de l'ASNp E328Q comporte le résidu D286 lié de manière covalente à un résidu β -D-glucosyle (code PDB : 1S46 (83)). Elle met en évidence les différentes interactions entre les résidus du site actif, le résidu β -D-glycosyle et une molécule d'eau (H₂O 800) qui serait impliquée dans la réaction d'hydrolyse (Figure 10). A partir de ce complexe Jensen et coll. (83) ont analysé l'origine de l'activité de transglucosylation en comparant cette structure à celle de l'intermédiaire covalent de la xylanase de *Streptomyces lividans* qui hydrolyse les liaisons β -(1,4) du xylane. De cette comparaison, ils ont déduit que le positionnement de la molécule d'eau H₂O 800, au sein du complexe covalent β -D-glucosyl-ASNp, ne se trouve pas en position optimale pour être un accepteur du résidu β -D-glucosyle. En effet, la distance entre la molécule d'eau et le carbone anomérique du résidu β -D-glucosyle est de 4,62 Å alors qu'elle devrait être beaucoup plus faible (3,7 Å, comme décrit pour la xylanase). Jensen et coll. ont alors suggéré que cette différence de positionnement de la molécule d'eau serait à l'origine de la prépondérance de la réaction de transglucosylation de l'ASNp au détriment de l'activité hydrolytique (86, 87).

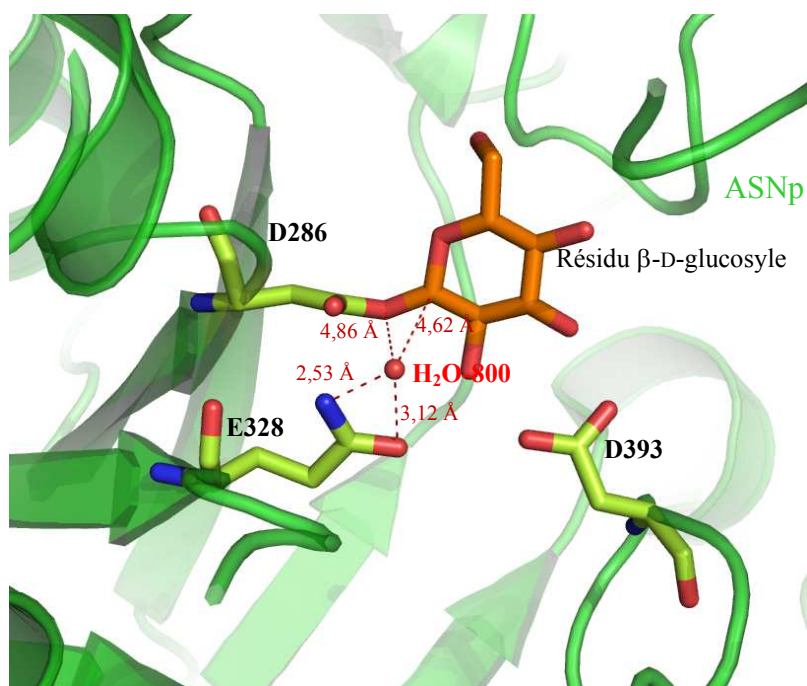


Figure 10 : Site actif de l'ASNp dont le résidu catalytique D286 est lié de façon covalente au résidu β-D-glucosyle (code PDB : 1S46)

En vert clair est représentée la triade catalytique. Le résidu β-D-glucosyle est coloré en orange et se trouve lié de manière covalente au résidu D286. La molécule d'eau H₂O 800, colorée en rouge, est représentée avec les distances la séparant des résidus catalytiques D286, E328 et du carbone anomérique du résidu β-D-glucosyle. D'après Jensen et coll. (83). Image générée avec Pymol (84)

II-4.4 Structure du complexe ASNp E328Q-maltoheptaose

Le complexe ASNp E328Q-maltoheptaose (code PDB : 1MW0 (81)) a permis de caractériser les différents sous-sites de fixation au niveau de la poche catalytique (site appelé OB1 pour Oligosaccharide Binding site 1, Figure 11). Le site actif forme une poche possédant 5 sous-sites allant de -1 à +4, les deux derniers sous-sites, +5 et +6, sont localisés en surface de la protéine. Les sous-sites sont nommés selon la nomenclature définie par Davies et coll. (47). La rupture de la liaison osidique s'opère entre les sous-sites -1 et +1, les sous-sites négatifs fixant l'extrémité non réductrice d'un maltooligosaccharide tandis que les sous-sites positifs fixent l'extrémité réductrice.

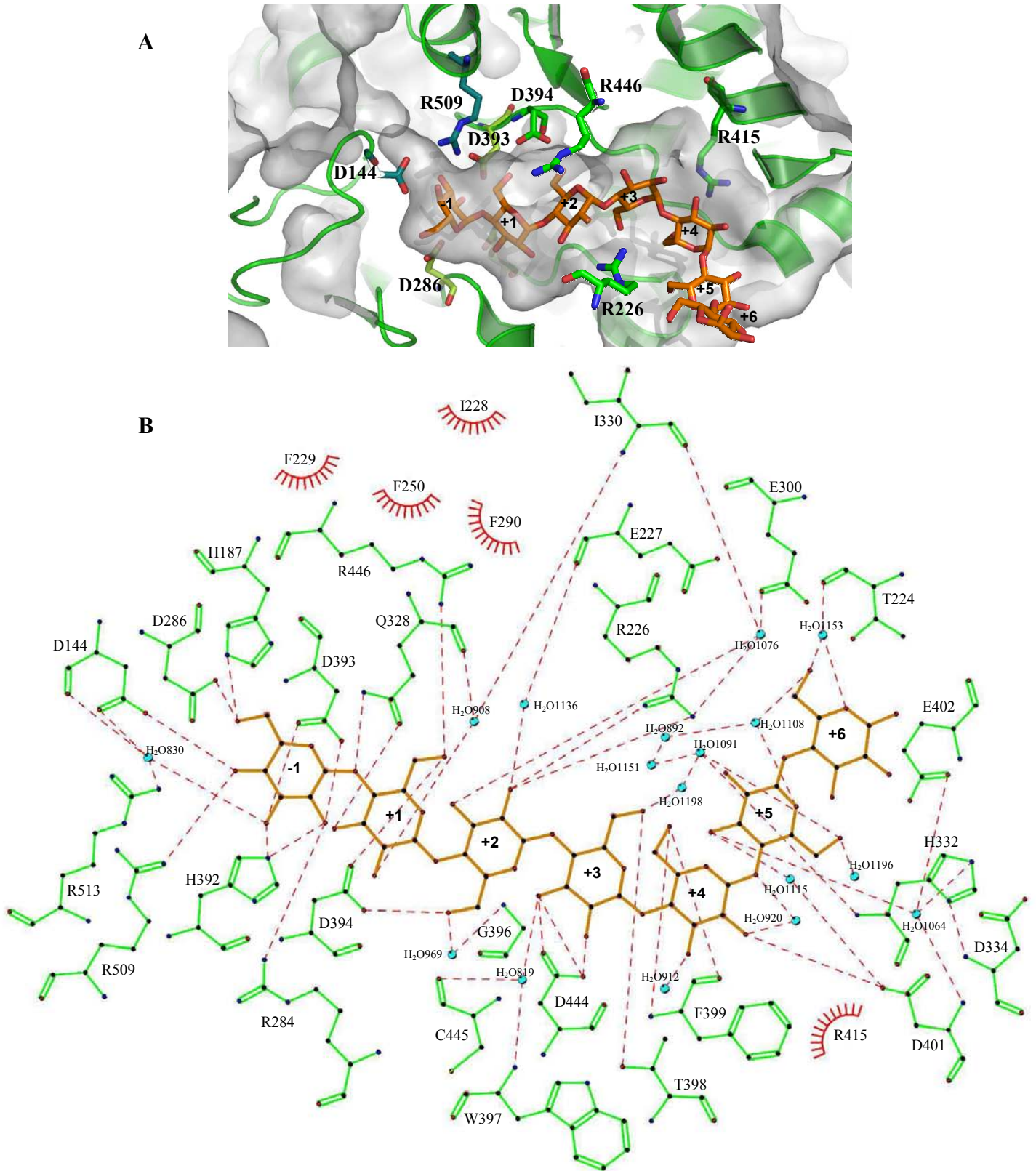


Figure 11 : Représentation schématique de la poche catalytique

(A) Vue en coupe du site actif de l'ASNp en complexe avec le maltoheptaose (code PDB : 1MW0) permettant de visualiser les 7 sous-sites de fixation des sucres dans l'enzyme. D'après Skov et coll. (81). Image générée avec Pymol (84)

(B) Détail du site de fixation du maltoheptaose (OB1). Image générée avec Ligplot (85)

Deux autres sites de liaison aux oligosaccharides sont aussi décrits en surface de l'ASNp (Figure 12A) : OB2 au niveau du domaine B' (Figure 12B) et OB3 (Figure 12C) au niveau du domaine C (81). Ces sites sont susceptibles de faciliter l'arrimage de longues chaînes oligosaccharidiques acceptrices permettant la stabilisation de complexes entre l'ASNp et une longue chaîne acceptrice lors des réactions de transglucosylation.

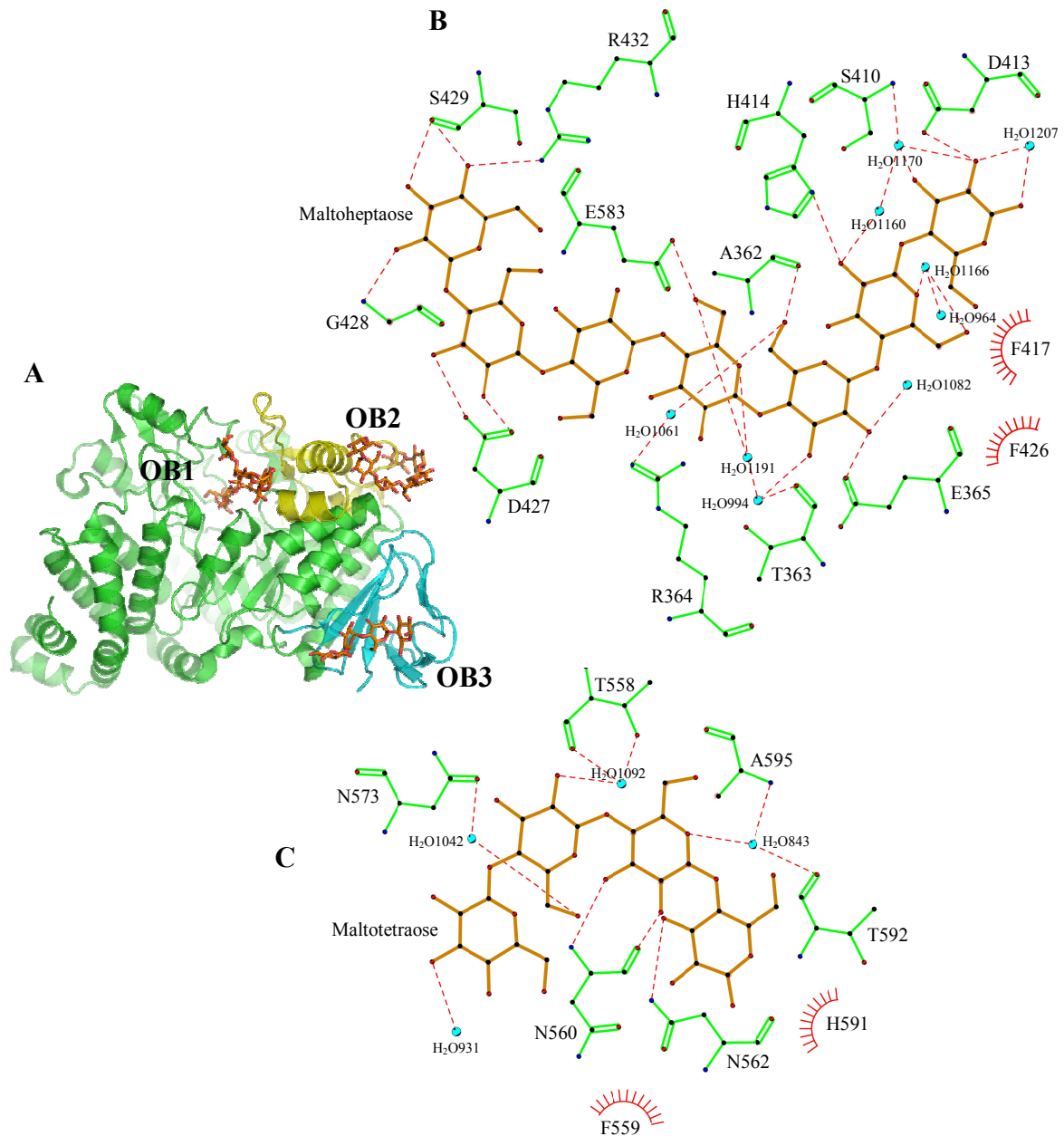


Figure 12 : Sites de fixation des oligosaccharides dans l'ASNp

(A) Vue générale de l'ASNp, en orange, les sites de fixation d'oligosaccharide (OB) d'après Skov et coll. (81).

Image générée avec Pymol (84)

(B) Détail du site de surface OB2 sur le domaine B', en jaune. Image générée avec Ligplot (85)

(C) Détail du site de surface OB3 sur le domaine C, en cyan. Image générée avec Ligplot (85)

II-5 Relations structure-activité

Afin de mieux comprendre les relations structure-activité de l'ASNp, Albenne et coll. (88) ont entrepris de sonder par mutagenèse dirigée, et plus particulièrement par alanine scanning (89), le rôle de certains résidus clefs du site actif sur l'activité enzymatique.

Des résidus du site OB1 (D394, R446, R226 et R415, Figure 11A) ont ainsi été individuellement substitués par une alanine. L'activité de ces mutants sur le saccharose a alors été étudiée en détail, et la composition des produits obtenus à partir de 100 mM de saccharose est donnée dans le Tableau 6.

Tableau 6 : Activité relative et profil des produits obtenus à partir des différents mutants de l'ASNp

d'après Albenne et coll. (88)

Enzyme Mutation	Sous-site affecté par la mutation, domaine concerné	Activité relative (%)	Saccharose consommé (%)
ASNp Sauvage		100 (1 000 U/g)	100
D394A	+1, B'	23,5	60
R446A	+1, B'	15	16
R415A	+4, B'	4,3	20
R226A	+2/+3, B	30	100

Enzyme Mutation	Concentration et rendement de production (pour 100 mM de saccharose initial)						
	Glucose	Turanose	Tréhalulose	Maltose	Maltotriose	Glucane soluble	Glucane insoluble
ASNp Sauvage	5 mM 5 %	14 mM 14 %	3 mM 3 %	2,3 mM 5 %	5,3 mM 16 %	<0,05 mM 12 %	nd 45 %
D394A	38 mM 63 %	5 mM 8,5 %	1 mM 1,5 %	4,2 mM 14 %	1,4 mM 7 %	<0,02 mM 6 %	0 %
R446A	11 mM 69 %	0,6 mM 4 %	0,2 mM 1 %	1,2 mM 15 %	0,5 mM 10 %	<0,01 mM 1 %	0 %
R415A	5 mM 25 %	1 mM 5 %	0,2 mM 1 %	1,8 mM 18 %	3,2 mM 48 %	<0,01 mM 3 %	0 %
R226A	3 mM 3 %	2,5 mM 2,5 %	0,5 mM 0,5 %	0,5 mM 1 %	1,5 mM 4,5 %	<0,01 mM 3,5 %	nd 85 %

nd : Non déterminé

La substitution par une alanine des résidus D394 et R446, deux résidus impliqués dans des interactions avec le substrat au niveau du sous-site +1 (Figure 11A), induit une augmentation drastique du taux de glucose final indiquant ainsi une activité "superhydrolase" de ces mutants. La mutation en alanine du résidu R226 impliqué dans des interactions au niveau du sous-site +2 (Figure 11A) conduit, à l'inverse, à une activité "superpolymérase" conduisant principalement à la formation de maltooligosaccharides insolubles. Quant à la mutation R415A, située à l'entrée de la poche catalytique, au niveau du sous-site +4 (Figure 11A), elle conduit à une augmentation significative du taux de glucose et de maltotriose final. Les résidus D394 et R446 appartiennent au domaine B', un domaine retrouvé exclusivement chez les amylosaccharases et sont clairement impliqués dans la réaction de transglucosylation qui se fait au détriment de la réaction d'hydrolyse du saccharose (observable en cas de mutation de ces résidus en alanine). Les résidus R415 et R226 semblent plutôt impliqués dans l'arrimage des chaînes polysaccharidiques, le résidu R226, un résidu encombrant, gênerait par sa taille la fixation du maltotriose ce qui freinerait la polymérisation. Lors de la réaction sur saccharose seul, la mutation R226A entraîne, une activation de la polymérisation au détriment des autres types de réactions (hydrolyse ou isomérisation). En outre, ce mutant montre une forte activation par le glycogène et catalyse essentiellement l'élongation des chaînes d' α -glucanes du polymère lorsqu'il est en présence de saccharose supplémenté de glycogène. De plus, les résultats rapportés par Albenne et coll. sur l'activité de l'enzyme sauvage en présence de saccharose seul, montrent que la polymérisation par l'ASNp se produit suivant un mécanisme de polymérisation non processif (ou multi-chaînes) (90) et que l'ASNp n'a pas besoin d'amorce pour commencer la réaction de polymérisation.

II-6 Réaction de disproportionnation

L'ASNp peut, à partir d'un maltooligosaccharide comme seul substrat, catalyser une réaction appelée réaction de disproportionnation (91). Cette réaction conduit à la formation de différents maltooligosaccharides à partir d'un maltoligosaccharide de taille donnée qui joue à la fois le rôle de donneur et d'accepteur de résidu glucosyle. Ainsi, à partir de maltohexaose, l'enzyme synthétise dans un premier temps du maltopentaose et du maltoheptaose de façon équimolaire. Dans un second temps, on voit apparaître, toujours de manière équimolaire, du maltotétraose et du maltooctaose. La réaction atteint un équilibre après 4 heures. Cependant, l'efficacité de cette réaction de disproportionnation dépend de la taille du

maltooligosaccharide, en effet, en dessous de 5 résidus glucosyle (maltopentaose), la vitesse de réaction est drastiquement diminuée.

Conclusion

Depuis le clonage de l'amylosaccharase de *Neisseria polysaccharea* en 1999, facilitant sa production et sa purification, de nombreuses études ont été menées, à commencer par la caractérisation des produits formés par ce catalyseur au cours de la réaction enzymatique en présence de saccharose. Par la suite, les études structurales sur l'ASNp ont permis de caractériser la structure sous forme *apo* et en complexe avec son substrat (saccharose) ou ses produits (glucose, maltoheptaose) ainsi qu'un intermédiaire covalent permettant d'aller plus loin dans la compréhension de son mécanisme réactionnel.

Depuis, de nouvelles amylosaccharases ont été caractérisées, montrant chacune des propriétés biochimiques et enzymatiques originales telles que la thermostabilité de l'ASDg ou des profils originaux de produits. L'origine de ces différences demeure inexplicée à ce jour.

Les amylosaccharases ont un véritable potentiel comme outils de glucosylation. Elles ont fait l'objet de nombreux brevets dont l'intérêt principal est l'utilisation du saccharose, une agroressource abondante et peu coûteuse, comme donneur d'unités glucosyle. Cependant, certaines limitations freinent leur exploitation industrielle, comme par exemple leur faible stabilité. Pour remédier à ces faiblesses, des travaux d'ingénierie ont été mis en œuvre sur les amylosaccharases, en particulier l'ASNp, afin d'améliorer leur stabilité, leur efficacité catalytique ou changer leur spécificité de substrat. Ces travaux sont exposés dans la troisième partie de ce chapitre.

Partie III : Ingénierie des Amylosaccharases

Depuis quelques années, des travaux d'ingénierie moléculaire ont été menés sur les amylosaccharases, essentiellement l'ASNp, afin, d'une part, d'améliorer les propriétés biochimiques et catalytiques, mais également d'utiliser leur potentiel de glycosylation de molécules exogènes dans le but de développer de nouvelles applications. Plusieurs approches ont été utilisées allant de l'ingénierie combinatoire pour l'amélioration de la stabilité et des performances catalytiques des enzymes à de l'ingénierie semi-rationnelle afin de modifier la spécificité vis-à-vis de l'accepteur.

III-1 Amélioration de l'activité et de la stabilité de l'ASNp

Les techniques d'évolution moléculaire dirigée permettent de créer de larges banques de mutants par mutations aléatoires et/ou recombinaison. Ces banques sont ensuite soumises à une sélection phénotypique ou un criblage permettant d'isoler les variants possédant la propriété recherchée. Ces approches qui miment en laboratoire et en accélèrent les phénomènes de l'évolution naturelle des protéines sont largement utilisées depuis les années 90 pour améliorer ou modifier les propriétés des enzymes (92, 93).

Nous décrivons ici les applications de ces méthodes pour l'ingénierie de l'ASNp.

Van der Veen et coll. ont généré une librairie de 100 000 mutants de l'ASNp par PCR à erreur et recombinaison aléatoire afin d'améliorer les propriétés catalytiques de l'enzyme pour les adapter à des conditions de synthèses industrielles (94). Cette librairie a été ensuite criblée par étalement des clones sur un milieu solide contenant du saccharose. Seuls les clones positifs produisent de l'amylose, facilement repérable visuellement. 7 clones ont ainsi été isolés et caractérisés. Deux mutants V389L-N503I et R20C-F598S ont révélé une amélioration d'un facteur ~4 de l'activité spécifique allant de 1,7 U.mg⁻¹ pour l'enzyme sauvage à 6,4 U.mg⁻¹ pour le double-mutant V389L-N503I et 7,2 U.mg⁻¹ pour le double-mutant R20C-F598S.

Afin d'améliorer la thermostabilité de l'ASNp, Emond et coll. ont mis en œuvre une stratégie d'évolution dirigée en utilisant la technique MutaGen (95) fondée sur l'utilisation des polymérase β et η humaines qui présentent la particularité d'être très peu fidèles. Une librairie de 30 000 clones a été créée (96). L'isolement des clones exprimant les amylosaccharases les plus prometteuses a été fait en deux étapes comprenant :

- ✓ la sélection des variants exprimant une amylosaccharase active sur saccharose à 30 °C sur milieu solide (4 500 clones sélectionnés),

- ✓ le criblage en milieu liquide des variants exprimant une amylosaccharase active sur saccharose à 37 °C après un traitement de 20 min à 50 °C (3 clones sélectionnés).

Cette technique a permis d'isoler des clones dont le temps de demi-vie à 50 °C a été amélioré d'un facteur 10, passant de 3 minutes à 32 minutes, notamment avec le double mutant R20C-A451T. Par modélisation moléculaire, il a été suggéré que le gain de stabilité induit par la mutation R20C serait dû à un réarrangement des ponts salins de la région environnant l'arginine 20 pouvant affecter la dynamique des mouvements de l'ASNp et/ou la flexibilité de la région mutée (96). En ce qui concerne la seconde mutation A451T, Emond et coll. ont émis l'hypothèse que l'introduction d'une fonction hydroxyle permettrait la création de liaisons hydrogène entre la thréonine 451 et l'aspartate 468, un résidu situé dans le domaine catalytique, ce qui pourrait ainsi stabiliser le réseau d'interactions de ce mutant thermostable.

III-2 Modification de la spécificité de l'ASNp

L'ingénierie semi-rationnelle combine les approches rationnelle et combinatoire et vise à construire des bibliothèques de variants sur des positions identifiées à partir des données structurales, variants qui sont ensuite triés par des techniques de criblage à moyen ou haut débit (97). Les bibliothèques obtenues sont, en effet, généralement de taille plus réduites que celles générées par les techniques aléatoires. Cette approche a été employée par Champion et coll. dans le cadre du projet ANR OptiGluc (2005-2008) qui avait pour objectif de concevoir de nouvelles voies de synthèse chimio-enzymatique permettant d'accéder à une gamme d'oligosaccharides antigéniques en vue du développement de vaccins contre la shigellose, une maladie diarrhéique meurtrière dans les pays en voie de développement, affectant principalement des enfants en bas âge (98). Ainsi, ce projet a consisté à dessiner de nouvelles voies de synthèse d'oligosaccharides mimant les sérotypes les plus prévalents de l'antigène O de *Shigella flexneri*. La majorité de ces sérotypes sont composés d'un tétrasaccharide linéaire composé d'unités L-rhamnose et N-acétyl-D-glucosamine, sur lequel sont greffées, par différentes liaisons osidiques, des unités D-glucosyle selon le sérotype. Le contrôle de la glucosylation 1,2-*cis* est une étape très délicate à réaliser par voie chimique. Pour lever ce verrou, le choix d'une glucosylation par voie enzymatique s'intégrant à une voie de synthèse chimio-enzymatique a été proposé.

Pour cela, deux molécules préparées pour leur élongation par voie chimique ont été ciblées : l' α -allyl-N-acétyl-D-glucosamine et l' α -méthyl-L-rhamnopyranoside, la première devant être

glucosylée en α -(1,4) et la seconde en α -(1,3). Aucune enzyme naturelle n'est capable de catalyser ces réactions avec la spécificité requise. Dès lors, l'option de l'ingénierie moléculaire a été choisie afin de créer de nouvelles enzymes capables de catalyser la glucosylation de ces accepteurs non-naturels en partant de l'ASNp comme enzyme candidate à l'évolution.

III-2.1 Stratégie de construction des bibliothèques de variants

Champion et coll. (3) ont réussi un tour de force en créant des mutants de l'ASNp dotés d'une nouvelle spécificité vis-à-vis de nouveaux accepteurs, naturellement non reconnus par l'enzyme sauvage. Guidée par une étude de modélisation moléculaire, une stratégie d'évolution rationnelle a été mise en œuvre afin de créer une banque de taille réduite de mutants de l'ASNp. Au total sept cibles de mutagenèse (I228, A289, F290, I330, V331, E394 and R446) ont été identifiées par arrimage moléculaire des produits d'intérêt dans le site actif (Figure 13) :

- ✓ I228, I330, E394 et R446 pour la synthèse de l' α -D-glucopyranosyl-1,3-méthyl-L-rhamnopyranoside (nouvelle spécificité),
- ✓ A289, F290, I330 et V331 pour la synthèse de l' α -D-glucopyranosyl-1,4- α -allyl-N-acétyl-D-glucosamine (amélioration de l'activité).

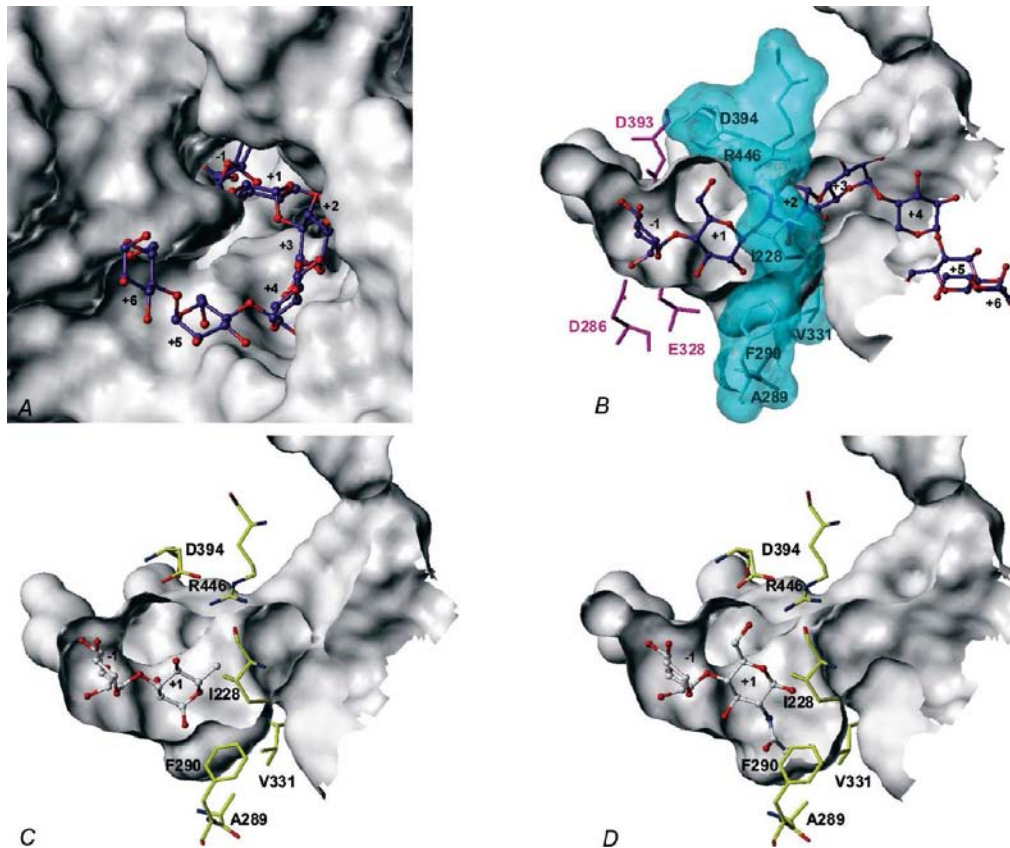


Figure 13 : Arrimage moléculaire des produits cibles dans le site actif de l'ASNp

d'après Champion et coll. (3)

- (A) Vue de la poche catalytique de l'ASNp en complexe avec une molécule de maltoheptaose
(B) Vue en coupe du site actif avec les résidus à l'interface des sous-sites -1 et +1. Les résidus ciblés pour la mutagenèse sont représentés en surface bleu
(C) Arrimage moléculaire de l' α -D-glucopyranosyl-1,3-méthyl-L-rhamnopyranoside
(D) Arrimage moléculaire de l' α -D-glucopyranosyl-1,4- α -allyl-N-acétyl-D-glucosamine

Afin d'explorer la diversité sur ces différentes positions, une banque de 133 monomutants de l'ASNp a été construite en remplaçant, par mutagenèse dirigée, chacune des positions par les 19 autres acides aminés possibles.

La banque de monomutants a ensuite été criblée sur leur capacité à glucosyler les deux accepteurs d'intérêt : l' α -méthyl-L-rhamnopyranoside et le N-acétyl-D-glucosamine.

L'ASNp sauvage n'est pas capable de glucosyler l' α -méthyl-L-rhamnopyranoside mais 15 mutants ont acquis une nouvelle spécificité de glucosylation et se sont avérés capables de synthétiser de l' α -D-glucopyranosyl-1,3-méthyl-L-rhamnopyranoside (Figure 14A). De plus, la position 228 semble être critique puisque 11 des 15 mutants sont des substitutions de l'isoleucine 228.

En ce qui concerne la synthèse de l' α -D-glucopyranosyl-1,4-*N*-acétyl-D-glucosamine, l'ASNp sauvage est capable de reconnaître l' α -allyl-*N*-acétyl-D-glucosamine, cependant le taux de glycosylation de l'accepteur est faible (2 %). Le criblage de la banque a révélé 30 mutants capables de glycosyler le *N*-acétyl-D-glucosamine avec un taux supérieur à 5 % (Figure 14B). Les positions 290, 289 et 228 semblent être impliquées dans la reconnaissance de l'accepteur, notamment les mutants de la phénylalanine 290 qui montrent les améliorations d'activité les plus spectaculaires, avec des taux de glycosylation proches de 100 % pour les mutants F290D et F290K.

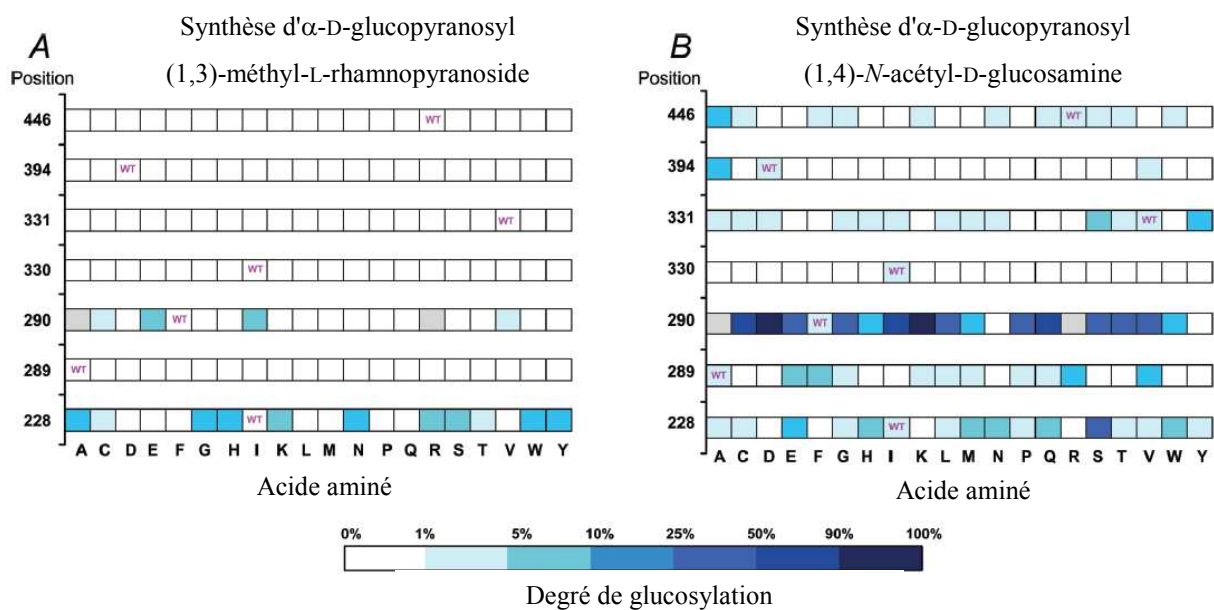


Figure 14 : Bilan de la glycosylation des accepteurs cibles par les monomutants non purifiés

d'après Champion et coll. (3)

- (A) Glycosylation de l' α -méthyl-L-rhamnopyranoside
(B) Glycosylation de la *N*-acétyl-D-glucosamine

III-2.2 Caractérisation des propriétés biochimiques des mutants de spécificité modifiée

Deux monomutants de l'ASNp, de spécificité améliorée (mutant F290K) ou nouvelle (mutant I228Y) vis-à-vis des accepteurs cibles ont ainsi été identifiés.

Mutant ASNp I228Y : l'analyse des produits de la réaction sur saccharose seul révèle que la réaction de transglucosylation est abolie au profit de l'hydrolyse du saccharose. Avec addition de l' α -méthyl-L-rhamnopyranoside, 41 % du saccharose initial est employé pour la synthèse de l' α -D-glucopyranosyl-1,3-méthyl-L-rhamnopyranoside, 3 % est employé pour la synthèse de formes di-glucosylées) et 53 % du saccharose initial est hydrolysé. L'efficacité catalytique de la réaction de transglucosylation sur l'accepteur est de $0,085 \text{ mM}^{-1} \cdot \text{min}^{-1}$.

Mutant ASNp F290K : contrairement au mutant I228Y, le mutant F290K présente une activité de polymérisation du glucose sur saccharose seul. En additionnant de l' α -allyl-N-acétyl-D-glucosamine en condition initiale, la réaction de polymérisation du glucose est réduite au profit de la réaction de transglucosylation sur l'accepteur avec une production de 56 % d' α -D-glucopyranosyl-1,4- α -allyl-N-acétyl-D-glucosamine et de 22 % de formes di-glucosylées. L'efficacité catalytique de la réaction de transglucosylation sur l'accepteur est de $15,9 \text{ mM}^{-1} \cdot \text{min}^{-1}$ contre $0,12 \text{ mM}^{-1} \cdot \text{min}^{-1}$ pour l'enzyme sauvage, soit une amélioration d'un facteur 132.

Au début de notre étude, les facteurs moléculaires à l'origine des performances des mutants isolés étaient encore méconnus.

Conclusion

Les outils de biologie moléculaire offrent aujourd'hui des possibilités inédites pour mener à bien l'ingénierie d'enzymes aux propriétés améliorées ou nouvelles, allant de la stabilité et de l'activité enzymatique à la spécificité de substrat. Néanmoins, des techniques de criblage toujours plus performantes sont nécessaires afin d'identifier les enzymes ayant les propriétés recherchées.

Malgré la puissance de ces outils de biologie moléculaire pour l'évolution dirigée, les origines moléculaires des nouvelles propriétés catalytiques acquises demeurent difficilement explicables sans une analyse approfondie des mécanismes et déterminants moléculaires régissant leur comportement cinétique, leur stabilité ou leur affinité pour un substrat donné. Les méthodes biophysiques permettant de rendre compte de ces phénomènes seront exposées en quatrième partie de ce chapitre.

*Partie IV : Structure, Stabilité des Protéines et
Interactions Protéine-Ligand*

L'analyse des structures secondaire et tertiaire des protéines et la détermination des facteurs moléculaires responsables de leur stabilité ou intervenant dans les interactions protéine-ligand s'appuient sur des techniques biophysiques et de modélisation moléculaire variées. Dans le cadre de nos travaux, nous avons mis en œuvre un certain nombre de ces techniques parmi lesquelles le dichroïsme circulaire, la fluorimétrie différentielle à balayage, la cristallographie des rayons X, la calorimétrie de titrage isotherme et des approches de modélisation moléculaire.

Ce chapitre a pour but de présenter les principes de ces différentes méthodes et, au travers de quelques exemples, de faire le point sur les informations qu'elles permettent d'obtenir.

IV.1 Le dichroïsme circulaire

Le dichroïsme circulaire repose sur l'absorption non équivalente des composantes gauche et droite de la lumière polarisée plane (99). La lumière polarisée passe à travers un modulateur pouvant la séparer en une composante gauche (L) et une composante droite (R) de magnitude égale mais l'une tournant dans le sens horaire (R) et l'autre dans le sens anti-horaire (L) (Figure 15).

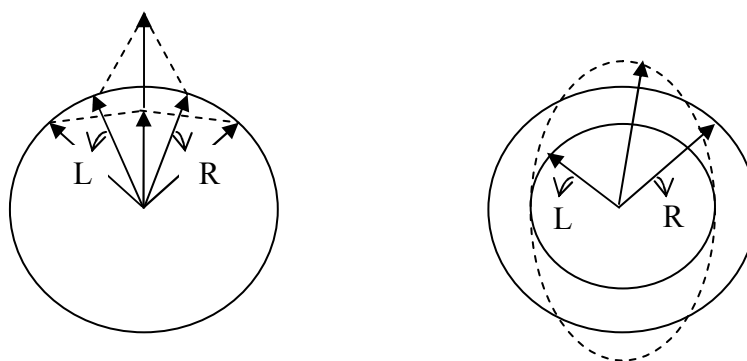


Figure 15 : Lumière polarisée

Ces deux composantes s'additionnent en traversant un environnement non chiral pour redonner une lumière polarisée plane circulaire. En revanche, si elles traversent un environnement chiral, une des deux composantes peut être absorbée préférentiellement. L'addition des deux composantes nous donne donc un signal non plus circulaire mais ellipsoïdal. L'expression du dichroïsme est appelée ellipticité (Equation 1). L'ellipticité étant

dépendante de la concentration en molécule chirale, elle est normalisée et appelée ellipticité molaire (Equation 2).

Equation 1 : Ellipticité

$$\theta = \frac{2,303*(A_l-A_r)*180}{4\pi} = 32,98*\Delta A$$

Avec :

- θ : ellipticité (deg)
- A_l : absorbance de la composante gauche
- A_r : absorbance de la composante droite
- ΔA : A_l-A_r

Equation 2 : Ellipticité molaire

$$[\theta] = \frac{100*\theta}{l*C} = 3298*\Delta\varepsilon$$

Avec :

- $[\theta]$: ellipticité molaire (deg.cm².dmol⁻¹)
- l : trajet optique
- C : concentration molaire
- $\Delta\varepsilon$: $\varepsilon_l-\varepsilon_r$
- ε_l : coefficient d'extinction molaire de la composante gauche
- ε_r : coefficient d'extinction molaire de la composante droite

Appliqué à l'étude des protéines, le dichroïsme circulaire permet d'étudier leur structure secondaire en utilisant une gamme de longueur d'onde qui se trouve dans l'UV lointain, entre 160 et 260 nm. A ces longueurs d'onde, l'absorption est essentiellement due aux liaisons peptidiques. Toutefois, les résidus aromatiques et les ponts disulfures peuvent absorber la lumière à 260 nm. L'ellipticité molaire observée est dépendante de la structure secondaire adoptée par le squelette peptidique. En effet, l'ellipticité molaire d'une solution contenant une protéine constituée exclusivement d'hélices α présentera un pic positif à 195 nm et deux pics négatifs à 208 et 222 nm. Par contre, dans le cas d'une protéine exclusivement constituée de brin β l'ellipticité molaire présentera un pic positif à 198 nm et un pic négatif à 218 nm (Figure 16). L'amplitude et la forme du spectre enregistré dans l'UV lointain vont dépendre de la proportion des différents éléments de structure secondaire présents dans la protéine. On pourra, par une analyse informatique, en déduire un pourcentage de chaque motif secondaire présent dans la protéine. Le principal inconvénient du dichroïsme circulaire est le choix du tampon de conditionnement de la protéine qui doit être optiquement inerte aux longueurs

d'ondes utilisées. La présence de sels ou additifs est à proscrire également (ou en concentration minimale). La technique nécessite peu d'échantillon biologique (~0,1 mg de protéine) et les temps de manipulation sont incompatibles avec des analyses à haut ou moyen débit.

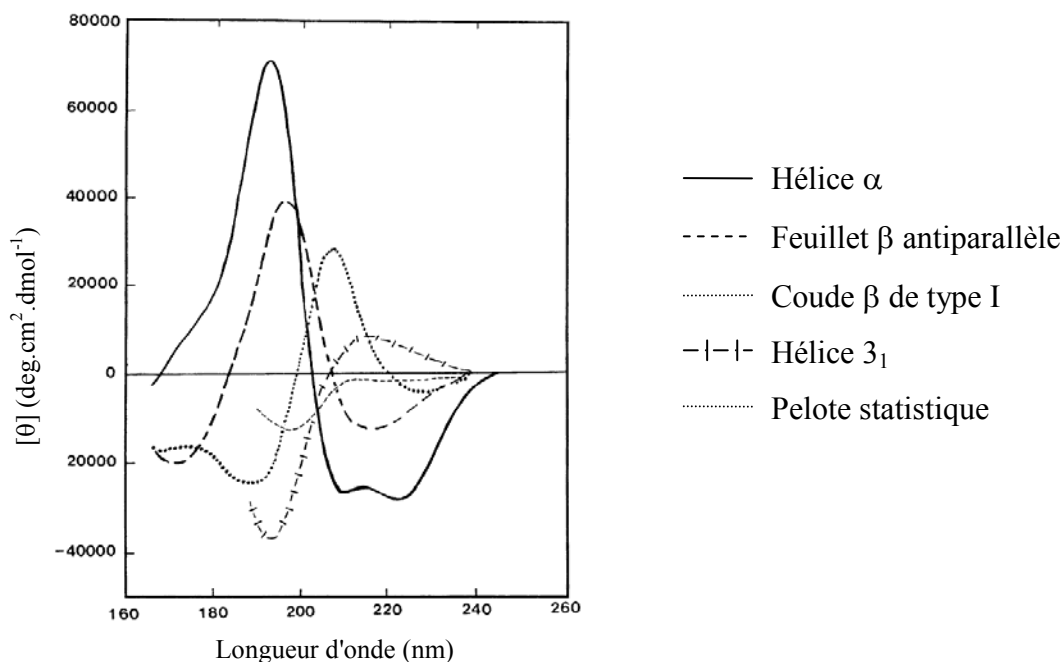


Figure 16 : Spectre d'absorption dans l'UV lointain en fonction de différents types de structure secondaire

D'après Kelly et coll. (99)

Le suivi de l'évolution des éléments de structure secondaire d'une protéine au cours de divers traitements est employé pour déterminer la température de dénaturation des protéines. Gatti-Lafranconi et coll. (100) ont suivi la dénaturation thermique de la lipase de *Pseudomonas fragi* par dichroïsme circulaire tandis que Chi et coll. ont suivi la dénaturation d'une α -amylase thermostable de *Bacillus* induite par des ions et divers solvant (101).

Le dichroïsme circulaire permet également d'étudier la structure tertiaire des protéines. Toutefois, ces études nécessitent des quantités plus importantes d'échantillon biologique, (~1 mg de protéine). Pour cela, la gamme de longueur d'onde employée se situe dans le proche UV, entre 260 et 320 nm. A ces longueurs d'onde, l'absorption est majoritairement due aux résidus aromatiques. La phénylalanine présente des pics d'absorption à 255 et 270 nm, la

tyrosine à 275 et 282 nm et le tryptophane à 295 nm (Figure 17). L'amplitude et la forme du spectre enregistré dans le proche UV va dépendre de nombreux facteurs : la nature et le nombre de résidus aromatiques, leur mobilité et environnement chimique (résidus voisins, interaction faible, ...). Le spectre est alors qualifié "d'empreinte" de la structure tertiaire de la protéine. La comparaison avec une autre "empreinte", telle que celle d'un mutant, permet de mettre en évidence des modifications dans la structure tertiaire.

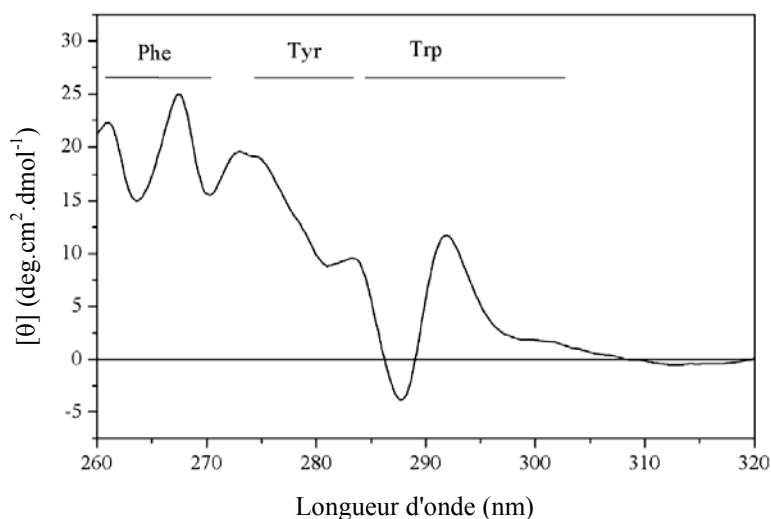


Figure 17 : Spectre d'absorption dans le proche UV

D'après Kelly et coll. (99)

Shizawa et coll. (102) ont montré l'effet de mutations sur une pyrophosphatase inorganique de *Bacillus subtilis* en étudiant des spectres de dichroïsme circulaire dans le proche UV (250-300 nm). Le mutant K297R montre un spectre complètement différent de l'enzyme sauvage indiquant un environnement différent des résidus aromatiques. Un contrôle dans le proche UV ne montre pas de différences dans les spectres, la structure secondaire n'est donc pas affectée par cette substitution, seule la structure tertiaire est influencée par la substitution K297R.

IV.2 La fluorimétrie différentielle à balayage

La fluorimétrie différentielle à balayage, en anglais DSF pour Differential Scanning Fluorimetry, est une technique permettant de suivre en temps réel le dépliement de la structure tertiaire d'une protéine (103). Cette méthode repose sur l'utilisation d'une sonde fluorescente, le Sypro Orange® (Invitrogen), dont la longueur d'onde d'absorption maximale est à 490 nm

et la longueur d'onde d'émission maximale est à 530 nm. De plus, il possède la particularité d'avoir un rendement quantique de fluorescence (Equation 3) faible en milieu hydrophile et élevé en milieu hydrophobe.

Equation 3 : Rendement quantique de fluorescence

$$\text{Rendement} = \frac{\text{Nombre de photons émis}}{\text{Nombre de photons reçus}}$$

Une solution de protéine (2 à 10 μM) est mise en présence du Sypro-orange (1 à 5 X) et est soumise à une augmentation de température graduelle qui dénature progressivement la protéine. La variation de fluorescence est enregistrée en fonction de la température. Généralement, une courbe d'allure sigmoïdale $\text{RFU}=\text{f}(\text{T})$ est obtenue. Elle peut être divisée en trois parties distinctes (Figure 18) :

- ✓ partie 1 : zone de stabilité de la protéine
- ✓ partie 2 : zone de dénaturation thermique de la protéine
- ✓ partie 3 : zone d'agrégation de la protéine

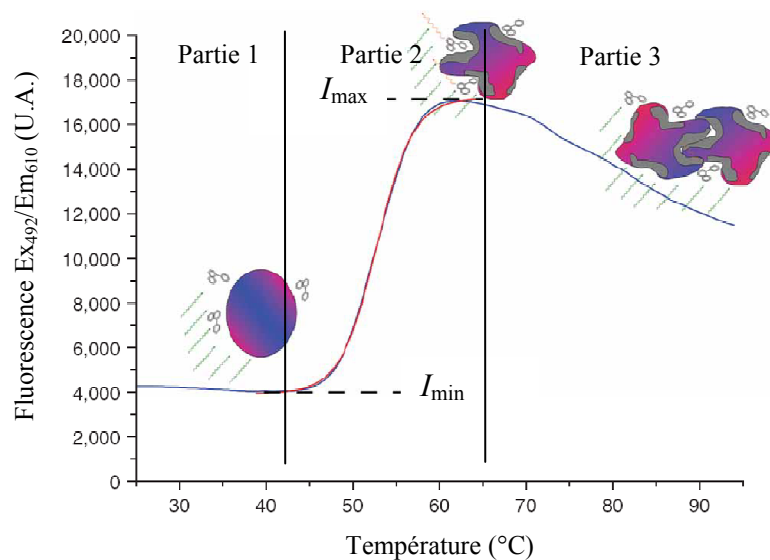


Figure 18 : Courbe de dénaturation

D'après Niesen et coll. (103)

Au cours de la dénaturation, le cœur hydrophobe des protéines est exposé au solvant aqueux, Le Sypro-orange entre alors en contact avec ces régions entraînant une augmentation de l'intensité de la fluorescence émise par cette sonde jusqu'à atteindre une intensité maximale où

la protéine est dénaturée. Les zones hydrophobes interagissent alors entre elles conduisant à un phénomène d'agrégation qui a pour conséquence de bloquer l'accès du Sypro-Orange à ces zones. L'intensité de fluorescence va alors décroître.

La température de fusion (T_m), caractéristique d'une protéine dans un milieu donné, peut être calculée à l'aide de l'équation de Boltzmann (Equation 4). Ce T_m est défini comme étant la température au point d'inflexion de la transition observée sur la courbe de dénaturation (partie 2 figure 18). Cette valeur caractéristique correspond à la température de demi-dénaturation.

Equation 4 : Equation de Boltzmann

$$y = I_{\min} + \frac{(I_{\max} - I_{\min})}{1 + \exp\left(\frac{T_m - x}{a}\right)}$$

Avec :

- I_{\min} : intensité minimale
- I_{\max} : intensité maximale
- T_m : température de dénaturation
- a : pente de la droite autour du T_m

En pratique, la valeur du T_m est obtenue en calculant la dérivée de la courbe de fluorescence $-d(\text{RFU})/(dT)=f(T)$. La valeur de T_m correspond à la température au minimum de cette fonction dérivée.

La DSF permet de rapidement déceler des différences de T_m de l'ordre de 0,5 °C, valeur dépendant principalement de la vitesse de montée en température. Elle est très employée pour l'étude de l'effet du conditionnement, de l'ajout de ligands, ou de mutations ponctuelles ou multiples sur la stabilité de la protéine.

En reportant sur un histogramme la différence entre le T_m d'une condition test et le T_m d'une condition de référence (ainsi que leurs écart-types), on peut évaluer le gain ou la perte de stabilité thermique de la protéine. L'un des avantages de la DSF est la cadence à laquelle peuvent être réalisés les essais. En effet, il est possible de cribler 96 conditions en seulement 2 h. Ericsson et coll. (104) ont mené une étude sur l'effet de nombreux tampons, agents additifs et atomes lourds sur un échantillon de 21 protéines d'origines variées. Ainsi ils ont démontré que l'environnement des protéines pouvait être rapidement criblé pour lancer des essais de cristallisation dans des conditions optimales de stabilité de la protéine. Pour aller plus loin dans l'analyse des interactions protéine-ligand par DSF, Lo et coll. (105) ont, non seulement étudié l'effet de différents ligands sur le T_m d'une protéase à aspartate, mais aussi

mesuré la valeur de la constante de dissociation (K_d) à $T = T_m$. Pour cela, ils ont utilisé le travail développé par Pantoliano et coll. (106) sur le développement de technique de criblage à haut débit par DSF. Cette technique permet de mesurer par le T_m ainsi que le K_d à $T = T_m$ grace aux équations 5 et 6, et ceux pour un grand nombre d'échantillon permettant des applications dans le drug design.

L'équation de Boltzmann (Equation 4) devient l'équation 5:

Equation 5 : Equation de Boltzmann modifiée

$$y_{(T)} = y_u + \frac{(y_f - f_u)}{1 + \exp \left[-\Delta H_u / R [1/T - 1/T_m] + \Delta C_{pu} / R [\ln(T/T_m) + T_m/T - 1] \right]}$$

Avec :

- $y_{(T)}$: Fluorescence en fonction du temps
- y_f : Fluorescence pré transitionnelle
- y_u : Fluorescence post transitionnelle
- T_m : température de dénaturation, $T_m = T_0$ en absence de ligand
- ΔH_u : Enthalpie de dénaturation protéique, $\Delta H_u^{T_0} = \Delta H_u$ en absence de ligand
- ΔC_{pu} : Capacité calorifique de dénaturation protéique, $\Delta C_{pu}^{T_0} = \Delta C_{pu}$ en absence de ligand
- R : Constante de gaz parfaits

La constante d'association à $T = T_m$ est calculée selon l'équation 6 :

Equation 6 : Détermination de la constante d'association

$$K_L^{T_m} = \frac{\exp \left[-\Delta H_u^{T_0} / R [1/T_m - 1/T_0] + \Delta C_{pu}^{T_0} / R [\ln(T_m/T_0) + T_0/T_m - 1] \right]}{[L_{T_m}]}$$

Avec :

- $K_L^{T_m}$: Constante d'association à $T = T_m$
- $[L_{T_m}]$: Concentration en ligand libre à $T = T_m$

L'effet d'une ou plusieurs mutation(s) peut aussi être analysé par DSF. Hopper et coll. (107) ont déterminé le T_m d'une aminotransférase et d'un variant allélique mineur afin d'analyser la stabilité de ces deux protéines en absence et en présence de ligand d'intérêt. Le variant allélique mineur est moins stable que le variant majeur mais tous les deux présentent une

agrégation en deux temps avec deux T_m caractéristiques (54 et 66 °C pour le variant majeur et 45 et 61 °C pour le variant mineur), ces deux variants sont grandement stabilisés par l'addition de PLP (Pyridoxal 5'-Phosphate) dont le T_m passe à 78 °C et 73 °C pour le variant majeur et le variant mineur respectivement avec une cinétique de dénaturation en un seul temps en présence de ligand.

Une variante intéressante de la DSF est la dénaturation isotherme ITD (pour IsoThermal Denaturation en anglais) développée par Senisterra et coll. (108), le principe d'analyse reste le même que pour la DSF excepté que l'expérience se fait à une température constante, généralement choisie quelques degrés en dessous du T_m . Froese et coll. (109) ont montré l'effet d'une mutation ponctuelle (R161Q) sur la protéine MMACHC (MethylMalonic Aciduria cblC type HomoCystinuria) avec un T_m passant de 39,3 °C à 37,1 °C en présence de la mutation, de plus, les courbes de RFU=f(T) montre une lente dénaturation de la protéine sauvage, ayant lieu sur un large incrément de température, alors que la protéine mutante se dénature très rapidement. Cet effet a été vérifié par ITD à 35 °C dont le résultat a montré une lente dénaturation du mutant R161Q contrairement à la protéine sauvage. Ce type d'analyse peut être très bénéfique pour détecter des conditions favorisant l'agrégation des protéines qui génèrerait l'analyse par ITC.

La DSF est donc une technique rapide à mettre en œuvre (quelques heures), permettant des cribles en microplaques et demandant très peu d'échantillon biologique (entre 40 et 200 pmol de protéine par test). Les résultats obtenus sont reproductibles et simples à analyser.

IV-3 Cristallographie des protéines par diffraction des rayons X

IV-3.1 Introduction à la résolution de structure de protéines par cristallographie des rayons X

Pour observer la structure d'une protéine à l'échelle atomique par un système optique, il est nécessaire d'utiliser une source de rayonnement de longueur d'onde compatible avec les dimensions de l'objet à observer. Les distances entre atomes au sein des molécules étant de l'ordre de l'angström, la source lumineuse la plus appropriée est donc le rayonnement X dont la longueur d'onde est comprise entre 0,1 et 100 Å. L'utilisation de ce rayonnement, outre le fait qu'il s'agisse d'un rayonnement ionisant, pose plusieurs problèmes :

- ✓ la diffusion des rayons X par les électrons d'une molécule isolée est trop faible pour être mesurée. Afin d'amplifier cette interaction, il faut disposer d'un grand nombre de molécules identiques pour pouvoir additionner leur contribution. Pour assurer une diffusion cohérente des rayons X, ces molécules doivent être empilées de manière périodique dans les trois dimensions de l'espace. Ceci correspond à la définition d'un cristal. La première étape consiste donc en l'obtention de cristaux de protéines. Ceci constitue l'étape limitante de la méthode car les conditions conduisant à l'obtention de cristaux ne peuvent être déterminées qu'empiriquement,
- ✓ les rayons X, une fois diffractés, doivent être focalisés afin de reconstituer une image agrandie de la protéine or il n'existe pas de système optique capable d'effectuer cette focalisation. Cependant, des capteurs CCD sont capables d'enregistrer les images issues de la diffraction des rayons X (appelées clichés de diffraction), qui permettent, grâce à une fonction mathématique nommée Transformée de Fourier, de reconstituer "l'image" de la protéine (la carte de densité électronique),
- ✓ la résolution de cette opération mathématique nécessite deux informations particulières définissant une onde sinusoïdale : l'amplitude et la phase de chaque onde diffractée. L'amplitude est aisément mesurable, elle est directement proportionnelle à l'intensité des taches de diffraction mesurée sur les clichés de diffraction. La phase est plus délicate à obtenir puisque cette information n'est pas mesurable expérimentalement, elle doit être obtenue de manière indirecte lors d'une étape appelée phasage.

IV-3.2 Cristallisation des protéines

Un cristal est un empilement périodique d'une même molécule dans les trois directions de l'espace appelé réseau. Ce réseau est construit par trois vecteurs **a**, **b** et **c** et trois angles α , β et γ (ce sont les paramètres de maille) définissant la maille du cristal (plus petit élément permettant la reconstruction du cristal par translation du réseau). Une maille peut contenir plusieurs éléments, ainsi l'unité asymétrique est le plus petit élément permettant de reconstruire la maille par des opérations de symétrie. Les paramètres de maille permettent de décrire sept systèmes cristallins, de plus, l'arrangement des nœuds du réseau peut se faire selon quatre types de réseaux. Ceci définit les 14 réseaux de Bravais qui sont constitués par les combinaisons entre les systèmes cristallins et les types de réseaux (Figure 19).

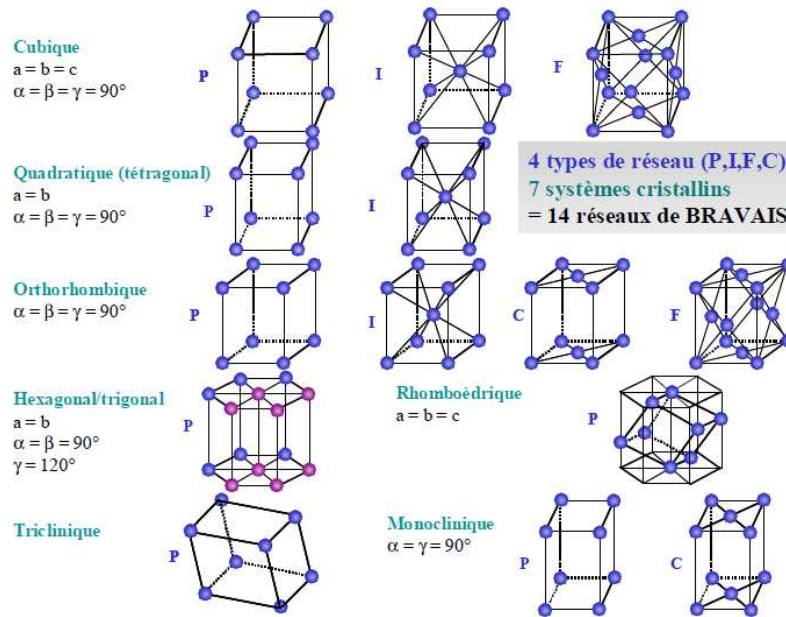


Figure 19 : Les 14 réseaux de Bravais

L'obtention d'un cristal de protéine est une des étapes limitantes de la cristallographie. En effet, de nombreuses conditions doivent être réunies afin d'obtenir un cristal. Il n'existe pas à ce jour, de moyen de prédire une condition de cristallisation. Ces conditions sont déterminées de manière empirique sur une échelle de temps s'étalant de quelques heures à quelques mois. Théoriquement, pour cristalliser, les protéines pures doivent se rapprocher les unes des autres, tout en conservant leur structure tertiaire, afin d'établir un ensemble d'interactions faibles assurant la cohésion de l'édifice (liaisons hydrogène, liaisons ioniques, forces de Van Der Walls). Ces interactions sont appelées contacts cristallins. La cristallisation d'une protéine repose sur un processus composé de trois étapes observable sur un diagramme de phase (Figure 20) :

- ✓ la sursaturation : un état où la protéine est en concentration supérieure à sa limite de solubilité. La concentration en protéine est de l'ordre de la dizaine de mg.mL^{-1} . L'état métastable de sursaturation va contraindre le rapprochement des protéines en solution et évoluer, soit vers la précipitation, soit vers la nucléation des protéines,
- ✓ la nucléation : dans certaines conditions, l'état de sursaturation va favoriser l'apparition de germes cristallins,
- ✓ la croissance cristalline : un retour à l'équilibre par l'incorporation des macromolécules protéiques à l'édifice cristallin. Une fois revenue à la limite de solubilité de la protéine, la croissance s'arrête.

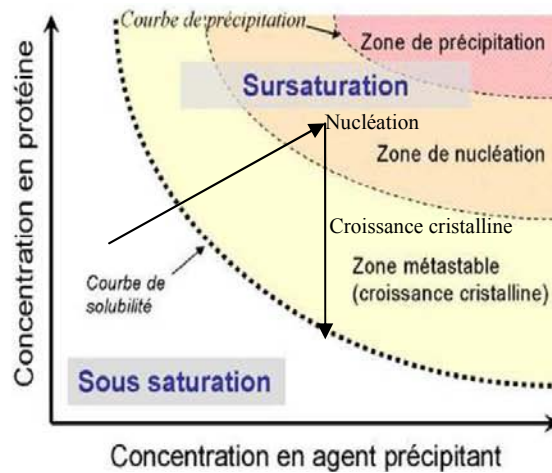


Figure 20 : Diagramme de phase

Le succès de la cristallisation d'une protéine repose sur un ensemble de paramètres physico-chimiques variés (110), tels que la nature chimique, la concentration, le pH du tampon utilisé et la nature de l'agent précipitant (qui permet de réduire la solubilité de la protéine), mais aussi de la température, la nature et la concentration des sels (effets salting-in et salting-out), ...

Il existe différentes techniques permettant d'amener les protéines à cristalliser (111), la plus utilisée est la technique de diffusion de vapeur. Une goutte contenant la solution de protéine mélangée à la solution de cristallisation est équilibrée contre un réservoir contenant la solution de cristallisation. La diffusion de vapeur va se faire de la goutte vers le réservoir, augmentant la concentration en protéine au sein de la goutte lentement jusqu'à l'entrée en sursaturation. Il existe deux variantes à cette technique :

- ✓ la diffusion de vapeur en goutte assise : généralement utilisée lors d'un criblage robotisé de conditions de cristallisation, minimisant le volume de la goutte (100 à 200 nL) et donc la quantité d'échantillon biologique nécessaire. Ceci permettant une recherche plus exhaustive de conditions permettant d'obtenir des objets cristallins,
- ✓ la diffusion de vapeur en goutte suspendue : essentiellement utilisée pour optimiser manuellement une condition de cristallisation préalablement identifiée par un criblage robotisé. Dans ce cas, les gouttes sont de l'ordre du μL décuplant ainsi les quantités de protéine nécessaires.

Une fois la structure d'une protéine obtenue sous forme *apo*, il peut être intéressant d'obtenir la structure d'un complexe entre la protéine et un ligand d'intérêt. A ce titre, deux stratégies sont envisageables : le trempage des cristaux dans une solution contenant le ligand d'intérêt ou bien la co-cristallisation, présentant chacune plusieurs variantes (112).

Le trempage nécessite, d'une part, de posséder un cristal de la protéine à étudier et, d'autre part, que le cristal résiste à son immersion dans une solution contenant le ligand en forte concentration. Cette méthode repose sur le fait qu'un cristal est en grande partie constitué de solvant (la proportion de solvant pouvant atteindre 80 % du volume cristallin), faisant ainsi apparaître des canaux de solvant qui permettent aux ligands de diffuser à l'intérieur du cristal jusqu'à un éventuel site de fixation sur la protéine. C'est la technique la plus souvent utilisée en premier lieu pour obtenir une structure de complexe protéine-ligand. Le trempage classique se fait en plongeant un cristal de protéine dans une solution de ligand solubilisé dans une condition la plus proche possible de la condition de cristallisation. En faisant varier le temps et/ou la concentration en ligand, deux approches peuvent être envisagées : un trempage bref (quelques secondes/minutes) dans une solution de concentration élevée en ligand ou un trempage plus long (quelques heures/jours) dans une solution contenant une faible concentration en ligand, le choix entre ces deux approches dépendant généralement de la solubilité du ligand, de son affinité et de la résistance du cristal au trempage. En cas de très faible solubilité du ligand, une approche par trempages successifs peut être entreprise, ce qui permet d'augmenter la concentration *in crystallo* en ligand. Elle consiste à tremper le cristal de protéine dans une solution de faible concentration en ligand pendant un temps donné et à répéter ce trempage plusieurs fois dans une solution de ligand fraîche en augmentant ou pas progressivement la concentration en ligand dans la goutte de trempage. Lorsque le cristal de protéine a été obtenu en présence d'un ligand ou d'un additif stabilisant la protéine, un trempage peut se révéler difficile si le site de fixation est le même que celui du ligand d'intérêt. Dans le cas d'un échange de ligand, l'affinité joue un rôle crucial, l'additif devant avoir dans les conditions cristallines une affinité moindre pour être "délogé".

Lorsque les résultats de trempage ne se révèlent pas optimum, la co-cristallisation est alors envisagée. Co-cristalliser la protéine en présence du ligand (qui peut être ajouté à différents stades : production, purification ou cristallisation, le ligand devant être conservé tout au long des étapes ultérieures à son introduction) peut se faire dans les mêmes conditions que la protéine *apo* ou dans des conditions radicalement différentes nécessitant de rechercher par criblage de nouvelles conditions dans le cas où la fixation du ligand empêche l'établissement

d'un ou de plusieurs contact(s) cristallin(s) ou entraîne un changement de conformation de la protéine. Là aussi des phénomènes de compétition entre différents ligands sont à prendre en compte, en cas d'ajout d'additifs.

IV-3.3 Diffraction des rayons X

Lorsqu'un cristal est soumis à un faisceau de rayons X, il se produit le phénomène de diffraction. En raison de la périodicité du cristal, les ondes diffractées interfèrent entre elles et ne seront constructives que dans certaines directions de l'espace comme défini par les équations de Laue. Soit \mathbf{u}_0 , le vecteur d'onde incidente et \mathbf{u} un vecteur d'onde diffusée. L'angle unissant \mathbf{u}_0 et \mathbf{u} est de 2θ . Le vecteur de diffusion est défini comme étant $\mathbf{S} = \mathbf{u} - \mathbf{u}_0$. Un réseau dont les vecteurs de base sont \mathbf{a} , \mathbf{b} et \mathbf{c} ne permettra d'observer que les ondes diffractées dans les directions de l'espace vérifiant les équations de Laue :

- ✓ $\mathbf{a} \cdot \mathbf{S} = h$ (définit les plans orthogonaux à \mathbf{a} et équidistants de $1/a$),
- ✓ $\mathbf{b} \cdot \mathbf{S} = k$ (définit les plans orthogonaux à \mathbf{b} et équidistants de $1/b$),
- ✓ $\mathbf{c} \cdot \mathbf{S} = l$ (définit les plans orthogonaux à \mathbf{c} et équidistants de $1/c$),

où h , k et l sont des entiers.

Les plans h , k et l définissent alors le réseau réciproque.

La loi de Bragg formalise différemment les conditions de Laue et introduit la sphère d'Ewald permettant de schématiser les conditions de diffraction. Le cristal est représenté comme un ensemble de miroirs appelés plans de Bragg et reliant des points équivalents dans le cristal réfléchissant les rayons X (c'est pourquoi une tache de diffraction peut également être appelée une réflexion). Les conditions de diffraction sont vérifiées par l'équation 7 : $2d \sin(\theta) = n\lambda$, où d est la distance séparant deux plans du réseau réciproque de la même famille (parallèles entre eux) et λ la longueur d'onde du faisceau incident. Seuls les nœuds du réseau réciproque vérifiant cette équation étant sur la sphère d'Ewald (une sphère centrée sur le cristal et de rayon $1/\lambda$, Figure 21) et seront en condition de diffraction.

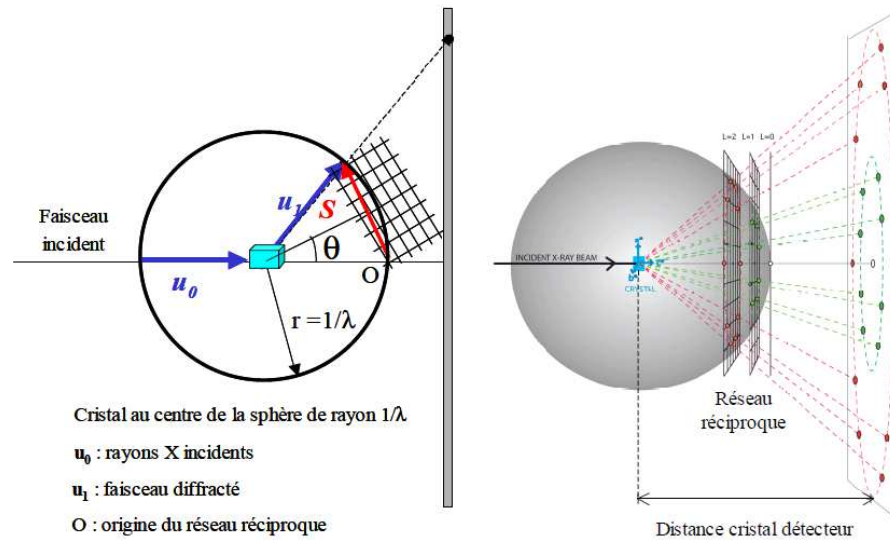


Figure 21 : Construction de la sphère d'Ewald

Afin d'enregistrer un jeu de données de diffraction, le cristal doit être irradié par un faisceau de rayons X. Les rayons X sont générés classiquement par deux manières :

- ✓ le générateur à anode tournante dont le principe réside dans le bombardement d'une anode métallique (généralement constituée de cuivre en biologie) par des électrons. Les électrons des atomes métalliques vont alors transiter par un état énergétique plus élevé et émettre des rayons X en retournant à leur état initial. Un monochromateur va sélectionner une raie caractéristique d'émission du métal : la raie $K\alpha$, dont la longueur d'onde est de 1,54 Å pour le cuivre,
- ✓ le synchrotron : le rayonnement est émis par des électrons accéléré à une vitesse proche de celle de la lumière et dont la trajectoire est déviée par un champ magnétique. Différents aimants sont situés sur le parcours des électrons dans l'anneau de stockage (aimant de courbure, éléments d'insertion) permettent d'obtenir un faisceau de rayons X intenses, polychromatiques et de petite taille. Un système de monochromateur permet de sélectionner une longueur d'onde donnée.

Pour une orientation du cristal donné, il n'y a que certains nœuds du réseau réciproque qui se trouvent sur la sphère d'Ewald. Ainsi, pour enregistrer l'ensemble des réflexions, la méthode du cristal tournant est utilisée et consiste à faire tourner le cristal autour d'un axe perpendiculaire au faisceau des rayons X. Différents paramètres sont à prendre en considération lors de l'enregistrement des données :

- ✓ La distance cristal-détecteur qui gouverne la résolution maximale enregistrable. Il faut utiliser au maximum toute la gamme dynamique du détecteur tout en permettant d'enregistrer les taches de diffraction les plus éloignées du centre du détecteur,
- ✓ L'angle de rotation total est l'angle entre le premier et le dernier cliché de diffraction et dépend de la géométrie du cristal (groupe d'espace et paramètre de maille). Il détermine la complétude d'un jeu de donnée (c'est-à-dire le nombre de réflexions observées par rapport au nombre de réflexions théoriquement observables),
- ✓ L'angle d'oscillation est l'angle de rotation durant lequel est enregistré un cliché de diffraction. Typiquement cet angle d'oscillation est de 1° ,
- ✓ Le temps d'exposition régit l'intensité des taches de diffraction, plus le temps d'exposition est long, plus les taches seront intenses. Cependant, les rayons X entraînent des dégradations radiatives sur le cristal pouvant diminuer drastiquement la limite de diffraction voire conduire à une perte complète de diffraction du cristal avant la fin de l'enregistrement. Il est donc important de trouver un compromis entre intensité et qualité de diffraction sur l'ensemble du jeu de données.

Les dégâts d'irradiation dus à l'impact du faisceau de rayons X sur le cristal (production de radicaux libres, échauffement du cristal) peuvent être limités en refroidissant le cristal à une température proche de 100 K lors de l'enregistrement des données. Cependant, ceci nécessite de cryoprotéger le cristal au préalable. Ceci se fait le plus souvent par trempage dans un agent cryoprotectant (éthylène glycol, glycérol, sucre, ...) qui limitera l'apparition de cristaux de glace et donc l'apparition d'anneaux sur les clichés de diffraction dus à la diffraction des rayons X par les cristaux de glace ainsi formés.

IV-3.4 Traitement des données de diffraction

L'indexation consiste, à partir de quelques clichés de diffraction (typiquement 2 ou 3), à assigner à chaque tache de diffraction les indices du nœud du réseau réciproque hkl (indice de Miller) lui correspondant. Ceci va permettre de rechercher un tenseur de 3 vecteurs non colinéaires dont la norme et l'angle fait entre eux vont servir à définir les paramètres de maille pour les 14 réseaux de Bravais, la solution choisie correspondant au réseau présentant la plus haute symétrie. Ceux-ci sont ensuite affinés en utilisant des données réparties sur l'ensemble du jeu de données afin de s'affranchir d'éventuelles erreurs sur les paramètres expérimentaux, notamment la distance cristal-détecteur.

L'intégration va permettre de mesurer l'intensité de chaque tache enregistrée. Pour cela une boîte de mesure est placée autour de chaque tache, l'intensité de chaque pixel définissant la tache est alors mesurée et sont intégrées afin de calculer l'intensité de la tache. L'intensité de la zone périphérique à la tache, constituant le bruit de fond, est également mesurée et retranchée à celle de la tache.

La mise à l'échelle va réduire l'ensemble des taches de diffraction à un jeu de réflexions uniques et moyennent l'intensité d'une même réflexion et de celles reliées par symétrie cristallographique. La variation d'intensité peut être due à des variations d'intensité du faisceau incident, ou à des dommages radiatifs sur le cristal abaissant son pouvoir diffractant. Enfin, les intensités sont converties en amplitude des ondes diffractées.

À l'issue de la mise à l'échelle, un certain nombre de paramètres permettent de juger de la qualité des jeux de données :

- ✓ Le R_{sym} (Equation 8) ou R_{merge} dans le cas de l'utilisation de plusieurs cristaux pour enregistrer un jeu complet, renseigne sur la dispersion des intensités entre les différentes observations d'une même réflexion et de ses équivalentes.

Equation 8 :

$$R_{\text{sym}} = \frac{\sum_{\text{hkl}} \sum_i | I_i(\text{hkl}) - \langle I(\text{hkl}) \rangle |}{\sum_{\text{hkl}} \langle I(\text{hkl}) \rangle}$$

Avec :

- $I_i(\text{hkl})$: intensité de la $i^{\text{ème}}$ mesure équivalente
- $\langle I(\text{hkl}) \rangle$: moyenne des intensités pondérée par toutes les mesures $I_i(\text{hkl})$

Ainsi, plus le R_{sym} tendra vers 0, plus la dispersion des intensités de réflexion équivalente sera faible.

- ✓ Le rapport signal sur bruit $I/\sigma(I)$ permet d'estimer l'intensité moyenne des réflexions par rapport à l'erreur sur les mesures d'intensités.
- ✓ La complétude traduit le fait qu'un jeu de données enregistré puisse être plus ou moins complet, cette valeur doit être voisine de 100 % car chaque réflexion indépendante manquante est un terme en moins pour calculer la carte de densité électronique.
- ✓ La multiplicité renseigne sur le nombre de réflexions équivalentes enregistrées pour une même réflexion unique. Plus cette valeur est élevée, meilleure est l'estimation de l'intensité de chaque réflexion et de l'erreur associée.

IV-3.5 Phasage

L'information de phase de chacune des ondes diffractées est perdue au cours de l'enregistrement des données de diffraction, elle est donc obtenue de manière indirecte. Il existe différentes techniques pour obtenir l'information de la phase :

- ✓ le remplacement moléculaire : utilisant la structure connue d'une molécule homologue comme source de phases,
- ✓ le remplacement isomorphe (SIR, MIR, SIRAS, MIRAS) : nécessitant l'introduction d'un ou de plusieurs atomes riches en électrons dans le réseau cristallin,
- ✓ les méthodes exploitant la diffusion anormale de certains atomes dans la gamme des rayons X (SAD et MAD).

Nous ne parlerons ici que du phasage par remplacement moléculaire qui est la seule méthode utilisée lors de ces travaux.

Le Phasage par remplacement moléculaire repose sur l'utilisation de la structure tridimensionnelle connue d'une protéine homologue afin de déterminer les phases à associer aux amplitudes obtenues expérimentalement. En effet, si les structures primaires de deux protéines sont conservées (20 à 30 % d'identité de séquence au minimum) alors leur repliement est très probablement similaire. La structure de la protéine homologue étant connue, il est possible de générer un édifice cristallin similaire à celui observé expérimentalement et de calculer les ondes diffractées théorique à la fois en amplitude et en phase. On peut alors en extraire l'information de phase. Afin de générer cet édifice cristallin à partir de la structure de la protéine "modèle", la méthode utilise la fonction de Patterson (Equation 9) qui permet de calculer des vecteurs interatomiques à partir des seules amplitudes expérimentales, s'affranchissant ainsi complètement de l'information de phase.

Equation 9 : Fonction de Patterson $P(\mathbf{u})$

$$P(\mathbf{u}) = \int \rho(\mathbf{r}) \cdot \rho(\mathbf{r} + \mathbf{u}) \cdot dV = \sum_{hkl} |F_{hkl}|^2 \cdot e^{2i\pi \cdot \mathbf{u} \cdot \mathbf{S}}$$

Avec :
➤ \mathbf{u} : vecteur interatomique

Le remplacement moléculaire correspond à une recherche à 6 dimensions (3 rotations et 3 translations). Cette recherche est décomposée en deux opérations successives à 3 paramètres :

dans un premier temps, une fonction de rotation (Equation 10) oriente la structure du modèle dans la même orientation que la protéine étudiée, puis, une fonction de translation (Equation 11) positionne la structure du modèle préalablement orienté dans la maille cristalline de la protéine étudiée.

La fonction de Patterson (représentant les vecteurs interatomiques) est calculée pour la protéine modèle et pour la protéine de structure inconnue. Le but étant d'obtenir des fonctions de Patterson observée et calculée similaires, ce qui est le cas lorsque le modèle est correctement orienté et placé à la bonne position dans la maille cristalline.

Les vecteurs interatomiques calculés par la fonction de Patterson peuvent être subdivisés en deux catégories :

- ✓ Les vecteurs interatomiques intramoléculaires dépendent uniquement de l'orientation de la molécule dans la maille, ce sont les vecteurs utilisés dans la fonction de rotation

Equation 10 : Fonction de rotation $R(\kappa, \phi, \psi)$

$$R(\kappa, \phi, \psi) = \int_{r_{\min}}^{r_{\max}} P_{\text{nat}}(U) \cdot P_{\text{mod}}(\kappa, \phi, \psi, U) \cdot dU$$

Avec :

- (κ, ϕ, ψ) : opération de rotation
- P_{nat} : fonction de Patterson de la structure à déterminer
- P_{mod} : fonction de Patterson de la structure modèle
- r_{\min} et r_{\max} : rayon d'intégration réduisant l'ensemble des vecteurs aux vecteurs intramoléculaires

Lorsque P_{mod} et P_{nat} se superposent, la fonction de rotation est à son maximum et indique la ou les orientation(s) optimales de la protéine dans la maille

- ✓ Les vecteurs interatomiques intermoléculaires dépendent de l'orientation et de la position de la molécule dans la maille, il est donc nécessaire de déterminer l'orientation de la molécule dans la maille avant d'effectuer la fonction de translation. La corrélation sera maximale pour un vecteur de translation \mathbf{t} permettant de relier deux molécules voisines

Equation 11 : Fonction de translation T(t)

$$T(\mathbf{t}) = \int_{V_{\text{maille}}} P_{\text{nat}}(\mathbf{U}) \cdot P_{\text{rot}}(\mathbf{U}-\mathbf{t}) \cdot d\mathbf{U}$$

Avec :

- V_{maille} : volume de la maille
- Prot : fonction de Patterson issue du modèle orientée

Ces deux types de vecteurs sont différenciables car statistiquement, les vecteurs interatomiques intramoléculaires sont plus courts que les vecteurs interatomiques intermoléculaires.

Une fois le modèle correctement positionné, une transformée de Fourier permet de calculer l'information de phase. Il est alors nécessaire d'évaluer la qualité de cette information. Il existe différents paramètres comme les scores utilisés pour juger les solutions des fonctions de rotation et translation qui doivent être maximisés, les meilleures solutions de ces fonctions présentant un écart de score important avec les suivantes. Les déviations standards, informant du respect des contraintes géométriques, le facteur d'accord cristallographique R (Equation 12) qui doit avoisiner les 50 % en sortie de remplacement moléculaire ou le coefficient de corrélation C (Equation 13) représentant la corrélation entre les facteurs de structure expérimentaux F_{obs} et les facteurs de structure calculés à partir d'un modèle positionné dans la maille F_{calc} , il doit être maximisé (une valeur supérieure à 30 % est considérée comme correcte).

Equation 12 : Calcul du facteur d'accord cristallographique R

$$R = \frac{\sum_{\text{hkl}} \left| |F_{\text{obs}}| - |F_{\text{calc}}| \right|}{\sum_{\text{hkl}} |F_{\text{obs}}|}$$

Avec :

- F_{obs} : facteur de structure observé
- F_{calc} : facteur de structure calculé à partir du modèle correctement orienté dans la maille

Equation 13 : Calcul du coefficient de corrélation C

$$C = \frac{\sum_{\text{hkl}} \left[|F_{\text{obs}}|^2 - \overline{|F_{\text{obs}}|^2} \right] * \left[|F_{\text{calc}}|^2 - \overline{|F_{\text{calc}}|^2} \right]}{\sqrt{\sum_{\text{hkl}} \left[|F_{\text{obs}}|^2 - \overline{|F_{\text{obs}}|^2} \right] * \left[|F_{\text{calc}}|^2 - \overline{|F_{\text{calc}}|^2} \right]}}$$

Il est alors possible de calculer une carte de densité électronique (Equation 14) issue du phasage par RM. Le modèle n'étant pas complet et étant entaché d'erreurs, il convient ensuite d'effectuer des cycles de construction manuelle et d'affinement de la structure ainsi construite afin de minimiser les erreurs sur les phases.

Equation 14 : Densité électronique

$$\rho(x,y,z) = \frac{1}{V_{\text{maille}}} \sum_{h=-\infty}^{\infty} \sum_{k=-\infty}^{\infty} \sum_{l=-\infty}^{\infty} F_{hkl} \cdot e^{-2.i.\pi.(hx+ky+lz)}$$

IV-3.7 Affinement de la structure

Une fois le phasage effectué, les phases calculées ne sont qu'approximatives. Les erreurs doivent donc être réduites par des cycles de construction/affinement successifs. Chaque atome est décrit par quatre paramètres pouvant être affinés, les coordonnées spatiales x,y,z et le facteur d'agitation thermique B. Concrètement, lors de corrections manuelles, les structures primaire, secondaire et tertiaire d'une protéine doivent être reconstruites en accord avec les cartes de densité électronique, il y a deux types de carte : la $2F_{\text{obs}}-F_{\text{calc}}$, qui représente la carte de densité électronique et la $F_{\text{obs}}-F_{\text{calc}}$, ou Fourier différence, qui permet de visualiser un excès ou un défaut de densité électronique. Une fois le cycle de construction manuelle achevé, un cycle d'affinement automatique est effectué où quatre paramètres par atome sont affinés : les trois coordonnées spatiales x, y, z et le facteur d'agitation thermique B.

L'affinement automatique peut être suivi par certains paramètres comme les déviations standards, informant du respect des contraintes géométriques, le facteur d'accord cristallographique R (Equation 12) qui doit diminuer au cours de l'affinement ou le coefficient de corrélation C (Equation 13) qui doit être maximisé.

IV-3.8 Validation de la structure

Le diagramme de Ramachandran donne la répartition des angles dièdres permis pour la chaîne principale (couples phi et psi d'une protéine) sur une grille colorée. Sur ce même diagramme est indiqué l'ensemble des couple ϕ et ψ de la structure d'intérêt. La structure est validée d'un

point de vue géométrique si la très grande majorité des couples φ et ψ se retrouve dans les zones permises (Figure 22).

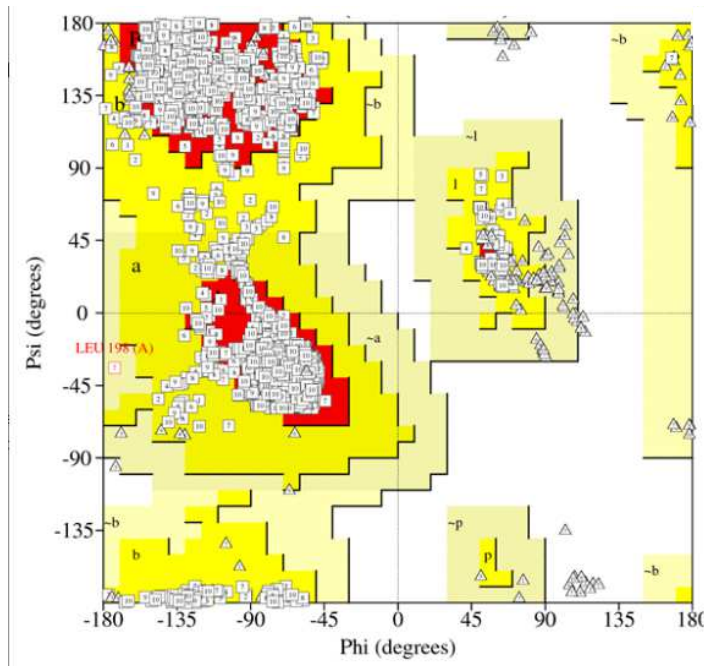


Figure 22 : Diagramme de Ramachandran

En rouge : zones favorables, en jaune : zones autorisées, en beige : zones limitées
et en blanc : zones interdites

Un autre critère de validation est le couple R et R_{libre} . En effet, le facteur d'accord cristallographique R se calcule avec des données utilisées lors de l'affinement qui a pour but de le minimiser. Sa valeur n'est donc pas objective pour juger de l'affinement d'une structure. Le R_{libre} est calculé de la même manière que le R mais avec des données écartées de l'affinement ($\sim 5\%$), la diminution concomitante du R et du R_{libre} est donc un bon critère pour suivre l'affinement et évaluer l'accord entre les données de diffraction calculées à partir du modèle et les données expérimentales.

IV-4 Modélisation moléculaire

Les techniques expérimentales ne permettent pas toujours d'accéder aux informations structurales et réarrangements conformationnels pouvant se produire au sein d'une molécule au cours du temps ou sous l'influence de son environnement (ligand, solvant, ...). Pour pallier l'insuffisance de données expérimentales, la modélisation moléculaire et les différentes techniques de simulations associées sont souvent utilisées pour visualiser, comprendre et

prédire les propriétés géométriques et physiques des systèmes moléculaires, et à terme, l'influence de modifications structurales sur ces propriétés.

Parmi les techniques les plus répandues, la mécanique moléculaire (MM) permet, par des calculs statiques, d'étudier la conformation de (macro)molécules et ses interactions mais sans donner toutefois d'information sur la dynamique des molécules et l'échelle de temps sur laquelle se déroulent les mouvements. Les techniques de simulation de dynamique moléculaire (DM) permettent quant à elles de simuler l'évolution d'un système moléculaire au cours du temps et d'acquérir ainsi une vision dynamique qui complète les données expérimentales fournies par cristallographie des rayons X ou diffusion des rayons X aux petits angles.

Nous développerons brièvement ci-après ces deux approches mises en œuvre au cours de la thèse.

IV-4.1 La mécanique moléculaire

La mécanique moléculaire est une méthode empirique de modélisation moléculaire dans laquelle seuls les mouvements des noyaux des atomes sont pris en compte, les électrons ne sont pas explicitement inclus dans le calcul. Dans cette approche, la molécule est considérée comme un ensemble de sphères interagissant entre elles par des forces décrites dans des fonctions d'énergie potentielle (E_{total}) définissant le "champ de force".

La somme des différents termes de E_{total} se décompose typiquement selon l'équation 15 :

Equation 15 : Fonction d'énergie potentielle

$$E_{\text{total}} = E_{\text{liaison}} + E_{\text{angle}} + E_{\text{torsion}} + E_{\text{impropre}} + E_{\text{vdW}} + E_{\text{electro}} + E_{\text{liaisons-H}}$$

Avec :

- E_{liaison} = Energie d'étirement et de compression de la liaison
- E_{angle} = Energie de déformation d'angle
- E_{torsion} = Energie de torsion
- E_{impropre} = Energie de déformation hors du plan (Energie impropre)
- E_{vdW} = Energie d'interaction de van der Waals
- E_{electro} = Energie électrostatique
- $E_{\text{liaisons-H}}$ = Energie due aux liaisons hydrogène

Les méthodes de mécanique classique reposent sur des paramètres de référence établis à partir de données expérimentales, telles que la diffraction des rayons X et les techniques spectroscopiques infrarouge et Raman. La forme analytique de la fonction d'énergie potentielle correspond toujours à un choix, les différentes méthodes empiriques se différenciant en fonction du degré de développement de la fonction d'énergie potentielle et par les paramètres utilisés.

Ces méthodes présentent des avantages indéniables puisqu'elles sont extrêmement rapides et ne nécessitent pas une très large puissance de calcul. Elles sont applicables à des molécules de taille importante (~100 000 atomes) et sont simples à mettre en œuvre.

Chacun des termes décomposant l'expression de l'énergie potentielle possède une valeur d'équilibre préférentielle (longueur de liaison, angle de valence, ...) pour laquelle l'énergie est minimale. La recherche du minimum énergétique local consiste, à partir d'une géométrie initiale, à rechercher le jeu de coordonnées cartésiennes qui réduit à son minimum la somme de toutes les interactions énergétiques. Les méthodes couramment utilisées reposent sur le calcul des dérivées premières (méthodes de la plus grande pente et du gradient conjugué) et seconde de l'énergie potentielle (méthode Newton-Raphson).

Parfois basées sur des techniques de mécanique moléculaire, les méthodes d'amarrage moléculaire (docking) permettent de prédire le mode de fixation d'un ligand (ou d'une protéine) sur une macromolécule cible de structure connue (ou modélisée) et d'élucider ainsi les interactions mises en jeu, au niveau atomique, dans la reconnaissance protéine-ligand.

IV-4.2 La dynamique moléculaire

La notion de mouvement peut se révéler indispensable dans la compréhension d'une fonction biologique précise. En effet, il existe tout un paysage énergétique (Figure 23A) pour une macromolécule donnée dont les différents minima représentent autant de conformations stables de la macromolécule. Les mouvements moléculaires se font sur différentes échelles de temps, allant de la fluctuation atomique (femtoseconde) aux larges mouvements tels que ceux des domaines protéiques ou de déplacement de ligand au sein d'un site actif (nanoseconde-seconde) (Figure 23B).

Cette notion de mouvement/flexibilité moléculaire est mal appréhendée par les calculs de mécanique moléculaire qui ne donnent aucune information sur le temps des mouvements et qui ne permettent pas de suivre les réarrangements conformationnels. Les mouvements peuvent en revanche être étudiés, au moins partiellement, par des techniques de dynamique

moléculaire qui prennent en compte les forces (dérivées de l'énergie potentielle) et les vitesses (associées à l'énergie cinétique) pour calculer la position des atomes en fonction du temps. Les simulations se font dans des conditions spécifiques de pression, de volume et de température et reposent sur l'échantillonnage de l'espace conformationnel du système étudié par l'intégration des équations du mouvement de Newton pour tous les atomes du système. Pour cela, les atomes du système étudié sont considérés comme des masses dont la mobilité est déterminée par un ensemble de forces exercées par les autres atomes en fonction du temps. Les intervalles de temps de chaque cycle de dynamique sont suffisamment courts (~ 1 fs) pour que l'on puisse considérer les forces interatomiques comme constantes, on peut alors intégrer les équations du mouvement de Newton et déterminer les positions des atomes de la molécule dans le temps. Pour cela, les simulations de dynamique moléculaire utilisent des algorithmes tels que celui de Verlet (113, 114) qui implique la conservation de l'énergie du système au cours de la simulation. Le potentiel d'énergie total utilisé comprend les contributions énergétiques de liaison, d'angle de valence, d'angle dièdre, des interactions de van der Waals et électrostatiques.

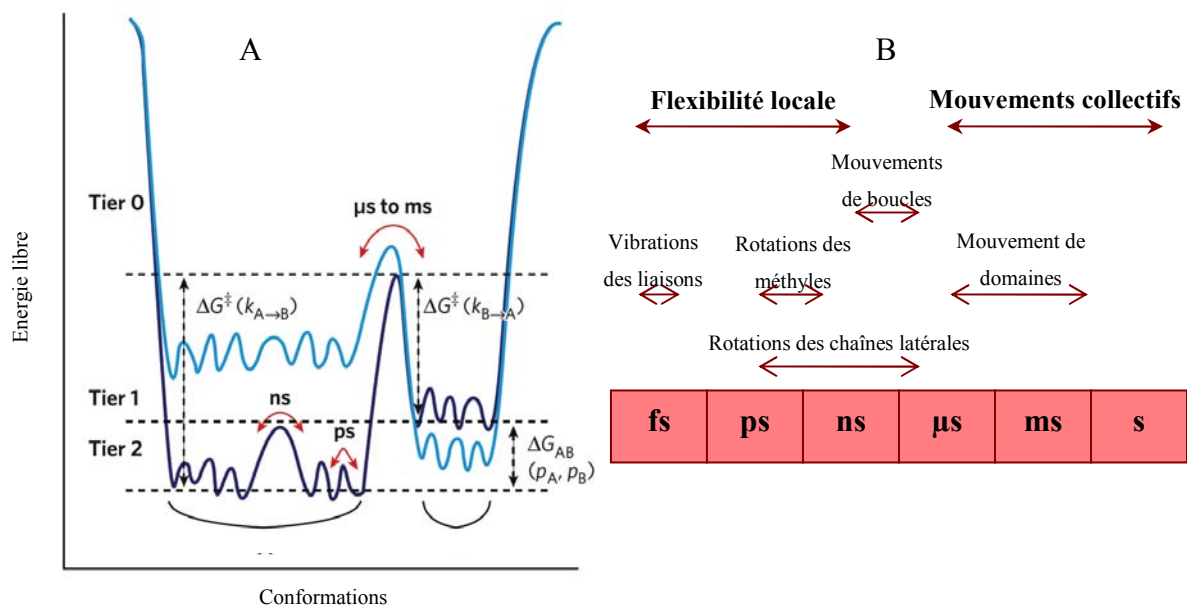


Figure 23 : Dynamique moléculaire

(A) Paysage énergétique
 (B) Echelle de temps des mouvements observés par dynamique moléculaire
 D'après Henzler-Wildman (115)

IV.5 La calorimétrie de titrage isotherme

L'ITC, (pour Isothermal Titration Calorimetry), est une technique permettant d'avoir accès aux constantes d'association lors d'interactions protéine-protéine, protéine-ligand (ADN, lipides, sucre, ...). Cette méthode permet de déterminer la stœchiométrie, l'énergie libre de Gibbs, l'enthalpie, l'entropie et la capacité calorifique associées à l'interaction protéine-ligand (116–118).

L'appareillage se compose de trois parties : une cellule de référence, une cellule de mesure toute les deux enfermées dans un caisson adiabatique et une seringue d'injection (Figure 24).

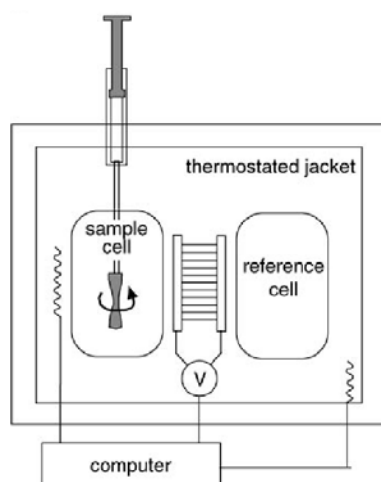
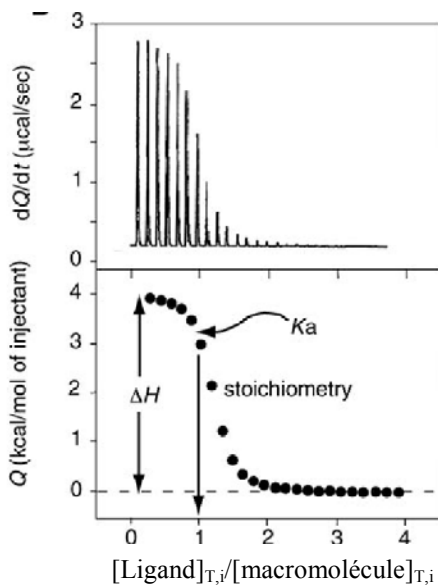


Figure 24 : Illustration schématique d'un microcalorimètre

D'après Velázquez et coll. (116)

Le produit titré (une macromolécule M par exemple) est généralement déposé dans la cellule de mesure, le titrant (un ligand L par exemple) est introduit dans la seringue d'injection, souvent 10 fois plus concentré que le titré (afin d'introduire le plus faible volume possible pour ne pas modifier le système par dilution) et en solution dans un tampon strictement identique. Le titrant est injecté petit à petit par palier dans le titré. La formation de complexe macromolécule-ligand (ML) s'accompagne d'un dégagement de chaleur (réaction exothermique) ou d'une absorption de chaleur (réaction endothermique) dans la cellule de mesure. Un rétrocontrôle assure un maintien de l'équilibre du système (ΔT entre les deux cellules = 0°C). L'énergie que doit fournir le système pour conserver le $\Delta T=0^\circ\text{C}$ est mesurée en $\mu\text{cal}\cdot\text{s}^{-1}$. Le signal de sortie est donné sous forme de courbe $dQ/dt=f(t)$ appelé thermogramme. L'aire de chaque pic est ensuite intégrée et normalisée par mole de ligand injectée, la courbe $Q=f([L]_{T,i}/[M]_{T,i})$ est alors tracée (Figure 25) où $[L]_{T,i}$ est la concentration

en ligand à l'injection i et à température T et $[M]_{T,i}$ est la concentration en macromolécule à l'injection i et à température T . Ces valeurs sont obtenues grâce aux équations 16-17-18. La concentration en complexe ML est donnée par l'équation 19. L'ITC est une technique très employée et fiable pour déterminer des constantes d'affinité. Les temps d'analyse sont raisonnables mais ne permettent pas de criblage à haut ou moyen débit. En fonction de l'affinité de la protéine pour son ligand, les quantités de matériel biologique peuvent se révéler importantes, avec une concentration pouvant dépasser la limite de solubilité, aussi bien pour la protéine que pour le ligand.



Plusieurs informations peuvent être déduites de la courbe $Q=f([Ligand]_{T,i}/[macromolécule]_{T,i})$:

- ✓ L'enthalpie de liaison ΔH : différence entre Q_{max} et Q_{min}
- ✓ La constante d'association K_a : pente autour du point d'inflexion
- ✓ La stœchiométrie n : valeur de l'abscisse au point d'inflexion

Figure 25 : Exemple d'un résultat obtenu

D'après Velázquez et coll. (116)

Equation 16 : Concentration en macromolécule à l'injection i

Avec :

$$[M]_{T,i} = [M]_0 \left(1 - \frac{v}{V} \right)^i$$

- $[M]_{T,i}$: concentration en macromolécule M après une injection i à température T
- $[M]_0$: concentration initiale en macromolécule M
- v : volume d'injection
- V : volume de la cellule de mesure

Equation 17 : Concentration en ligand à l'injection i

Avec :

$$[L]_{T,i} = [L]_0 \left(1 - \left(1 - \frac{v}{V} \right) \right)^i$$

- $[L]_{T,i}$: concentration en ligand L après une injection i à température T
- $[L]_0$: concentration initiale en ligand L
- v : volume d'injection
- V : volume de la cellule de mesure

Equation 18 : Quantité de chaleur

Avec :

$$q_i = V \Delta H_a \left([ML]_i - [ML]_{i-1} \left(1 - \frac{v}{V} \right) \right)$$

- q_i : quantité de chaleur libéré à l'injection i
- ΔH_a : Enthalpie d'association
- $[ML]_i$: concentration en complexe macromolécule-ligand ML après l'injection i

Equation 19 : Concentration en complexe à l'injection i

$$[ML]_i = \frac{1 + n[M]_{T,i} K_a + K_a[L]_{T,i} - \sqrt{(1 + n[M]_{T,i} K_a + K_a[L]_{T,i})^2 - 4n[M]_{T,i} K_a^2 [L]_{T,i}}}{2K_a}$$

Avec :

- n : stœchiométrie
- K_a : constante d'association

La principale application de l'ITC est la mesure de la constante d'association et des paramètres thermodynamiques d'une interaction protéine-ligand (sans passer par une étape de fixation d'un des deux partenaires sur un support, ce qui est un avantage indéniable en regard du biacor). Nous pouvons distinguer deux cas particuliers : macromolécule-ligand et macromolécule-macromolécule.

Les interactions protéine-ligand sont idéalement mesurées lorsque le K_a est compris entre $10^4 < K_a < 10^8 \text{ M}^{-1}$ (Figure 26). Pour une interaction trop faible ($K_a \ll 10^4 \text{ M}^{-1}$) ou trop forte ($K_a \gg 10^8 \text{ M}^{-1}$), le K_a ne peut pas être mesuré directement, une stratégie de déplacement doit être adoptée afin de rendre "visible" l'interaction (cela implique de disposer de deux ligands dont un est d'une affinité mesurable). Pour étudier la faisabilité d'un projet, le paramètre c (=

$nK_a[M]_T$) peut être approximé si l'on dispose d'un ordre de grandeur du K_a . Pour une affinité modérée, le paramètre c devrait se trouver entre 1 et 1000.

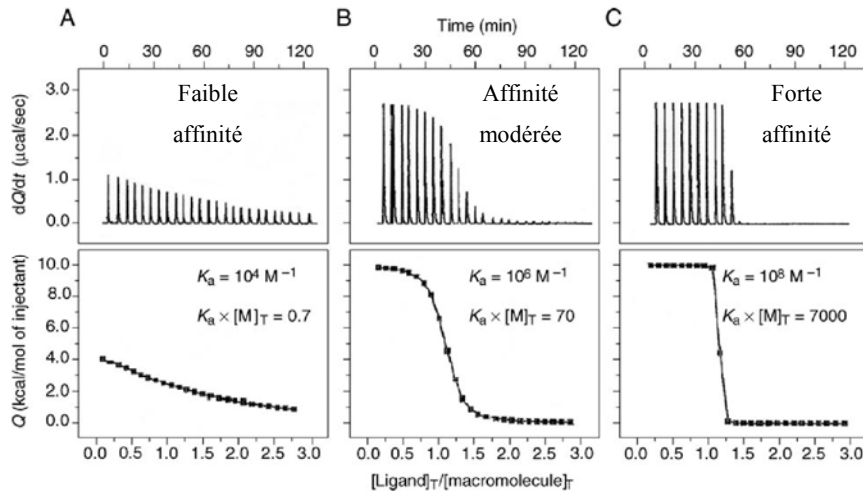


Figure 26 : Exemple de résultats en fonction de l'affinité

D'après Velázquez et coll. (116)

Ohnuma et coll. (119) ont déterminé par ITC la constante de dissociation de la chitinase de *Bryum coronatum* (GH-19) pour des oligosaccharides de chitine allant de 3 à 6 unités *N*-acetyl-D-glucosamine. Ils ont ainsi mis en évidence que la constante de dissociation diminue lorsque le nombre d'unités augmente.

Zolotnitsky et coll. (120) ont employé l'ITC pour cartographier le site actif de la xylanase de *Geobacillus stearothermophilus* (GH-10). Des analyses thermodynamiques ont été menées sur différents xylooligosaccharides allant de 2 à 6 unités. Couplées à l'analyse structurale de complexes cristallographiques entre la xylanase et un di ou trixylooligosaccharides, les données d'ITC ont montré l'importance des sites de stacking de trois sous-sites (+1, +2 et +3), chacun contribuant fortement à la variation de capacité calorifique à hauteur de -100 à -150 cal.mol⁻¹.K. Afin de valider ces résultats, des mutants sur un des résidus impliqués dans ces interactions d'empilement des sous-sites +1 et +2 ont été créés et ont montré une perte de capacité calorifique de -105 et -165 cal.mol⁻¹.K respectivement, confirmant ainsi l'importance de ces résidus responsable d'interactions d'empilement dans la fixation du ligand.

En ce qui concerne l'analyse d'interaction de type protéine-protéine, le schéma opérationnel reste le même que dans l'étude des interactions protéine-ligand, une des deux macromolécules jouant le rôle du ligand. Une variante intéressante est l'étude de l'homodimérisation des protéines. Dans ce cas, la protéine (sous-entendant une population monomère-dimère fixe) est placée dans la seringue, la cellule, quant à elle, est remplie avec du tampon de

conditionnement. L'injection de la protéine dans la cellule provoque un déplacement de l'équilibre monomère-dimère dans le sens de la monomérisation due à la dilution. Le thermogramme nous donne donc la quantité de chaleur dégagée ou absorbée lors de la monomérisation. Contreras et coll. (121) ont déterminé l'état d'oligomérisation de la β -xylosidase de *Geobacillus stearothermophilus* (GH-52) de manière indirecte en analysant les paramètres thermodynamiques de la fixation d'un inhibiteur compétitif (122) de cette enzyme : le DEO (1-deoxynojirimycine). Ils ont montré que la stœchiométrie de la fixation est d'une molécule de ligand pour deux protéines, suggérant la présence du site catalytique à l'interface de dimérisation. Ce résultat est corroboré par deux autres techniques biochimiques : la gel filtration et l'ultracentrifugation.

L'ITC peut être employée pour d'autres applications telles que la mesure des paramètres cinétiques d'une réaction enzymatique (123, 124) ou l'étude de la stabilité d'une protéine en présence d'un ligand. Ainsi Olsen et coll. (125) ont montré par ITC l'effet des ions calcium Ca^{2+} sur la stabilité de l' α -amylase de *Bacillus halmapalus*. L'ajout dans la cellule de mesure, après équilibration, d'EDTA (déplétant ainsi complètement le milieu en ion calcium) induit une forte réaction exothermique due à la complexation des ions calcium avec l'EDTA puis, dans un second temps, une réaction endothermique diminuant graduellement le signal montrant une lente dénaturation de la protéine en absence d'ion calcium. Cheleski et coll. (126) ont analysé par ITC les paramètres enzymatiques de la dihydroorotate dehydrogenase de *Trypanosoma cruzi*, étant donné que cette enzyme fonctionne avec deux substrats : le DHO (dihydroorotate) et le fumarate, les expériences de titrage ont été faites pour chaque substrat en présence d'un large excès du second substrat. Le flux de chaleur libéré lors de la réaction catalytique est proportionnel à la vitesse de réaction et est donnée par l'équation 20 :

Avec :

$$V = \frac{d[P]_t}{\Delta t} = \frac{1}{\Delta H_{\text{app}} V_0} \frac{dQ}{dt}$$

- $[P]_t$: concentration en produit
- V_0 : volume de la cellule
- dQ/dt : vitesse du flux de chaleur mesurée durant l'expérience
- ΔH_{app} : enthalpie molaire apparente, cette donnée dépend de la réaction et est déterminée à part par ITC

La concentration en substrat $[S]_t$ est donnée par l'expression :

$$[S]_t = [S]_0 - \frac{\int_0^t (dQ/dt)dt}{\Delta H_{app} V_0}$$

Avec :
➤ $[S]_0$: concentration initiale en substrat

Ensuite les données calorimétriques sont intégrées à l'équation de Michaelis et Menten et les paramètres cinétiques sont déduits de la courbe $1/V=f(1/[S]_t)$ avec un K_m de 8,56 μM pour le DHO et un K_m de 120 μM pour le fumarate.

L'ITC est donc une méthode performante possédant un large champ d'application permettant non seulement, de rentrer dans les détails thermodynamiques d'interactions protéine-ligand, mais aussi dans des interactions entre acides nucléiques, lipides, polysaccharides (127).

Conclusion

Les techniques biophysiques permettant de caractériser la structure et la stabilité des protéines ainsi que les interactions protéine-ligand sont nombreuses, chacune possédant ses avantages et inconvénients. Les études portant sur les relations entre structure et activité d'une enzyme vont généralement combiner plusieurs de ces approches complémentaires afin de décrire le plus fidèlement possible les paramètres physico-chimiques gouvernant l'activité et la stabilité des enzymes, mais aussi les interactions entre l'enzyme et son ligand (qu'il soit substrat, produit ou inhibiteur) ainsi que quantifier les forces d'interaction les unissant. L'influence de la dynamique est aussi un paramètre important à prendre en compte lors de l'analyse d'une fonction, les changements conformationnels pouvant être importants et essentiels à la catalyse enzymatique. De telles informations sont capitales à la compréhension des relations structure-dynamique-activité d'une enzyme notamment lorsqu'elles peuvent être recoupées avec des données biochimiques.

Objectifs de la Thèse

L'identification de nouvelles amylosaccharases par criblage de la diversité naturelle et l'ingénierie de l'amylosaccharase de *Neisseria polysaccharea* (ASNp) ont permis d'isoler des enzymes présentant des caractéristiques fonctionnelles ou physico-chimiques très distinctes de celles de l'ASNp sauvage. En particulier, l'amylosaccharase de *Deinococcus geothermalis* (ASDg), caractérisée au LISBP, est l'amylosaccharase la plus thermostable décrite dans la littérature. De plus, elle présente un profil de produits différent de celui obtenu avec l'ASNp. Jusqu'à nos travaux, aucune structure tridimensionnelle de l'ASDg n'était disponible afin d'identifier les facteurs structuraux potentiellement responsables de la stabilité mais aussi de la spécificité de cette enzyme. Nous avons également vu que l'ingénierie semi-rationnelle de l'ASNp a conduit à l'isolement de mutants ponctuels déployant des performances remarquables pour la reconnaissance d'accepteurs non naturels. Là encore, l'absence de données structurales et d'analyse biophysique de ces mutants ne permettait pas de proposer une analyse approfondie des facteurs responsables des modifications d'activité.

Pour palier ces déficiences et apporter des réponses aux questions soulevées par les particularités et les variations de comportement de ces différentes amylosaccharases, mes travaux de thèse réalisés entre l'équipe de Catalyse et Ingénierie Moléculaire Enzymatiques du LISBP (spécialiste des glucane-saccharases et de leur ingénierie) et l'équipe de Biophysique structurale de l'IPBS (spécialiste en analyse structurale de macromolécules biologiques) ont tout d'abord été focalisés sur la détermination de la structure tridimensionnelle de ces protéines. Nous avons centré nos expériences de cristallographie des rayons X sur l'ASDg ainsi que sur les simple- ou double-mutants de l'ASNp les plus performants pour la glycosylation d'accepteurs cibles. Différents essais de cristallisation ont été réalisés permettant la formation de cristaux sous formes *apo* (sans ligand) et/ou *holo* de ces protéines (en complexe avec différents substrats, produits ou tout autre analogue). Certains mono-mutants inactifs ou double-mutants ont été construits spécifiquement afin d'obtenir des complexes.

Pour affiner nos interprétations, nous avons réalisé en parallèle des caractérisations cinétiques complétées par des expériences de fluorimétrie différentielle à balayage, dichroïsme circulaire et microcalorimétrie. Enfin, des simulations par mécanique et dynamique moléculaires ont également été mises en œuvre pour explorer la dynamique de ces enzymes et apporter des éléments d'interprétation sur les liens entre structure, activité et dynamique de ces protéines.

Références Bibliographiques

1. 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**, D233–238
2. Potocki de Montalk, G., Remaud-Simeon, M., Willemot, R. M., Sarçabal, P., Planchot, V., and Monsan, P. (2000) Amylosucrase from *Neisseria polysaccharea*: novel catalytic properties. *FEBS Lett.* **471**, 219–223
3. Champion, E., André, I., Moulis, C., Boutet, J., Descroix, K., Morel, S., Monsan, P., Mulard, L. A., and Remaud-Siméon, M. (2009) Design of alpha-transglucosidases of controlled specificity for programmed chemoenzymatic synthesis of antigenic oligosaccharides. *J. Am. Chem. Soc.* **131**, 7379–7389
4. Hehre, E. J., and Hamilton, D. M. (1946) Bacterial synthesis of an amylopectin-like polysaccharide from sucrose. *J. Biol. Chem.* **166**, 777
5. Hehre, E. J., and Hamilton, D. M. (1948) The Conversion of Sucrose to a Polysaccharide of the Starch-Glycogen Class by *Neisseria* from the Pharynx. *J. Bacteriol.* **55**, 197–208
6. Riou, J. Y., Guibourdenche, M., and Popoff, M. Y. (1983) A new taxon in the genus *Neisseria*. *Ann. Microbiol. (Paris)* **134B**, 257–267
7. Büttcher, V., Welsh, T., Willmitzer, L., and Kossmann, J. (1997) Cloning and characterization of the gene for amylosucrase from *Neisseria polysaccharea*: production of a linear alpha-1,4-glucan. *J. Bacteriol.* **179**, 3324–3330
8. Riou, J. Y., Guibourdenche, M., Perry, M. B., MacLean, L. L., and Griffith, D. W. (1986) Structure of the exocellular D-glucan produced by *Neisseria polysaccharea*. *Can. J. Microbiol.* **32**, 909–911
9. Henrissat, B. (1991) A classification of glycosyl hydrolases based on amino acid sequence similarities. *Biochem. J.* **280 (Pt 2)**, 309–316
10. Henrissat, B., and Davies, G. (1997) Structural and sequence-based classification of glycoside hydrolases. *Curr. Opin. Struct. Biol.* **7**, 637–644
11. Henrissat, B., and Bairoch, A. (1993) New families in the classification of glycosyl hydrolases based on amino acid sequence similarities. *Biochem. J.* **293 (Pt 3)**, 781–788
12. Henrissat, B., and Bairoch, A. (1996) Updating the sequence-based classification of glycosyl hydrolases. *Biochem. J.* **316 (Pt 2)**, 695–696
13. Stam, M. R., Danchin, E. G. J., Rancurel, C., Coutinho, P. M., and Henrissat, B. (2006) Dividing the large glycoside hydrolase family 13 into subfamilies: towards improved functional annotations of alpha-amylase-related proteins. *Protein Eng. Des. Sel.* **19**, 555–562
14. Gabrisko, M., and Janecek, S. (2009) Looking for the ancestry of the heavy-chain subunits of heteromeric amino acid transporters rBAT and 4F2hc within the GH13 alpha-amylase family. *FEBS J.* **276**, 7265–7278
15. Vujcic-Zagar, A., Pijning, T., Kralj, S., López, C. A., Eeuwema, W., Dijkhuizen, L., and Dijkstra, B. W. (2010) Crystal structure of a 117 kDa glucansucrase fragment provides insight into evolution and product specificity of GH70 enzymes. *Proc. Natl. Acad. Sci. U.S.A.* **107**, 21406–21411
16. 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
17. Brison, Y., Pijning, T., Malbert, Y., Fabre, E., Mourey, L., Morel, S., Potocki-Véronèse, G., Monsan, P., Tranier, S., Remaud-Siméon, M., and Dijkstra, B. W. (2012) Functional and structural characterization of an α -(1->2) branching sucrose derived from DSR-E glucansucrase. *The Journal of Biological Chemistry*

18. Monsan, P., Remaud-Siméon, M., and André, I. (2010) Transglucosidases as efficient tools for oligosaccharide and glucoconjugate synthesis. *Curr. Opin. Microbiol.* **13**, 293–300
19. 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
20. Ficko-Blean, E., Stuart, C. P., and Boraston, A. B. (2011) Structural analysis of CPF_2247, a novel α -amylase from *Clostridium perfringens*. *Proteins* **79**, 2771–2777
21. Turkenburg, J. P., Brzozowski, A. M., Svendsen, A., Borchert, T. V., Davies, G. J., and Wilson, K. S. (2009) Structure of a pullulanase from *Bacillus acidopullulyticus*. *Proteins* **76**, 516–519
22. Gourlay, L. J., Santi, I., Pezzicoli, A., Grandi, G., Soriani, M., and Bolognesi, M. (2009) Group B streptococcus pullulanase crystal structures in the context of a novel strategy for vaccine development. *J. Bacteriol.* **191**, 3544–3552
23. Wind, R. D., Uitdehaag, J. C., Buitelaar, R. M., Dijkstra, B. W., and Dijkhuizen, L. (1998) Engineering of cyclodextrin product specificity and pH optima of the thermostable cyclodextrin glycosyltransferase from *Thermoanaerobacterium thermosulfurigenes* EM1. *J. Biol. Chem.* **273**, 5771–5779
24. Harata, K., Haga, K., Nakamura, A., Aoyagi, M., and Yamane, K. (1996) X-ray structure of cyclodextrin glucanotransferase from alkalophilic *Bacillus* sp. 1011. Comparison of two independent molecules at 1.8 Å resolution. *Acta Crystallogr. D Biol. Crystallogr.* **52**, 1136–1145
25. Fritzsche, H. B., Schwede, T., and Schulz, G. E. (2003) Covalent and three-dimensional structure of the cyclodextrinase from *Flavobacterium* sp. no. 92. *Eur. J. Biochem.* **270**, 2332–2341
26. Yamamoto, K., Miyake, H., Kusunoki, M., and Osaki, S. (2010) Crystal structures of isomaltase from *Saccharomyces cerevisiae* and in complex with its competitive inhibitor maltose. *FEBS J.* **277**, 4205–4214
27. Jung, T.-Y., Li, D., Park, J.-T., Yoon, S.-M., Tran, P. L., Oh, B.-H., Janecek, S., Park, S. G., Woo, E.-J., and Park, K.-H. (2012) Association of a novel domain in the active site of archaic hyperthermophile maltogenic amylase from *Staphylothermus marinus*. *The Journal of Biological Chemistry*
28. Hondoh, H., Kuriki, T., and Matsuura, Y. (2003) Three-dimensional structure and substrate binding of *Bacillus stearothermophilus* neopullulanase. *J. Mol. Biol.* **326**, 177–188
29. Shirai, T., Hung, V. S., Morinaka, K., Kobayashi, T., and Ito, S. (2008) Crystal structure of GH13 alpha-glucosidase GSJ from one of the deepest sea bacteria. *Proteins* **73**, 126–133
30. Morishita, Y., Hasegawa, K., Matsuura, Y., Katsube, Y., Kubota, M., and Sakai, S. (1997) Crystal structure of a maltotetraose-forming exo-amylase from *Pseudomonas stutzeri*. *J. Mol. Biol.* **267**, 661–672
31. Song, H.-N., Jung, T.-Y., Park, J.-T., Park, B.-C., Myung, P. K., Boos, W., Woo, E.-J., and Park, K.-H. (2010) Structural rationale for the short branched substrate specificity of the glycogen debranching enzyme GlgX. *Proteins* **78**, 1847–1855
32. Katsuya, Y., Mezaki, Y., Kubota, M., and Matsuura, Y. (1998) Three-dimensional structure of *Pseudomonas* isoamylase at 2.2 Å resolution. *J. Mol. Biol.* **281**, 885–897
33. Hondoh, H., Saburi, W., Mori, H., Okuyama, M., Nakada, T., Matsuura, Y., and Kimura, A. (2008) Substrate recognition mechanism of alpha-1,6-glucosidic linkage hydrolyzing enzyme, dextran glucosidase from *Streptococcus mutans*. *J. Mol. Biol.* **378**, 913–922

34. Kanai, R., Haga, K., Akiba, T., Yamane, K., and Harata, K. (2004) Biochemical and crystallographic analyses of maltohexaose-producing amylase from alkalophilic *Bacillus* sp. 707. *Biochemistry* **43**, 14047–14056
35. Noguchi, J., Chaen, K., Vu, N. T., Akasaka, T., Shimada, H., Nakashima, T., Nishi, A., Satoh, H., Omori, T., Kakuta, Y., and Kimura, M. (2011) Crystal structure of the branching enzyme I (BEI) from *Oryza sativa* L with implications for catalysis and substrate binding. *Glycobiology* **21**, 1108–1116
36. Pal, K., Kumar, S., Sharma, S., Garg, S. K., Alam, M. S., Xu, H. E., Agrawal, P., and Swaminathan, K. (2010) Crystal structure of full-length *Mycobacterium tuberculosis* H37Rv glycogen branching enzyme: insights of N-terminal beta-sandwich in substrate specificity and enzymatic activity. *J. Biol. Chem.* **285**, 20897–20903
37. Kobayashi M., M, K., and Y, M. (2003) Refined Structure and Functional Implications of Trehalose Synthase from *Sulfolobus Acidocaldarius*. *J Appl Glycosci* **50**, 1–8
38. Roujeinikova, A., Raasch, C., Sedelnikova, S., Liebl, W., and Rice, D. W. (2002) Crystal structure of *Thermotoga maritima* 4- α -glucanotransferase and its acarbose complex: implications for substrate specificity and catalysis. *J. Mol. Biol.* **321**, 149–162
39. 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 alpha-amylase family. *J. Biol. Chem.* **276**, 25273–25278
40. Sprogøe, D., van den Broek, L. A. M., Mirza, O., Kastrup, J. S., Voragen, A. G. J., Gajhede, M., and Skov, L. K. (2004) Crystal structure of sucrose phosphorylase from *Bifidobacterium adolescentis*. *Biochemistry* **43**, 1156–1162
41. Timmins, J., Leiros, H.-K. S., Leonard, G., Leiros, I., and McSweeney, S. (2005) Crystal structure of maltooligosyltrehalose trehalohydrolase from *Deinococcus radiodurans* in complex with disaccharides. *J. Mol. Biol.* **347**, 949–963
42. Zhang, D., Li, N., Lok, S.-M., Zhang, L.-H., and Swaminathan, K. (2003) Isomaltulose synthase (Pall) of *Klebsiella* sp. LX3. Crystal structure and implication of mechanism. *J. Biol. Chem.* **278**, 35428–35434
43. Kim, M.-I., Kim, H.-S., Jung, J., and Rhee, S. (2008) Crystal structures and mutagenesis of sucrose hydrolase from *Xanthomonas axonopodis* pv. glycines: insight into the exclusively hydrolytic amylosucrase fold. *J. Mol. Biol.* **380**, 636–647
44. Champion, E., Remaud-Simeon, M., Skov, L. K., Kastrup, J. S., Gajhede, M., and Mirza, O. (2009) The apo structure of sucrose hydrolase from *Xanthomonas campestris* pv. *campestris* shows an open active-site groove. *Acta Crystallogr. D Biol. Crystallogr.* **65**, 1309–1314
45. Mótýán, J. A., Gyémánt, G., Harangi, J., and Bagossi, P. (2011) Computer-aided subsite mapping of α -amylases. *Carbohydr. Res.* **346**, 410–415
46. Sukanuma, T., Matsuno, R., Ohnishi, M., and Hiromi, K. (1978) A study of the mechanism of action of Taka-amylase A1 on linear oligosaccharides by product analysis and computer simulation. *J. Biochem.* **84**, 293–316
47. 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
48. Kumar, V. (2010) Analysis of the key active subsites of glycoside hydrolase 13 family members. *Carbohydr. Res.* **345**, 893–898
49. Skov, L. K., Mirza, O., Henriksen, A., Potocki de Montalk, G., Remaud-Simeon, M., Sarcabal, P., Willemot, R. M., Monsan, P., and Gajhede, M. (2000) Crystallization and preliminary X-ray studies of recombinant amylosucrase from *Neisseria polysaccharea*. *Acta Crystallogr. D Biol. Crystallogr.* **56**, 203–205

50. Jespersen, H. M., MacGregor, E. A., Henrissat, B., Sierks, M. R., and Svensson, B. (1993) Starch- and glycogen-debranching and branching enzymes: prediction of structural features of the catalytic (beta/alpha)₈-barrel domain and evolutionary relationship to other amylolytic enzymes. *J. Protein Chem.* **12**, 791–805
51. Sarçabal, P., Remaud-Simeon, M., Willemot, R., Potocki de Montalk, G., Svensson, B., and Monsan, P. (2000) Identification of key amino acid residues in *Neisseria polysaccharea* amylosucrase. *FEBS Lett.* **474**, 33–37
52. Kumar, V. (2011) Identification of the sequence motif of glycoside hydrolase 13 family members. *Bioinformatics* **6**, 61–63
53. Crooks, G. E., Hon, G., Chandonia, J.-M., and Brenner, S. E. (2004) WebLogo: a sequence logo generator. *Genome Res.* **14**, 1188–1190
54. Janecek, S., Svensson, B., and Henrissat, B. (1997) Domain evolution in the alpha-amylase family. *J. Mol. Evol.* **45**, 322–331
55. Robert, X., Haser, R., Gottschalk, T. E., Ratajczak, F., Driguez, H., Svensson, B., and Aghajari, N. (2003) The structure of barley alpha-amylase isozyme 1 reveals a novel role of domain C in substrate recognition and binding: a pair of sugar tongs. *Structure* **11**, 973–984
56. Bozonnet, S., Jensen, M. T., Nielsen, M. M., Aghajari, N., Jensen, M. H., Kramhøft, B., Willemoës, M., Tranier, S., Haser, R., and Svensson, B. (2007) The “pair of sugar tongs” site on the non-catalytic domain C of barley alpha-amylase participates in substrate binding and activity. *FEBS J.* **274**, 5055–5067
57. MacGregor, E. A., Janecek, S., and Svensson, B. (2001) Relationship of sequence and structure to specificity in the alpha-amylase family of enzymes. *Biochim. Biophys. Acta* **1546**, 1–20
58. Davies, G., and Henrissat, B. (1995) Structures and mechanisms of glycosyl hydrolases. *Structure* **3**, 853–859
59. Withers, S. (2001) Mechanisms of glycosyl transferases and hydrolases. *Carbohydrate Polymers* **44**, 325–337
60. Pizzut-Serin, S., Potocki-Véronèse, G., van der Veen, B. A., Albenne, C., Monsan, P., and Remaud-Simeon, M. (2005) Characterisation of a novel amylosucrase from *Deinococcus radiodurans*. *FEBS Lett.* **579**, 1405–1410
61. Emond, S., Mondeil, S., Jaziri, K., André, I., Monsan, P., Remaud-Siméon, M., and Potocki-Véronèse, G. (2008) Cloning, purification and characterization of a thermostable amylosucrase from *Deinococcus geothermalis*. *FEMS Microbiol. Lett.* **285**, 25–32
62. Ha, S.-J., Seo, D.-H., Jung, J.-H., Cha, J., Kim, T.-J., Kim, Y.-W., and Park, C.-S. (2009) Molecular cloning and functional expression of a new amylosucrase from *Alteromonas macleodii*. *Biosci. Biotechnol. Biochem.* **73**, 1505–1512
63. De Montalk, G. P., Remaud-Simeon, M., Willemot, R. M., Planchot, V., and Monsan, P. (1999) Sequence analysis of the gene encoding amylosucrase from *Neisseria polysaccharea* and characterization of the recombinant enzyme. *J. Bacteriol.* **181**, 375–381
64. Jung, J.-H., Seo, D.-H., Ha, S.-J., Song, M.-C., Cha, J., Yoo, S.-H., Kim, T.-J., Baik, N.-I., Baik, M.-Y., and Park, C.-S. (2009) Enzymatic synthesis of salicin glycosides through transglycosylation catalyzed by amylosucrases from *Deinococcus geothermalis* and *Neisseria polysaccharea*. *Carbohydr. Res.* **344**, 1612–1619
65. Schneider, J., Fricke, C., Overwin, H., and Hofer, B. (2011) High level expression of a recombinant amylosucrase gene and selected properties of the enzyme. *Appl. Microbiol. Biotechnol.* **89**, 1821–1829

66. Seo, D.-H., Jung, J.-H., Ha, S.-J., Cho, H.-K., Jung, D.-H., Kim, T.-J., Baek, N.-I., Yoo, S.-H., and Park, C.-S. (2012) High-yield enzymatic bioconversion of hydroquinone to α -arbutin, a powerful skin lightening agent, by amylosucrase. *Applied Microbiology and Biotechnology*
67. Ferreira, A. C., Nobre, M. F., Rainey, F. A., Silva, M. T., Wait, R., Burghardt, J., Chung, A. P., and da Costa, M. S. (1997) *Deinococcus geothermalis* sp. nov. and *Deinococcus murrayi* sp. nov., two extremely radiation-resistant and slightly thermophilic species from hot springs. *Int. J. Syst. Bacteriol.* **47**, 939–947
68. Lichtenthaler, F. W., and Rönninger, S. (1990) α -D-Glucopyranosyl-D-fructoses: distribution of furanoid and pyranoid tautomers in water, dimethyl sulphoxide, and pyridine. Studies on ketoses. Part 4. *J. Chem. Soc., Perkin Trans. 2*, 1489–1497
69. Potocki de Montalk, G., Remaud-Simeon, M., Willemot, R. M., and Monsan, P. (2000) Characterisation of the activator effect of glycogen on amylosucrase from *Neisseria polysaccharea*. *FEMS Microbiol. Lett.* **186**, 103–108
70. Potocki-Veronese, G., Putaux, J.-L., Dupeyre, D., Albenne, C., Remaud-Siméon, M., Monsan, P., and Buleon, A. (2005) Amylose synthesized in vitro by amylosucrase: morphology, structure, and properties. *Biomacromolecules* **6**, 1000–1011
71. Putaux, J.-L., Potocki-Véronèse, G., Remaud-Simeon, M., and Buleon, A. (2006) Alpha-D-glucan-based dendritic nanoparticles prepared by in vitro enzymatic chain extension of glycogen. *Biomacromolecules* **7**, 1720–1728
72. Rolland-Sabaté, A., Colonna, P., Potocki-Véronèse, G., Monsan, P., and Planchot, V. (2004) Elongation and insolubilisation of α -glucans by the action of *Neisseria polysaccharea* amylosucrase. *Journal of Cereal Science* **40**, 17–30
73. Topping, D. L., Fukushima, M., and Bird, A. R. (2003) Resistant starch as a prebiotic and synbiotic: state of the art. *Proc Nutr Soc* **62**, 171–176
74. Thompson, J., and Pikis, A. Metabolism of sugars by genetically diverse species of oral *Leptotrichia*. *Molecular Oral Microbiology*
75. Kim, J.-H., Wang, R., Lee, W.-H., Park, C.-S., Lee, S., and Yoo, S.-H. (2011) One-pot synthesis of cycloamyloses from sucrose by dual enzyme treatment: combined reaction of amylosucrase and 4- α -glucanotransferase. *J. Agric. Food Chem.* **59**, 5044–5051
76. Cho, H.-K., Kim, H.-H., Seo, D.-H., Jung, J.-H., Park, J.-H., Baek, N.-I., Kim, M.-J., Yoo, S.-H., Cha, J., Kim, Y.-R., and Park, C.-S. (2011) Biosynthesis of (+)-catechin glycosides using recombinant amylosucrase from *Deinococcus geothermalis* DSM 11300. *Enzyme Microb. Technol.* **49**, 246–253
77. Seo, D.-H. (2009) Highly selective biotransformation of arbutin to arbutin- α -glucoside using amylosucrase from *Deinococcus geothermalis* DSM 11300. *Journal of Molecular Catalysis. B, Enzymatic* **60**, 113
78. Schneider, J., Fricke, C., Overwin, H., Hofmann, B., and Hofer, B. (2009) Generation of amylosucrase variants that terminate catalysis of acceptor elongation at the di- or trisaccharide stage. *Appl. Environ. Microbiol.* **75**, 7453–7460
79. Wang, R., Kim, J.-H., Kim, B.-S., Park, C.-S., and Yoo, S.-H. (2011) Preparation and characterization of non-covalently immobilized amylosucrase using a pH-dependent autoprecipitating carrier. *Bioresour. Technol.* **102**, 6370–6374
80. Mirza, O., Skov, L. K., Remaud-Simeon, M., Potocki de Montalk, G., Albenne, C., Monsan, P., and Gajhede, M. (2001) Crystal structures of amylosucrase from *Neisseria polysaccharea* in complex with D-glucose and the active site mutant Glu328Gln in complex with the natural substrate sucrose. *Biochemistry* **40**, 9032–9039
81. Skov, L. K., Mirza, O., Sprogøe, D., Dar, I., Remaud-Simeon, M., Albenne, C., Monsan, P., and Gajhede, M. (2002) Oligosaccharide and sucrose complexes of

- amylosucrase. Structural implications for the polymerase activity. *J. Biol. Chem.* **277**, 47741–47747
82. Skov, L. K., Mirza, O., Sprogøe, D., Van Der Veen, B. A., Remaud-Simeon, M., Albenne, C., Monsan, P., and Gajhede, M. Crystal structure of the Glu328Gln mutant of *Neisseria polysaccharea* amylosucrase in complex with sucrose and maltoheptaose. *Biocatalysis and biotransformation* **24**, 99–105
83. Jensen, M. H., Mirza, O., Albenne, C., Remaud-Simeon, M., Monsan, P., Gajhede, M., and Skov, L. K. (2004) Crystal structure of the covalent intermediate of amylosucrase from *Neisseria polysaccharea*. *Biochemistry* **43**, 3104–3110
84. *The PyMOL Molecular Graphics System, Version 1.5.0.4 Schrödinger, LLC.*
85. Wallace, A. C., Laskowski, R. A., and Thornton, J. M. (1995) LIGPLOT: a program to generate schematic diagrams of protein-ligand interactions. *Protein Eng.* **8**, 127–134
86. Derewenda, U., Swenson, L., Green, R., Wei, Y., Morosoli, R., Shareck, F., Kluepfel, D., and Derewenda, Z. S. (1994) Crystal structure, at 2.6-Å resolution, of the *Streptomyces lividans* xylanase A, a member of the F family of beta-1,4-D-glycanases. *J. Biol. Chem.* **269**, 20811–20814
87. Ducros, V., Charnock, S. J., Derewenda, U., Derewenda, Z. S., Dauter, Z., Dupont, C., Shareck, F., Morosoli, R., Kluepfel, D., and Davies, G. J. (2000) Substrate specificity in glycoside hydrolase family 10. Structural and kinetic analysis of the *Streptomyces lividans* xylanase 10A. *J. Biol. Chem.* **275**, 23020–23026
88. Albenne, C., Skov, L. K., Mirza, O., Gajhede, M., Feller, G., D'Amico, S., André, G., Potocki-Véronèse, G., van der Veen, B. A., Monsan, P., and Remaud-Simeon, M. (2004) Molecular basis of the amylose-like polymer formation catalyzed by *Neisseria polysaccharea* amylosucrase. *J. Biol. Chem.* **279**, 726–734
89. Cunningham, B. C., and Wells, J. A. (1989) High-resolution epitope mapping of hGH-receptor interactions by alanine-scanning mutagenesis. *Science* **244**, 1081–1085
90. 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
91. Albenne, C., Skov, L. K., Mirza, O., Gajhede, M., Potocki-Véronèse, G., Monsan, P., and Remaud-Simeon, M. (2002) Maltooligosaccharide disproportionation reaction: an intrinsic property of amylosucrase from *Neisseria polysaccharea*. *FEBS Lett.* **527**, 67–70
92. Wong, T. S., Roccatano, D., and Schwaneberg, U. (2007) Steering directed protein evolution: strategies to manage combinatorial complexity of mutant libraries. *Environ. Microbiol.* **9**, 2645–2659
93. Kang, S., and Saven, J. G. (2007) Computational protein design: structure, function and combinatorial diversity. *Curr Opin Chem Biol* **11**, 329–334
94. van der Veen, B. A., Potocki-Véronèse, G., Albenne, C., Joucla, G., Monsan, P., and Remaud-Simeon, M. (2004) Combinatorial engineering to enhance amylosucrase performance: construction, selection, and screening of variant libraries for increased activity. *FEBS Lett.* **560**, 91–97
95. Emond, S., Mondon, P., Pizzut-Serin, S., Douchy, L., Crozet, F., Bouayadi, K., Kharrat, H., Potocki-Véronèse, G., Monsan, P., and Remaud-Simeon, M. (2008) A novel random mutagenesis approach using human mutagenic DNA polymerases to generate enzyme variant libraries. *Protein Eng. Des. Sel.* **21**, 267–274
96. Emond, S., André, I., Jaziri, K., Potocki-Véronèse, G., Mondon, P., Bouayadi, K., Kharrat, H., Monsan, P., and Remaud-Simeon, M. (2008) Combinatorial engineering to enhance thermostability of amylosucrase. *Protein Sci.* **17**, 967–976

97. Lutz, S. (2010) Beyond directed evolution--semi-rational protein engineering and design. *Curr. Opin. Biotechnol.* **21**, 734–743
98. Kotloff, K. L., Winickoff, J. P., Ivanoff, B., Clemens, J. D., Swerdlow, D. L., Sansonetti, P. J., Adak, G. K., and Levine, M. M. (1999) Global burden of *Shigella* infections: implications for vaccine development and implementation of control strategies. *Bull. World Health Organ.* **77**, 651–666
99. Kelly, S. M., Jess, T. J., and Price, N. C. (2005) How to study proteins by circular dichroism. *Biochim. Biophys. Acta* **1751**, 119–139
100. Gatti-Lafranconi, P., Natalello, A., Rehm, S., Doglia, S. M., Pleiss, J., and Lotti, M. (2010) Evolution of stability in a cold-active enzyme elicits specificity relaxation and highlights substrate-related effects on temperature adaptation. *J. Mol. Biol.* **395**, 155–166
101. Chi, M.-C., Wu, T.-J., Chuang, T.-T., Chen, H.-L., Lo, H.-F., and Lin, L.-L. (2010) Biophysical characterization of a recombinant α -amylase from thermophilic *Bacillus* sp. strain TS-23. *Protein J.* **29**, 572–582
102. Shizawa, N., Uchiumi, T., Taguchi, J., Kisseleva, N. A., Baykov, A. A., Lahti, R., and Hachimori, A. (2001) Directed mutagenesis studies of the C-terminal fingerprint region of *Bacillus subtilis* pyrophosphatase. *Eur. J. Biochem.* **268**, 5771–5775
103. Niesen, F. H., Berglund, H., and Vedadi, M. (2007) The use of differential scanning fluorimetry to detect ligand interactions that promote protein stability. *Nat Protoc* **2**, 2212–2221
104. Ericsson, U. B., Hallberg, B. M., Detitta, G. T., Dekker, N., and Nordlund, P. (2006) Thermofluor-based high-throughput stability optimization of proteins for structural studies. *Anal. Biochem.* **357**, 289–298
105. Lo, M.-C., Aulabaugh, A., Jin, G., Cowling, R., Bard, J., Malamas, M., and Ellestad, G. (2004) Evaluation of fluorescence-based thermal shift assays for hit identification in drug discovery. *Anal. Biochem.* **332**, 153–159
106. Pantoliano, M. W., Petrella, E. C., Kwasnoski, J. D., Lobanov, V. S., Myslik, J., Graf, E., Carver, T., Asel, E., Springer, B. A., Lane, P., and Salemme, F. R. (2001) High-density miniaturized thermal shift assays as a general strategy for drug discovery. *J Biomol Screen* **6**, 429–440
107. Hopper, E. D., Pittman, A. M. C., Fitzgerald, M. C., and Tucker, C. L. (2008) *In vivo* and *in vitro* examination of stability of primary hyperoxaluria-associated human alanine:glyoxylate aminotransferase. *J. Biol. Chem.* **283**, 30493–30502
108. Senisterra, G. A., Soo Hong, B., Park, H.-W., and Vedadi, M. (2008) Application of high-throughput isothermal denaturation to assess protein stability and screen for ligands. *J Biomol Screen* **13**, 337–342
109. Froese, D. S., Healy, S., McDonald, M., Kochan, G., Oppermann, U., Niesen, F. H., and Gravel, R. A. (2010) Thermolability of mutant MMACHC protein in the vitamin B12-responsive cblC disorder. *Mol. Genet. Metab.* **100**, 29–36
110. McPherson, A. (2004) Introduction to protein crystallization. *Methods* **34**, 254–265
111. Chayen, N. E., and Saridakis, E. (2008) Protein crystallization: from purified protein to diffraction-quality crystal. *Nat. Methods* **5**, 147–153
112. Hassell, A. M., An, G., Bledsoe, R. K., Bynum, J. M., Carter, H. L., 3rd, Deng, S.-J. J., Gampe, R. T., Grisard, T. E., Madauss, K. P., Nolte, R. T., Rocque, W. J., Wang, L., Weaver, K. L., Williams, S. P., Wisely, G. B., Xu, R., and Shewchuk, L. M. (2007) Crystallization of protein-ligand complexes. *Acta Crystallogr. D Biol. Crystallogr.* **63**, 72–79
113. Verlet, L. (1967) Computer “Experiments” on Classical Fluids. I. Thermodynamical Properties of Lennard-Jones Molecules. *Phys. Rev.* **159**, 98–103

114. Omelyan, I. P., Mryglod, I. M., and Folk, R. (2002) Optimized Verlet-like algorithms for molecular dynamics simulations. *Phys Rev E Stat Nonlin Soft Matter Phys* **65**, 056706
115. Henzler-Wildman, K., and Kern, D. (2007) Dynamic personalities of proteins. *Nature* **450**, 964–972
116. Velázquez-Campoy, A., Ohtaka, H., Nezami, A., Muzammil, S., and Freire, E. (2004) Isothermal titration calorimetry. *Curr Protoc Cell Biol* **Chapter 17**, Unit 17.8
117. Indyk, L., and Fisher, H. F. (1998) Theoretical aspects of isothermal titration calorimetry. *Meth. Enzymol.* **295**, 350–364
118. Liang, Y. (2008) Applications of isothermal titration calorimetry in protein science. *Acta Biochim. Biophys. Sin. (Shanghai)* **40**, 565–576
119. Ohnuma, T., Sørлие, M., Fukuda, T., Kawamoto, N., Taira, T., and Fukamizo, T. (2011) Chitin oligosaccharide binding to a family GH19 chitinase from the moss *Bryum coronatum*. *FEBS J.* **278**, 3991–4001
120. Zolotnitsky, G., Cogan, U., Adir, N., Solomon, V., Shoham, G., and Shoham, Y. (2004) Mapping glycoside hydrolase substrate subsites by isothermal titration calorimetry. *Proc. Natl. Acad. Sci. U.S.A.* **101**, 11275–11280
121. Contreras, L. M., Gómez, J., Prieto, J., Clemente-Jiménez, J. M., Las Heras-Vázquez, F. J., Rodríguez-Vico, F., Blanco, F. J., and Neira, J. L. (2008) The family 52 beta-xylosidase from *Geobacillus stearothermophilus* is a dimer: structural and biophysical characterization of a glycoside hydrolase. *Biochim. Biophys. Acta* **1784**, 1924–1934
122. Bravman, T., Zolotnitsky, G., Belakhov, V., Shoham, G., Henrissat, B., Baasov, T., and Shoham, Y. (2003) Detailed kinetic analysis of a family 52 glycoside hydrolase: a beta-xylosidase from *Geobacillus stearothermophilus*. *Biochemistry* **42**, 10528–10536
123. Todd, M. J., and Gomez, J. (2001) Enzyme kinetics determined using calorimetry: a general assay for enzyme activity? *Anal. Biochem.* **296**, 179–187
124. Wiggers, H. J., Cheleski, J., Zottis, A., Oliva, G., Andricopulo, A. D., and Montanari, C. A. (2007) Effects of organic solvents on the enzyme activity of *Trypanosoma cruzi* glyceraldehyde-3-phosphate dehydrogenase in calorimetric assays. *Anal. Biochem.* **370**, 107–114
125. Olsen, S. N., Andersen, K. B., Øgøndal, L. H., and Westh, P. (2009) A calorimetric study of solute effects on the kinetic stability of α -amylase. *Thermochimica Acta* **484**, 32–37
126. Cheleski, J., Wiggers, H. J., Citadini, A. P., da Costa Filho, A. J., Nonato, M. C., and Montanari, C. A. (2010) Kinetic mechanism and catalysis of *Trypanosoma cruzi* dihydroorotate dehydrogenase enzyme evaluated by isothermal titration calorimetry. *Anal. Biochem.* **399**, 13–22
127. Falconer, R. J., and Collins, B. M. (2011) Survey of the year 2009: applications of isothermal titration calorimetry. *J. Mol. Recognit.* **24**, 1–16

*Chapitre 2 : Etude de la
Thermostabilité et de la Spécificité
de Produits de l'Amylosaccharase
de Deinococcus geothermalis*

Les travaux présentés dans ce chapitre ont porté sur l'amylosaccharase de *Deinococcus geothermalis* (ASDg), une amylosaccharase caractérisée dans l'équipe de Catalyse et Ingénierie Moléculaire Enzymatiques du LISBP de l'INSA de Toulouse, montrant de nombreuses différences biochimiques et enzymatiques en comparaison à l'amylosaccharase de *Neisseria polysaccharea* (ASNp). En effet, les profils de produits synthétisés par les deux enzymes sont différents. En particulier, elles conduisent à la formation de quantités distinctes d'isomères du saccharose. Les deux enzymes se distinguent aussi par leur stabilité, l'ASDg possédant une stabilité à la température plus forte que celle de l'ASNp dont les origines moléculaires n'ont toujours pas été expliquées.

Ce chapitre propose une étude structurale de l'ASDg afin de comprendre l'origine des différences observées avec l'ASNp (code PDB : 1G5A). Nous avons, dans un premier temps, résolu la structure cristallographique de l'ASDg à 1,97 Å de résolution, déposée à la Protein Data Bank sous le code 3UCQ. L'analyse de la structure a révélé un certain nombre de particularités, dont un domaine N radicalement différent de celui de l'ASNp, ou encore, des insertions au sein de boucles (comme, par exemple, la boucle reliant l'hélice $\alpha 5$ et le brin $\beta 6$ du tonneau catalytique) rendant possibles des contacts intermoléculaires entre deux protomères d'ASDg qui ne peuvent se produire pour l'ASNp. De plus, une étude par dynamique moléculaire de ces deux enzymes a révélé un comportement singulièrement différent des deux protéines, notamment au niveau du domaine N. Enfin, l'étude en solution par SEC-MALLS a montré pour l'ASDg ainsi que pour l'ASDr (amylosaccharase de *Deinococcus radiodurans*) un état dimérique en solution se distinguant de l'état monomérique observé pour l'ASNp.

Pour comprendre les différences observées au niveau des profils de produits de l'ASDg et de l'ASNp, nous avons entrepris d'étudier l'activité d'isomérisation du saccharose en turanose (α -D-glucopyranosyl-1,3-D-fructose) et en tréhalulose (α -D-glucopyranosyl-1,1-D-fructose). Cette réaction étant minoritaire, elle a été forcée en rajoutant du fructose avec le saccharose en condition initiale afin d'amplifier les différences sur les profils de produits obtenus pour les deux enzymes. Les résultats ont montré une synthèse équimolaire des deux isomères pour l'ASDg, alors que l'ASNp synthétise préférentiellement du turanose. Afin d'aller plus loin dans la compréhension de ces résultats, nous avons déterminé la structure cristallographique des complexes de ces deux protéines avec le turanose. Le complexe ASDg-turanose a été résolu à 2,10 Å de résolution et déposé à la Protein Data Bank sous le code 3UER, tandis que

l'ASNp-turanose a été résolu à 1,85 Å de résolution et déposé sous le code 3UEQ. Au sein du complexe ASDg-turanose, au niveau du site actif, le résidu fructosyle du turanose se trouve sous forme furanose avec un mélange des anomères α et β tandis que dans le complexe ASNp-turanose, ce même résidu se trouve sous forme linéaire, traduisant un mode de fixation du fructose dans le site actif différent pour les deux enzymes.

Grâce aux informations structurales collectées, cette étude offre de nouvelles perspectives pour moduler l'activité enzymatique et/ou altérer le profil de produits par ingénierie rationnelle ou semi-rationnelle de l'amylosaccharase thermostable de *Deinococcus geothermalis*.

Ces travaux ont fait l'objet d'une publication à paraître dans le volume 287 de *The Journal of Biological Chemistry* p.6642-6654.

STRUCTURAL INVESTIGATION OF THE THERMOSTABILITY AND PRODUCT SPECIFICITY OF AMYLOSUCRASE FROM THE BACTERIUM *DEINOCOCCUS GEOTHERMALIS*

Frédéric Guérin, Sophie Barbe, Sandra Pizzut-Serin, Gabrielle Potocki-Véronèse, David Guieysse, Valérie Guillet, Pierre Monsan, Lionel Mourey, Magali Remaud-Siméon, Isabelle André and Samuel Tranier

Summary

Amylosucrases are sucrose-utilizing α -transglucosidases that naturally catalyze the synthesis of α -glucans, linked exclusively through α -1,4-linkages. Side-products and in particular sucrose isomers such as turanose and trehalulose are also produced by these enzymes. Here, we report the first structural and biophysical characterization of the most thermostable amylosucrase identified so far, the amylosucrase from *Deinococcus geothermalis* (DgAS). The 3D-structure revealed a homodimeric quaternary organization, never reported before for other amylosucrases. A sequence signature of dimerization was identified from the analysis of the dimer interface and sequence alignments. By rigidifying DgAS structure, the quaternary organization is likely to participate in the enhanced thermal stability of the protein. Amylosucrase specificity with respect to sucrose isomer formation (turanose or trehalulose) was also investigated. We report the first structures of the amylosucrases from *Deinococcus geothermalis* and *Neisseria polysaccharea* in complex with turanose. In the amylosucrase from *N. polysaccharea* (NpAS), key residues were found to force the fructosyl moiety to bind in an open state with the O3' ideally positioned to explain the preferential formation of turanose by NpAS. Such residues are either not present or not similarly placed in DgAS. As a consequence, DgAS binds the furanoid tautomers of fructose through a weak network of interactions to enable turanose formation. Such topology at subsite +1 is likely favoring other possible fructose binding modes in agreement with the higher amount of trehalulose formed by DgAS. Our findings help to understand the inter-relationships between amylosucrase structure, flexibility, function and stability and provide new insight for amylosucrase design.

Introduction

Amylosucrases (AS, E.C. number 2.4.1.4) are glucansucrases belonging to Glycoside-Hydrolase Family 13 of the Carbohydrate-Active EnZymes classification (1-3). The main reaction catalyzed by these enzymes is the formation of an amylose-like glucan with a concomitant release of fructose from sucrose substrate (Fig. 27) (4). Side-reactions including sucrose hydrolysis and sucrose isomer synthesis (i.e. turanose: α -D-Glcp(1 \rightarrow 3)- β -D-Fru and trehalulose: α -D-Glcp(1 \rightarrow 1)- β -D-Fru) also occur (4,5) although other GH13 enzymes are known to be more specific for these types of reaction (6-8). These molecules are known to be less cariogenic than sucrose and are thus of interest for nutritional applications (9). AS are also able to transfer the glucosyl moieties from sucrose to exogenous acceptors such as maltose, glycogen, maltodextrins, arbutin and others (10,11). AS ability to convert sucrose (an abundant and inexpensive substrate) into valuable derivatives is a great advantage in comparison to Leloir glucosyltransferases which use nucleotide-activated sugars as glucosyl donor substrates (12). Moreover, protein engineering techniques also enabled the creation of novel AS with tailored specificity toward unnatural acceptor molecules (13).

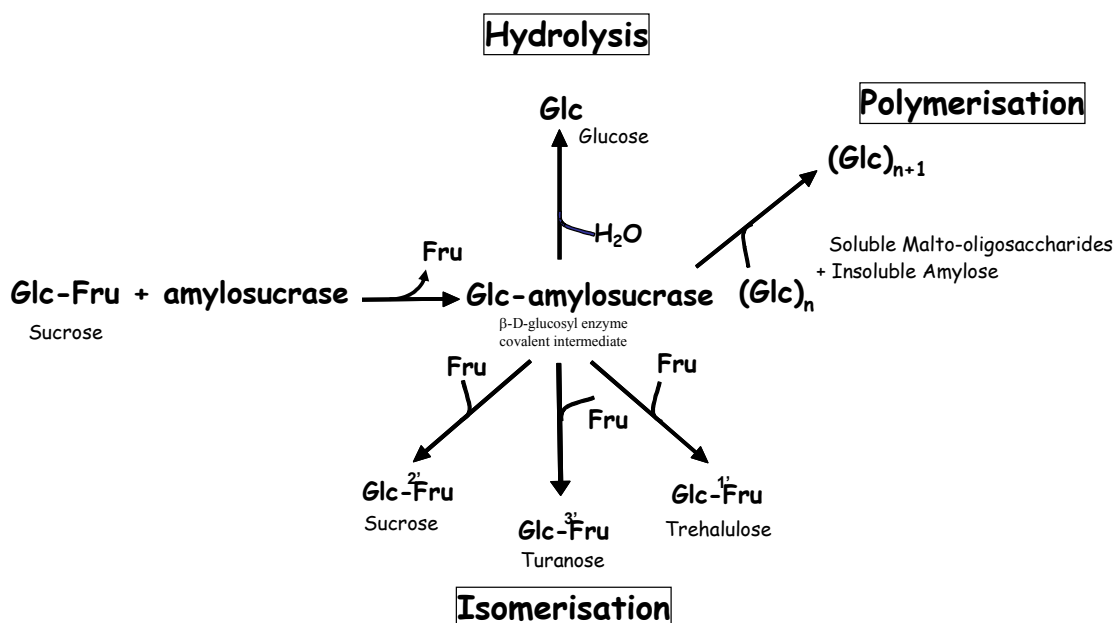


Figure 27: Reactions catalyzed by amylosucrases from sole sucrose

Glc: Glucose; Fru: Fructose; Sucrose: α -D-glucopyranosyl-1,2- β -D-fructofuranoside; Turanose: α -D-glucopyranosyl-1,3- β -D-fructose; Trehalulose: α -D-glucopyranosyl-1,1- β -D-fructose

AS that were characterized so far are produced by various species from the genus *Neisseria* (4), *Deinococcus* (14,15) and *Alteromonas* (16). However, *Neisseria polysaccharea* amylosucrase (NpAS) is the only AS for which several structures, alone or in complex with sucrose substrate or products are available to date (17-20). The three-dimensional structure of NpAS is organized in five domains, namely A, B, C, N and B'. B' and N domains are only found in AS whereas the three other ones are conserved amongst GH13 enzymes. Although the potential of NpAS for glycodiversification is large, this enzyme suffers from a low catalytic efficiency and a weak thermostability, limiting its industrial development (15). Directed evolution has been attempted to improve NpAS catalytic efficiency and thermostability (21). Searching for more thermostable and efficient enzymes in the natural diversity is another alternative that has motivated the biochemical characterization of the amylosucrases from *Deinococcus geothermalis* (DgAS) (15), *Deinococcus radiodurans* (DrAS) (14) and *Alteromonas macleodii* (AmAS) (15).

With a specific activity of 44 U.mg⁻¹ at the optimal temperature of 50 °C, the recombinant DgAS is the most thermostable AS characterized to date (15). The size distribution of the α -glucan chains produced by DgAS and NpAS differs and the two enzymes do not synthesize equivalent amounts of turanose and trehalulose. Indeed, DgAS produces significantly higher amounts of trehalulose than NpAS (15). Glucosylation of unnatural acceptors such as salicin also revealed that DgAS only produces a monoglucosylated form whereas NpAS is able to produce a di-glucosylated compound (22).

Here, we report the first 3D-structure of DgAS in its apo form that reveals an unusual homodimeric arrangement, thereby allowing identification of determinants possibly controlling thermostability. Furthermore, the structures of both DgAS and NpAS in complex with turanose (a fructose acceptor reaction product) were also determined and their analysis highlighted major differences upon fructose binding which are discussed with regard to the enzyme product specificities.

Experimental procedures

Expression and purification. Expression and purification of recombinant glutathione *S*-transferase-amylosucrase fusion proteins (GST-DgAS, GST-DrAS and GST-NpAS) were performed as previously described by Emond *et al.* (15), Pizzut *et al.* (14) and Potocki de Montalk *et al.* (23), respectively. The GST-AS proteins were purified by affinity

chromatography using Glutathion-Sepharose-4B support (GE Healthcare, Little Chalfont, UK). The GST tag was then removed using Prescission protease (GE Healthcare, Little Chalfont, UK) which left 5 residues (GPLGS) of the cutting site at the N-ter extremity.

Thermostability of amylosucrases. The melting point (T_m) of amylosucrases was assayed by differential scanning fluorimetry (DSF). A mix of enzyme (2 μ M), Sypro-orange (5 X) (Invitrogen, Paisley, UK) and 50 mM Tris, pH 7.0, 150 mM NaCl, 1 mM DTT, 1 mM EDTA were incubated using a temperature gradient from 20 to 80 °C with a 0.3 °C increment. The thermal transition was monitored using a RTQ-PCR CFX96 Real-Time System (Biorad, Marnes-la-Coquette, France). T_m was given by the inflexion point of the curve RFU=f(T), with RFU standing for Relative Fluorescence Unit. Circular dichroism spectra were recorded on a JASCO J815 spectropolarimeter equipped with a Peltier cell temperature controller. T_m values were obtained by heating the sample (enzyme 2 μ M in 50 mM Tris, pH 7.0, 150 mM NaCl, 70 μ M DTT, 70 μ M EDTA) at 1 °C/min and recording the ellipticity value at 220 nm from 25 °C to 80 °C with delays of 30 seconds. The SigmaPlot 10.0 software was used for all graphic analyses, T_m determination and statistics.

The half-life ($t_{1/2}$) at 50 °C of DgAS and NpAS preparations were also determined by incubating AS pure enzymes (300 mg.L⁻¹) in 50 mM Tris, pH 7.0, 150 mM NaCl, 1 mM DTT, 1 mM EDTA at 50 °C. At various intervals, aliquots were taken and the enzyme activity was determined as previously described (22).

Size exclusion chromatography multiangle laser light scattering (SEC-MALLS) experiments. NpAS, DgAS and DrAS protein samples buffered in 50 mM Tris pH 7.5, 150 mM NaCl were analyzed on a Shodex KW-803 column (Showa Denko Europe GmbH, Munich, Germany) with multiangle laser light scattering (MALLS). The column was equilibrated in a 0.1 μ m filtered 50 mM Tris pH 7.5, 150 mM NaCl, 0.02 % (w/v) NaN₃ buffer on a Agilent 1260 Infinity LC chromatographic system (Agilent Technology, Massy, France). Data were collected using a DAWN HELEOS-II 18-angle and Optilab T-rEX refractive index detector (Wyatt technology Corp., Toulouse France). Sample concentrations were 1.55 g.L⁻¹, 2.1 g.L⁻¹ and 1.85 g.L⁻¹ for NpAS, DgAS and DrAS solutions respectively. Protein samples were prepared in the Tris buffer used as the mobile phase to equilibrate the column. 20 μ l of each protein sample was loaded on the column and the separation was

performed at a flow rate of $0.5 \text{ mL}\cdot\text{min}^{-1}$ at $22 \text{ }^\circ\text{C}$. Results were analyzed using the ASTRA V software (Wyatt technology Corp., Toulouse France).

Native-PAGE. Native-PAGE was performed with a two-phase gel composed of a 4 % stacking gel (4 % acrylamide, 125 mM Tris, pH 6.8, 0.1 % APS, 0.15 % TEMES) and a 10 % separating gel (10 % acrylamide, 375 mM Tris, pH 8.8, 0.1 % APS, 0.15 % TEMED). $15 \mu\text{L}$ of the protein samples ($2 \text{ g}\cdot\text{L}^{-1}$) mixed with 2 X loading buffer (40 mM Tris, pH 6.8, 50 % glycerol (v/v), 0.04 % bromophenol blue (w/v)) were loaded onto the gel. Electrophoresis was performed at room temperature using a mini-PROTEAN system (BioRad, Marnes-la-Coquette, France) at a voltage of 100 V with an electrophoresis buffer composed of $15 \text{ g}\cdot\text{L}^{-1}$ Tris, pH 8.3 and $72 \text{ g}\cdot\text{L}^{-1}$ glycine. The gels were stained with Coomassie Brilliant Blue. Molecular weight estimation was performed using NativeMark native protein markers (Invitrogen, Paisley, UK) as standards.

Crystallization. NpAS was crystallized using conditions previously described by Skov *et al.* (17,18). DgAS crystallization experiments were carried out at $12 \text{ }^\circ\text{C}$ using the hanging drop vapour diffusion method. Best crystals were obtained with a 1:1 (v/v) ratio of protein ($6 \text{ mg}\cdot\text{mL}^{-1}$ in 20 mM Tris, pH 8.0) to precipitant solution (1.5 M sodium Acetate, 0.1 M sodium cacodylate, pH 7.0). Lens shape crystals appeared after two weeks and grew to a maximal size of $140\times 60\times 40 \mu\text{m}^3$.

Soaking experiments. Crystals of NpAS were soaked for 20 min in the reservoir solution supplemented by 250 mM turanose (Sigma, Saint-Martin-d'Hères, France). DgAS crystals quickly fractured in the presence of such a high concentration of turanose, so they were soaked for few seconds in the reservoir solution supplemented with only 14 mM of turanose.

Data collection and structure determination. X-ray experiments were carried out at 100 K. Prior to flash cooling, native crystals of DgAS were soaked for few seconds in the reservoir solution supplemented with 20 % (v/v) glycerol to avoid ice formation. Conversely, due to the cryoprotection effect of turanose, crystals of AS-turanose complexes were intrinsically cryoprotected. Native DgAS and NpAS-turanose diffraction datasets were collected to a maximum resolution of 1.97 \AA and 1.85 \AA , respectively on beamline ID14-1 at

the European Synchrotron Radiation Facility (ESRF, Grenoble, France). DgAS-turanose complex dataset was collected to 2.10 Å on the ESRF beamline ID29. Diffracted intensities were integrated using iMOSFLM (24) and scaled with SCALA (25) from the CCP4 software suite (26,27) and 5 % of the scaled amplitudes were randomly selected and excluded from the refinement procedure. Crystals of DgAS in *apo* form and in complex with turanose belong to the C222₁ spacegroup with 1 molecule per asymmetric unit giving a Matthews coefficient of 2.3 Å³/Da. Crystal of NpAS-turanose belongs to the P2₁2₁2 spacegroup with 1 molecule per asymmetric unit and a Matthews coefficient of 2.4 Å³/Da. Data collection statistics are given in Table 1. The native structure of DgAS was solved by the molecular replacement method using PHASER (28) and the structure of NpAS (PDB code: 1G5A) (18) as a search model. The translation function Z-score was 16.5 and R and R_{free} of the refined molecular replacement solution were 0.39 and 0.47, respectively. Structures of the complexes NpAS-turanose or DgAS-turanose were straight refined from their native structures using refmac5 (29).

Building and refinement. Structure refinement was performed with refmac5 from the CCP4 GUI (29) and models were manually reconstructed in SigmaA weighted electron density maps using COOT (30). Water molecules were automatically assigned and ligand molecules were manually fitted in residual maps. Final DgAS structures in *apo* form and in complex with turanose contain 651 residues out of the 655 theoretical residues with four missing residues at the C-terminal extremity. The final model of NpAS in complex with turanose contains 628 residues out of the 632 theoretical residues with four missing residues at the N-ter extremity. Refinement statistics are given in Table 1.

Table 1: *Data collection and refinement statistics*

	NpAS-turanose	DgAS	DgAS-turanose
Data collection			
Space group	P2 ₁ 2 ₁ 2	C222 ₁	C222 ₁
a, b, c (Å)	96.0, 116.3, 60.5	105.3, 110.2, 115.5	104.7, 110.4, 115.3
α, β, γ (°)	90.0, 90.0, 90.0	90.0, 90.0, 90.0	90.0, 90.0, 90.0
Resolution (Å)	29.27-1.85 (1.95-1.85)*	33.38-1.97 (2.08-1.97)	63.45-2.10 (2.21-2.10)
R_{sym}	0.096 (0.332)	0.086 (0.376)	0.099 (0.324)
I/σI	6.5 (2.2)	8.3 (2.0)	4.9 (2.1)
Completeness (%)	99.6 (100)	99.6 (99.8)	100 (98.9)
Redundancy	3.6 (3.6)	4.0 (3.9)	6.3 (4.2)
Nb of molecule / AU	1	1	1
Matthews coefficient (Å³/Da)	2.4	2.3	2.3
Refinement			
Resolution (Å)	29.10-1.85	33.22-1.97	13-2.10
No. of unique reflections	58,309 (8,452)	47,355 (6,837)	39,319 (5,648)
R_{work} / R_{free} (%)	14.7/18.6	14.4/18.7	14.6/20.7
Total number of atoms	5,867	5,956	5,712
Number of residues in the protein	632	651	651
Number of ligand molecules	2 Turanoses 1 PEG	1 Tris 13 Glycerols	1 Turanose
Number of water molecules	767	602	394
B-factors (Å²)			
Protein	11.6	17.9	20.7
Ligand	19.9	36.9	26.4
Water	21.2	27.5	26.1
Rmsd			
Bond lengths (Å)	0.022	0.024	0.023
Bond angles (°)	1.8	1.8	1.9

*Values in parentheses are for the outer resolution shell

Coordinates. Coordinates have been deposited at the protein data bank (PDB codes: 3UCQ, 3UER and 3UEQ for DgAS, DgAS-turanose complex and NpAS-turanose complex, respectively).

In vitro synthesis of sucrose isomers. Synthesis of sucrose isomers (turanose and trehalulose) was performed with 100 mM sucrose in the presence of 0 or 100 mM fructose at 30 °C or at enzyme optimum temperature (50 °C for DgAS and 37 °C for NpAs). At the end of the reaction (24 h), samples were centrifuged at 12,000 rpm for 10 min. Concentration of sucrose isomers was determined by HPAEC-PAD using an analytical CarboPAC™ PA100 (4x250 mm) column with a CarboPAC™ PA-100 Guard (4x50 mm). Detection was performed using a Dionex ED40 module with a gold working electrode and an Ag/AgCl pH reference electrode (Dionex, Sonnyvale, CA, USA).

Molecular Dynamics (MD) simulations. All MD simulations were carried out using the AMBER 9 suite of programs and the all-atom ff03 force field (31,32). The starting models were derived from the high resolution crystal structures of DgAS and NpAS (PDB code: 1G5A) (18). Fifteen and twenty Na⁺ cations were added to neutralize the DgAS and the NpAS monomers, respectively. Each protein together with its counter-ions was embedded in a rectangular parallelepipedal solvent box that left a space of 0.12 nm around the solute. TIP3P water molecules (approximately 28,000) were added using the LEaP module integrated in the AMBER9 package (33). Simulation preparation consisted in different phases of minimization, heating, equilibration under different type of restraints. The simulation was carried out at constant temperature (303 K) and pressure (1 bar) conditions over 60 ns. The temperature and pressure were controlled using Langevin thermostat (34) and Berendsen barostat (35) with a collision frequency (2 ps⁻¹) and pressure relaxation time (2 ps). Long-range electrostatic forces were handled by using the particle-mesh Ewald method (36). The time step of the simulations was 2.0 fs and the SHAKE algorithm was used to constrain the lengths of all chemical bonds involving hydrogen atoms to their equilibrium values (37).

To avoid artefacts, MD simulations were run three times with different starting velocity distribution. The resulting trajectories were analyzed using the Ptraj module of the AMBER9 package. The RMSD was calculated for the protein backbone atoms using least squares fitting. Distances between protein loops were calculated with respect to their center of mass. Atomic positional fluctuations (Δr_i^2) of protein backbone were calculated using the

coordinates of the 60 ns trajectories. A mass-weighted average value was then calculated for each residue. These parameters are related to the B-factors through the following relationship:

$$B_i = \frac{8\pi^2}{3} \langle \Delta r_i^2 \rangle$$

Sequence and structure alignments. Multiple sequence alignment was performed using Clustal W2 (38) and represented using ESPript (39). Structural alignments were generated using SwissPDB-viewer (40) and subsequently served as a basis to align sequences of homologous proteins of unknown 3D-structures. Multiple sequence alignment was then manually corrected. The root mean-square deviations (RMSDs) among three-dimensional structures were calculated using Superpose (41) from the CCP4 suite of programs (26,27).

Results and discussion

Overall structure and dynamics

Tertiary structure of DgAS. The three-dimensional structure of DgAS is organized into the five distinct domains already described for NpAS (18) (Fig. 28): the all helical N-terminal domain (residues 1 to 83), the central catalytic domain A (residues 84-177, 259-398 and 468-565), two regions protruding from domain A, constituting respectively domains B (178-258) and B' (399-467), and finally the β -stranded C-terminal domain (566-646) adopting a Greek key motif. DgAS and NpAS share 38 % sequence identity. Amongst the five regions, domains A, B and B' are the most conserved with a sequence identity of 43 %, 58 % and 52 %, respectively. Domains N and C only share 21 % and 19 % sequence identity, respectively.

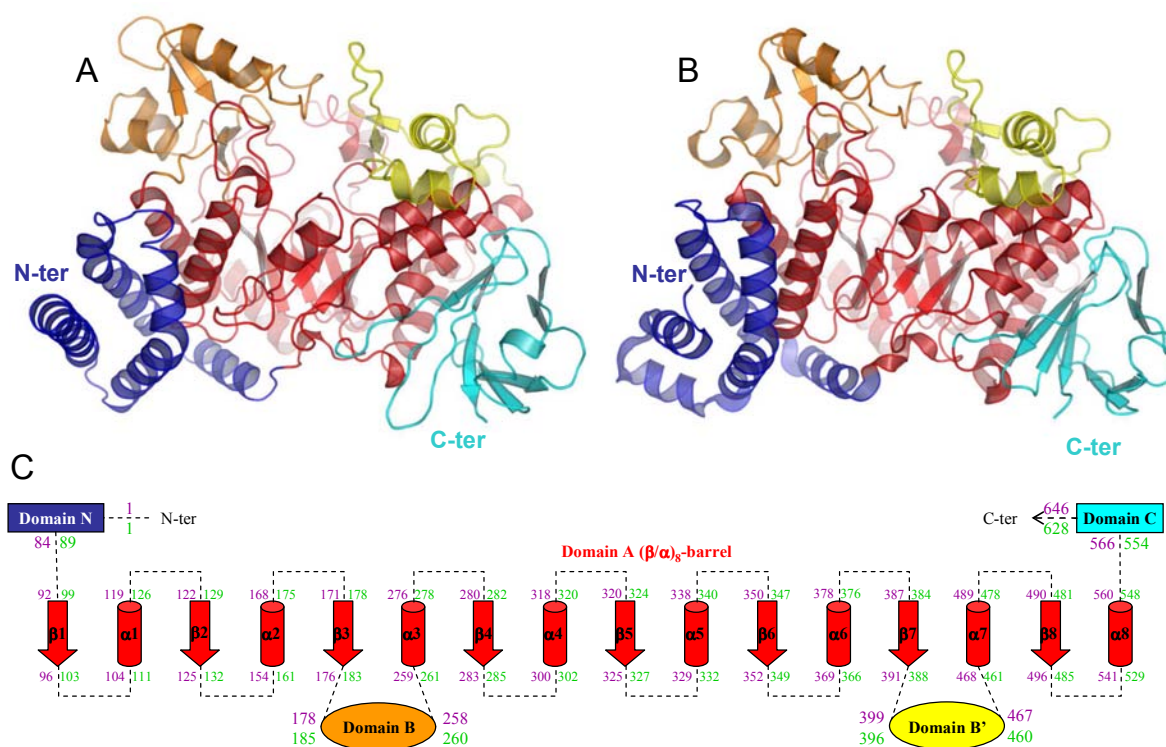


Figure 28: *Comparison of the three dimensional architecture of DgAS and NpAS*

Panel A. DgAS (code PDB: 3UCQ)

Panel B. NpAS (code PDB 1G5A (18))

Panel C. Schematic view of the DgAS and NpAS primary and secondary structures. The limits of the different domains are given in purple for DgAS and in green for NpAS. The same color code is used throughout the different views: catalytic domain A (red), domain N (blue), domain C (cyan), domain B (orange) and domain B' (yellow)

The overall RMSD value obtained after secondary structure superimposition of DgAS with NpAS was 1.3 Å. Lower RMSD values, i.e. 1.0 Å, 1.1 Å and 0.9 Å, were observed for domains A, B and B', respectively. More significant structural differences occurred between the less conserved domains N and C leading to larger RMSD values of 1.9 Å and 1.5 Å, respectively.

The catalytic domain A of DgAS adopts the typical $(\beta/\alpha)_8$ -barrel of GH family 13 (2). The active site is highly conserved with respect to that of NpAS (Fig. 29A). Residues E326 and D284 of DgAS were identified as the general acid/base and the nucleophile, respectively, involved in the formation of the β -glucosyl intermediate occurring in the α -retaining mechanism (corresponding to E328 and D286 in NpAS (42)). Three additional important residues conserved in GH13 family and known to assist in catalysis are D396, H180 and H395 (corresponding to D393, H187 and H392 in NpAS, respectively). A salt bridge formed by residues R520 and D137 (equivalent to R509 and D144 in NpAS) blocks the bottom of the

catalytic pocket. All the residues defining the catalytic site are conserved between NpAS and DgAS with the exception of R226, which is substituted by P219 in DgAS. As a result, the active site pocket of DgAS shows an enlarged aperture around subsite +2 compared to NpAS that may facilitate ligand accessibility and binding into the active site (Fig. 29B). Two ligands were found in the active site (Fig. 29A), a glycerol molecule (used as cryoprotectant for X-ray experiments) at subsite -1 and a Tris molecule at subsite +1, which are interacting together through a polar contact involving Glycerol-O2 and Tris-O3. Both B and B' domains are found longer in DgAS than in NpAS. Domain B is composed of 81 amino acid residues in DgAS (vs 76 residues in NpAS), organized in two antiparallel β -strands and two α -helices. As for domain B', it is composed of four α -helices comprising 69 amino acid residues (vs 65 in NpAS). The domain C of DgAS is composed of 82 amino acid residues folded as an eight-stranded β -sandwich as previously observed in the NpAS structure. The domains C of both enzymes differ in the short α -helix between strands β 4 and β 5 observed only in DgAS. Additionally, the loop interconnecting strands β 2 and β 3 is three residues longer in DgAS than in NpAS. The main difference between the DgAS and NpAS helical domains N resides in the N-terminal extremity which is composed by a long α -helix in DgAS and two small α -helices in NpAS (Fig. 29C). The remaining α -helices of the N-terminal domain (H3, H4 and H5) are structurally conserved in DgAS and NpAS.

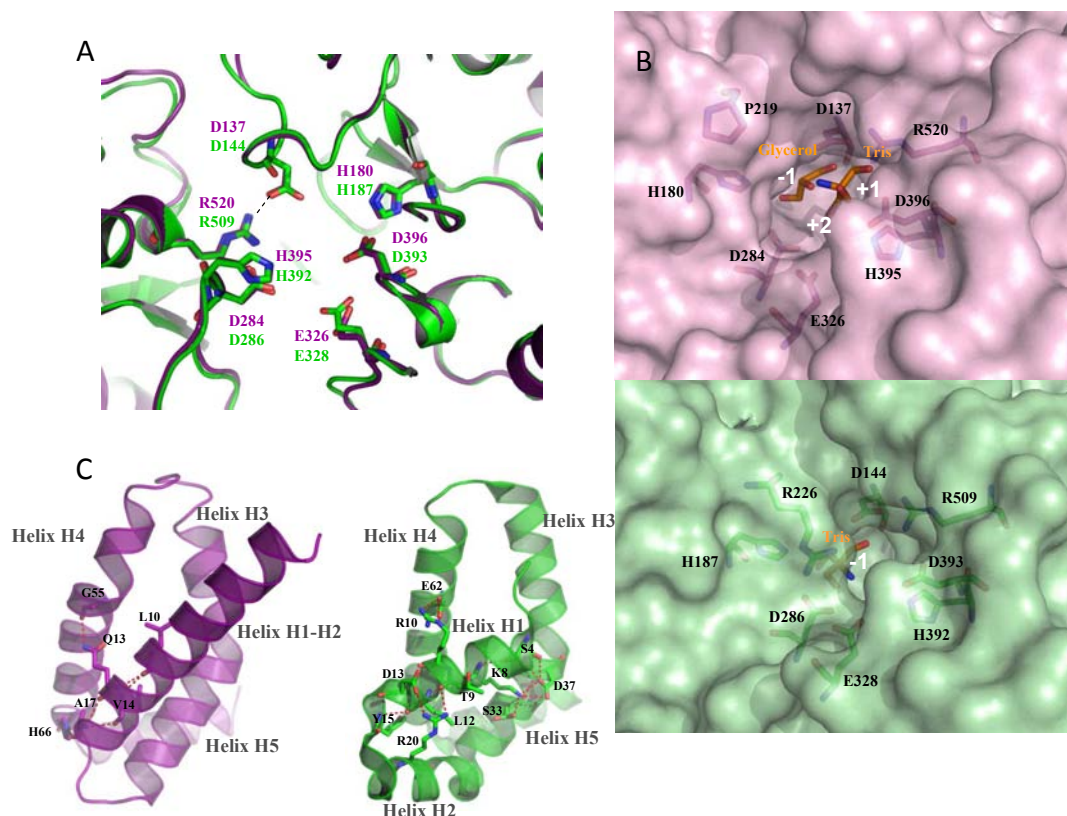


Figure 29: Comparison of DgAS and NpAS crystallographic structures

Panel A. Superimposition of the active sites of DgAS (purple, PDB code: 3UCQ) and NpAS (green, PDB code: 1G5A (18)). Catalytic residues are shown in stick. Ligands (Tris and Glycerol) were omitted for clarity purpose

Panel B. Molecular surface representation of the DgAS (purple) and NpAS (green) catalytic sites

Panel C. Representation of the polar interactions involved in the stabilization of the first helix of domain N in DgAS (purple) and NpAS (green).

Insight into enzyme plasticity through Molecular Dynamics (MD) simulations. To complete our structural comparison of amylosucrases, we undertook the investigation of their dynamic properties. Large scale MD simulations (60 ns) were thus carried out in explicit water for DgAS and NpAS.

As a check for the stability of protein structures during the course of MD simulations, the time evolution of the backbone atoms RMSD was calculated after least square fitting (Supplementary Fig. S1A). Variations of backbone RMSD indicated a distinct conformational behaviour of the two enzymes. The RMSD of DgAS rapidly increased by 3 Å within the first 12 ns of the simulation and then stabilized for the rest of the simulation. As for NpAS, the RMSD slowly increased by 2.2 Å over the first 30 ns before reaching a plateau (Supplementary Fig. S1A). Thus, faster (12 ns vs 30 ns) and larger-amplitude (3 Å vs 2.2 Å) changes were observed for DgAS compared to NpAS. MD simulations further revealed that

RMSD deviation of DgAS mainly resulted from a high mobility of the N-terminal part of the protein (residues 1-83). As shown in Supplementary Fig. S1B, the RMSD of the domain N increased by as much as 5.2 Å over the first 12 ns of the simulation for DgAS whereas it remained unchanged for NpAS. The high flexibility of the N-extremity was also shown by the large calculated B-factor values (Supplementary Fig. S1C). Such large conformational fluctuations of DgAS are due to the rearrangement of the first long helix (residues 1-24) into two shorter helices (composed of residues [1-14] and [15-24], respectively). Such conformational changes are not observed for NpAS as its first α -helix is already composed of two short helices in the X-ray structure. These helices are strongly stabilized through a large network of ionic interactions involving charged residues (R, K, D, E) from the three first α -helices belonging to the N-terminal region of NpAS (Fig. 29C). Noteworthy, there are no equivalent charged residues in DgAS. Even if other structural regions also appear flexible, as indicated by the calculated B-factors, differences between both enzymes are, however, less significant. Most flexible regions correspond to loops of the $(\beta/\alpha)_8$ -barrel that confer a pocket topology to the active site. In particular, loops 7 and 3 (encompassed in domains B' and B, respectively) moved away from each other during the MD simulations (Supplementary Fig. S1D) resulting in the opening of the catalytic pocket. With the exception of the N domain in DgAS and surface loops (2, 3, 4, 7 and 8) in both enzymes, remaining parts of DgAs and NpAS did not exhibit any significant fluctuation (Supplementary Fig. S1E and S1F).

Oligomerization and thermostability

The first dimeric structure of an amylosucrase. While the asymmetric unit of DgAS is composed of a single peptide chain, exploration of symmetry related molecular interfaces using the PISA server (43) indicated a stable dimeric organization with a buried interface of 2,950 Å² out of a total accessible surface area of 45,330 Å² (Fig. 30A). The buried surface represents 6.1 % of the solvent accessible area. The interface is defined by a set of residues located in the catalytic domain (A318-A320; L334-H344; A378-S385) and also in domains N (D21-G35; R73-R88) and C (T560-E568; R584-G590) (Fig. 30B). Stabilization between the protomer units of DgAS benefits from strong salt bridge interactions formed between R74 and E25 and between R341 and D84, and from a network of direct and water-mediated hydrogen-bonding interactions (Supplementary Fig. S2). The residues involved in the interface stabilization of DgAS are not conserved in NpAS. In particular, R341 which appears to play a

key role in stabilizing the DgAS interface, is one of the five additional residues of the $\alpha 5$ - $\beta 6$ interconnecting loop from the $(\beta/\alpha)_8$ -barrel, which is found longer compared to NpAS (Fig. 28C and 30B). Of note, such insertion region in DgAS is also found to interact through polar interactions with loop R584-M591 from the domain C of the second protomer, which is four residues shorter in NpAS (Fig. 30C).

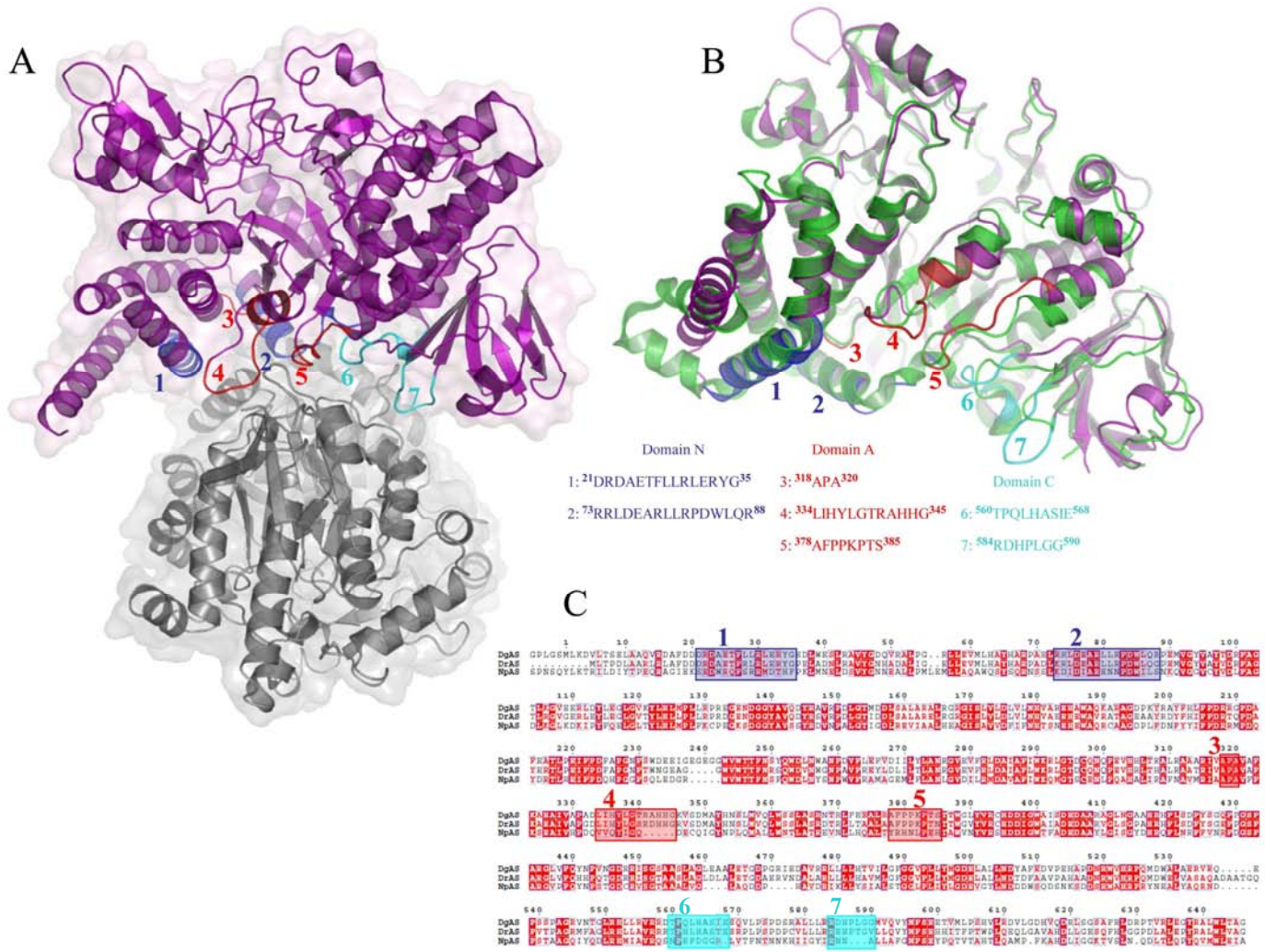


Figure 30: Quaternary structure organization of DgAS

Panel A. Visualization of the DgAS dimer (PDB code: 3UCQ). The protomers are symmetry-related following a two-fold crystallographic symmetry

Panel B. Structural superimposition of DgAS (purple) and NpAS (PDB code 1G5A (18)) (green) monomeric units. Structural elements involved in the dimeric interface of DgAS involve 7 regions of domains N, A and C. The amino acid sequences composing the 7 regions are listed

Panel C. Sequence alignment of DgAS, DrAS and NpAS. The 7 regions are shown on the alignment using the same color code as in panel B. A high resolution alignment is providing in Supplementary Fig. S6

The dimerization interface displays a strong shape complementarity between the protomers. First, a bundle of hydrophobic residues from the N domain involving L80 and L81 from α -helix H5 of the first protomer fits nicely into the groove formed by the two α -helices, H3 and H5, from the second protomer (Supplementary Fig. S2). Residues L80, L81 and L29, belonging to these α -helices and only present in DgAS, are involved in these hydrophobic contacts (Supplementary Fig. S2A). In addition, the extremity of the loop R584-G590 (insertion region situated between strands β 3 and β 4 of the C-terminal domain in DgAS as described earlier) is found interlocked between two loops, 339-344 (another insertion region in DgAS) and 379-385 (highly conserved in the *Deinococcus* genus). Interestingly, in such contacting amino acid residues, we observe a bunch of prolines, which are generally known to rigidify the polypeptide backbone and that may constrain the polypeptide chain conformation to facilitate the dimer formation. P587 from the first protomer faces up P380, P381 and P383 from the second one. Although oligomeric states have been previously observed for other members of the GH13 family, such as cyclomaltodextrinases (CDase; EC 3.2.1.54, PDB code: 1H3G (44)), maltogenic amylase (MAase; EC 3.2.1.133, PDB code: 1SMA (45)), and neopullulanase (NPase, EC, 3.2.1.135, PDB code: 1J0H (46)), the structure of DgAS is, to our knowledge, the first homodimeric amylosucrase reported to date. Comparison of electrophoresis under denaturing (SDS-PAGE) and native conditions argue for a dimeric assembly of DgAS in solution (Fig. 31A and 31B).

The dimeric conformation of DgAS in solution was confirmed by SEC-MALLS/RI experiments. As shown in Fig 31C, the elution volume of DgAS differed significantly from that of NpAS, corresponding to a molecule of molecular weight of 151.7 kDa compatible with a dimeric assembly of DgAS (theoretical molecular weight of 146.3 kDa). As expected, the NpAS mainly eluted as a monomer with a molecular weight of 74.5 kDa (theoretical molecular weight of 71.5 kDa). The monomeric NpAS and the dimeric DgAS remained stable over few weeks at 4 °C.

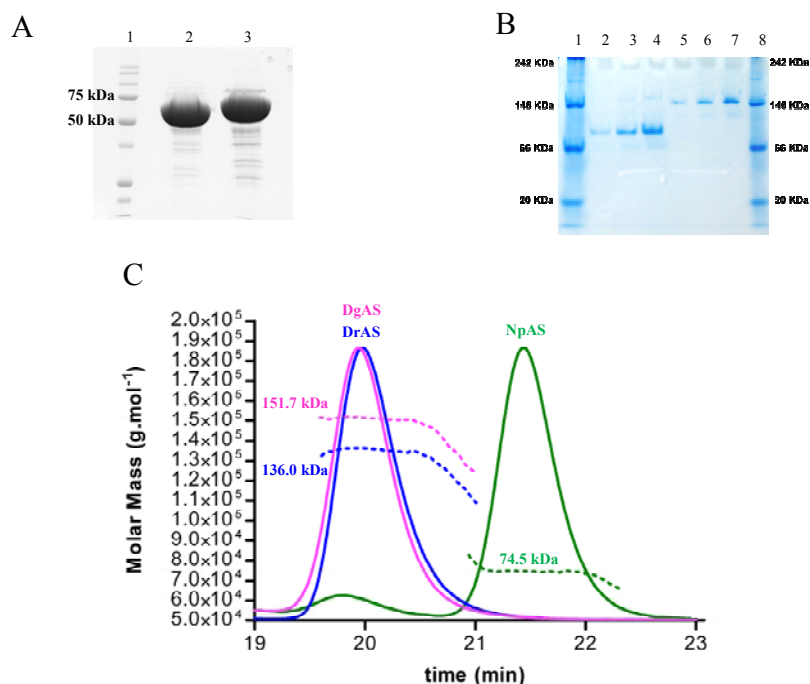


Figure 31: Biochemical analysis of DgAS and NpAS

Panel A. SDS-PAGE of protein markers (lane 1), DgAS (lane 2) and NpAS (lane 3)

Panel B. Native PAGE of protein markers (lanes 1 and 8), NpAS (lanes 2, 3, 4) and DgAS (lanes 5, 6, 7); for both enzymes, no effect of dilution is visible (concentration in proteins are 0.5, 1 and 2 g.L⁻¹, respectively)

Panel C. SEC-MALLS/RI experiments for NpAS (green line), DgAS (purple line) and DrAS (blue line).

Continuous lines represent the light scattering curve. The experimentally measured molecular weight distribution (as horizontal lines) and the average molecular weight are indicated for each elution peaks

To search for a sequence signature of dimeric forms in other amylosucrases, we performed a sequence alignment of DgAS, NpAS and DrAS (Fig. 30C). Among the 7 regions found to be involved in DgAS dimerization from structural analysis (listed in Fig. 30B), regions 4, 6 and 7 are well conserved in DrAS. Notably, the corresponding loops in NpAS are shorter and could prevent dimer formation. To get some evidence of DrAS dimerization, SEC-MALLS analysis of DrAS was performed and it confirmed the homodimeric organization of this enzyme with a molecular weight of 136.0 kDa (theoretical molecular weight of 143.3 kDa) (Fig. 31C). The pattern composed of regions 1 to 7 could thus be proposed as a signature of dimerization. We screened the identified pattern against sequences of GH13 enzymes (sequence alignment is provided in Supplementary Fig. S3). The proposed dimerization pattern was found in *Deinococcus desertii* DdAS, *Deinococcus maripensis* DmAS, α -amylases from *Meiothermus rubber*, *Meiothermus silvanus* and *Truepera radiovictrix* sequences, suggesting a similar dimeric organization to DgAS. On the other hand,

sequences of amylosucrases from *Neisseria* and *Alteromonas* groups did not contain this molecular pattern. Like NpAS, one could infer that they might adopt a monomeric form (23).

DgAS thermostability. The optimum temperature of DgAS is 50 °C whereas an optimal temperature of 37 °C was found for NpAS (15). The $t_{1/2}$ of freshly purified DgAS and NpAS was 69 h and 15 min at 50 °C, respectively. The higher stability of DgAS was confirmed by T_m measurements which gave 58.9 +/- 0.2 °C for DgAS versus 49.6 +/- 0.4 °C for NpAS when measured by DSF and 62.0 +/- 0.2 °C versus 51.3 +/- 0.2 °C when determined by circular dichroism. Close inspection of the primary sequences of DgAS and NpAS gave rise to the following observation: DgAS contains i) a higher proportion of charged residues (K, R, H, D, E), which represent 28.8 % of the total number of residues compared to only 22.9 % in NpAS, ii) a lower number of polar uncharged residues (G, S, T, N, Q, Y, C) which represent 24.3 % of the total number of residues for DgAS and 33.4 % for NpAS, iii) more hydrophobic residues (L, M, I, V, W, P, A, F) representing 46.9 % of the total number of residues compared to 43.7 % for NpAS and iv) six additional proline residues. All aforementioned features are usually observed in proteins from thermophilic organisms and could contribute to the higher thermostability of DgAS (47,48). In addition, we have shown using MD simulations on DgAS and NpAS monomers that the N-terminal extremity of DgAS exhibits a higher flexibility. DgAS dimerization might then constrain the movement of helix H1-H2 at the N-terminus of the protein and one could thus assume that dimerization is likely to contribute significantly to enzyme stabilization of this enzyme.

Specificity toward sucrose isomer formation

Previous characterization of DgAS and NpAS revealed that both enzymes catalyze turanose and trehalulose synthesis as side-products when incubated with sucrose but the ratio of turanose *versus* trehalulose production was radically different for the two enzymes (15). Other enzymes catalyzing exclusively sucrose isomerization reactions were previously characterized (7). Their primary structures contain a highly conserved motif named “charged isomerization motif” composed of residues RLDRD which are involved in the product specificity (49). Additionally, these enzymes contain two conserved phenylalanine forming the so-called “aromatic clamp” motif, which is involved in the tautomerization of fructose occurring during the isomerization reaction (8). Both sites are neither conserved in primary nor

in tertiary structures of amylosucrases what might help to explain differences observed between product profiles of amylosucrases and sucrose isomerases. To further investigate the specificity for sucrose isomerization of DgAS and NpAS, experiments were carried out in the presence of 100 mM sucrose and of exogenous fructose acceptor at either 30 °C or at the optimal temperature of each enzyme, *i.e.* 50 °C and 37 °C for DgAS and NpAS, respectively. The results reported in Table 2 show that, for all operational conditions, NpAS preferentially synthesized turanose. In contrast, equivalent amounts of turanose and trehalulose were produced by DgAS. Increasing initial fructose concentration enabled sucrose isomer production yield to be enhanced. Acceptor reaction onto fructose was forced at the detriment of α -glucan chain formation. At 30 °C, 69 % and 64 % of the glucosyl units coming from sucrose were transferred onto D-Fructose by DgAS and NpAS, respectively. However, the ratio of turanose *versus* trehalulose remained unchanged compared to that obtained in the absence of exogenous fructose acceptor. Similarly, increasing the temperature had no significant effect on sucrose isomer ratio. To go further in the understanding of this feature, the crystal structures of DgAS and NpAS in complex with turanose were determined. These are the first complexes of amylosucrases with sucrose isomers and electron density maps showed in both cases density which corresponds to turanose.

Table 2: Comparison of turanose and trehalulose formation by DgAS and NpAS expressed as the percentage of the glucosyl units coming from sucrose transferred onto fructose

Enzyme	NpAS				DgAS			
	30 °C		37 °C (Optimum temperature)		30 °C		50 °C (Optimum temperature)	
Reaction conditions	Suc 100 mM	Suc 100 mM Fru 100 mM	Suc 100 mM	Suc 100 mM Fru 100 mM	Suc 100 mM	Suc 100 mM Fru 100 mM	Suc 100 mM	Suc 100 mM Fru 100 mM
% of glucosyl residues in turanose	20.6	51.9	19.4	44.0	12.9	34.1	12.4	33.4
% of glucosyl residues in trehalulose	4.5	11.9	4.1	9.0	12.8	35.2	10.6	26.5

The NpAS-turanose complex was first compared to the structure of inactive E328Q NpAS mutant in complex with sucrose (PDB code: 1JGI) and no difference exceeding the structure coordinate errors was observed between the two structures (19). The structure of the NpAS-turanose complex revealed two turanose-binding sites, which were named TB1 and

TB2 for the site located in the catalytic pocket and at the surface of the C domain, respectively. The electron density corresponding to the turanose bound at subsite -1 and +1 of the catalytic pocket is shown in Figure 32A. The glucosyl moiety occupies the catalytic subsite -1 in a 4C_1 chair conformation. Its pyranosyl ring can be nicely superimposed onto the sucrose glucosyl ring from X-ray structure (PDB code: 1JGI) and is maintained by the same dense network of hydrogen bonding interactions involving residues D144, H187, E328, R284, D286, H392, D393, R509 (Supplementary Table S1). Glucosyl binding is also reinforced by a stacking interaction with Y147 and hydrogen bonds with R513 and D144 mediated by a water molecule (Supplementary Fig. S4). The most striking observation arose from the fructosyl moiety binding conformation. Indeed, at subsite +1, the electron density of the ligand unequivocally revealed that the fructose unit of turanose is bound in the open state. This is quite surprising, considering that in water solution, turanose adopts an equilibria with a predominant proportion of β pyranoid (47 %) and β furanoid (37 %) tautomers after 20 min at 20 °C as determined by 1H NMR spectroscopy (50). The hydroxyl groups O4', O5' and O6' of the open state fructose form direct hydrogen bonds with residues A287, E328 and I330, respectively. Noteworthy, a dense network of additional hydrogen bonds mediated by water molecules is also observed around fructose (Table 3). Within this network, the water molecule Wat1134 mediates two hydrogen bonds between R226 and both O1' and O6'. In addition O6' is also within hydrogen bond distance of the main chain NH of I330 and water molecule Wat1366. These interactions involving O1' and O6' of the fructose unit probably favor fructose binding in the open state by preventing hemi-acetal formation. Such a binding mode allows a perfect positioning of the O3' for the nucleophilic attack of the β -glucosyl intermediate and the formation of turanose. It is likely responsible for the predominant turanose formation by NpAS. In the turanose molecule found in TB2, the fructose ring adopts a β -D-furanosyl conformation (Supplementary Fig. S5).

Table 3: Interactions of fructose unit from sucrose and turanose with NpAS and DgAS

Fructosyl moiety in NpAS sucrose complex (PDB: 1JGI) (19)	Fructose unit of NpAS turanose complex	α -Fructose unit of DgAS turanose complex	β -Fructose unit of DgAS turanose complex
	A287(N)•••Fru(O4') 3.2 Å		
Q328(OE1)•••Fru(O1') 2.6 Å	E328(OE1)•••Fru(O4') 2.6 Å	E326(OE1)•••Fru(O1') 3.4 Å	E326(OE1)•••Fru(O2') 3.4 Å
Q328(NE2)•••Fru(O1) 3.2 Å	E328(OE2)•••Fru(O5') 3.1 Å	E326(OE2)•••Fru(O2') 2.6 Å	E326(OE2)•••Fru(O2') 3.2 Å
Q328(NE2)•••Fru(O1') 2.8 Å			
Q328(NE2)•••Fru(O3') 3.4 Å			
	I330(N)•••Fru(O6') 3.0 Å		
D393(OD2)•••Fru(O3') 2.8 Å		D396(OD2)•••Fru(O4') 3.5 Å	D396(OD2)•••Fru(O4') 3.5 Å
D394(OD1)•••Fru(O6') 2.8 Å			
D394(OD2)•••Fru(O6') 3.4 Å			
R446(NH2)•••Fru(O6') 2.8 Å			
Water molecules			
	Wat808•••Fru(O6') 2.8 Å	Wat808•••Fru(O1') 3.2 Å	
	Wat838•••Fru(O2') 3.1 Å	Wat838•••Fru(O4') 2.6 Å	Wat838•••Fru(O4') 2.6 Å
Wat1102•••Fru(O3') 3.5 Å			
Wat1111•••Fru(O3') 3.0 Å			
Wat1134•••Fru(O4') 2.8 Å	Wat1134•••Fru(O1') 3.0 Å		
	Wat1134•••Fru(O6') 2.6 Å		
		W1151•••Fru(O6') 3.0 Å	W1151•••Fru(O6') 3.0 Å
Wat1273•••Fru(O1') 2.7 Å			
Wat1379•••Fru(O4') 2.7 Å	Wat1379•••Fru(O1') 2.8 Å		
	Wat1447•••Fru(O1') 2.7 Å		
	Wat1447•••Fru(O5') 2.8 Å		
	Wat1448•••Fru(O5') 3.1 Å		

In DgAS, only one turanose-binding site, equivalent to TB1 of NpAS, was observed (Fig. 32B). The absence of TB2 is likely due to the fact that residues corresponding to F559, N560 and N562 from domain C of NpAS are substituted by three hydrophobic residues: L572, P573 and P575, respectively. Such substitutions prevent stacking and hydrogen bonding interactions with turanose. In the TB1 site of DgAS, the glucosyl moiety of turanose occupies subsite -1 and adopts a 4C_1 configuration with a very clear electron density. Interactions between the glucopyranosyl ring and the protein are similar to those described for the NpAS-turanose complex (Supplementary Table S1). The electron density of the fructosyl ring is less clear at subsite +1. However, it defines an envelope that fits best with the α -anomer of fructofuranose (Fig. 32B). The β -anomer also probably bound subsite +1 but to a lesser extent. In addition, we cannot exclude the presence of fructose either in open or pyranoid forms. The O3' glucosylated- α -D-fructofuranose is reported to be one of the less

prevalent tautomer of turanose in water solution (1 % at 20 °C (50)). Accommodation of this conformation in subsite +1 involves H-bonds with the catalytic residues E326 and D396 and two water-mediated interactions with O1', O4' and O6'. DgAS binds the furanoid tautomers of fructose to form turanose but *via* a considerably weaker network of interactions than that observed in NpAS. Further comparisons of DgAS and NpAS subsite +1 revealed structural traits that prevent open state conformation of fructose to bind DgAS subsite +1 (Fig. 32C). Indeed, residue R226 of NpAS, which interacts *via* a water molecule with both O1' and O6' of the open form of fructose, is replaced by P219 in DgAS. Furthermore, residue I328 corresponding to I330 in NpAS, which was directly involved in H-bond interaction with O6' of fructose in open state, cannot play the same role in DgAS due to a slight motion of the loop bearing this residue. The discrimination between the various tautomers of turanose is less drastic in DgAS than in NpAS. DgAS subsite +1 accommodates various fructose tautomers in order to optimally arrange O1' or O3' of fructose unit for nucleophilic attack of the glucosyl enzyme intermediate therefore yielding equivalent amounts of trehalulose and turanose.

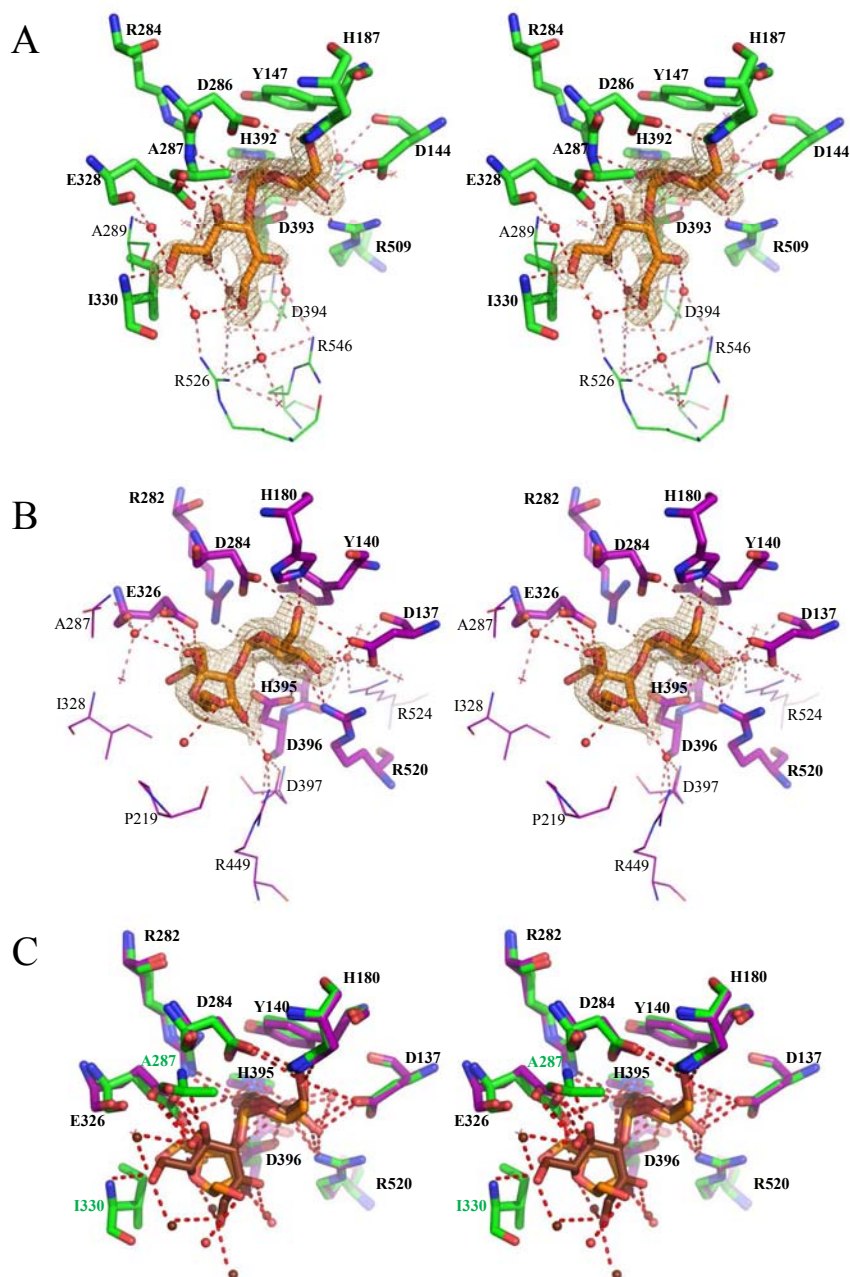


Figure 32: Stereoview of turanose conformation bound to NpAS and DgAS

with a sigmaA weighted $2F_o - F_c$ electron density map contoured at 1.0σ around the ligand (in orange).

Hydrogen bonding interactions are shown as red dashed line and water molecules as red spheres

Panel A. TB1 turanose-binding site of NpAS (in green, PDB code: 3UEQ).

Panel B. TB1 turanose-binding site of DgAS (in purple, PDB code: 3UER), the two alternative conformations of the ligand corresponding to the two turanose anomers α and β are represented.

Panel C. Stereoview of the structural superimposition of TB1 turanose binding sites of NpAS and DgAS using the same colour code except for the turanose-NpAS complex where the turanose and water molecules are brown coloured for clarity purpose. Residue numbering has been done according to DgAS sequence numbering with the exception of I330 and A287 (in green) that are only present in the interaction between NpAS and turanose.

DgAS- and NpAS-turanose complex analysis thus enabled us to explain differences observed in DgAS and NpAS product specificity. This work opens new perspectives for the rational and/or semi-rational redesign of this thermostable amylosucrase in order to modulate sucrose isomer synthesis with the view of reducing side-product formation or controlling sucrose isomer profile for the development of novel syrups enriched in sucrose substitutes.

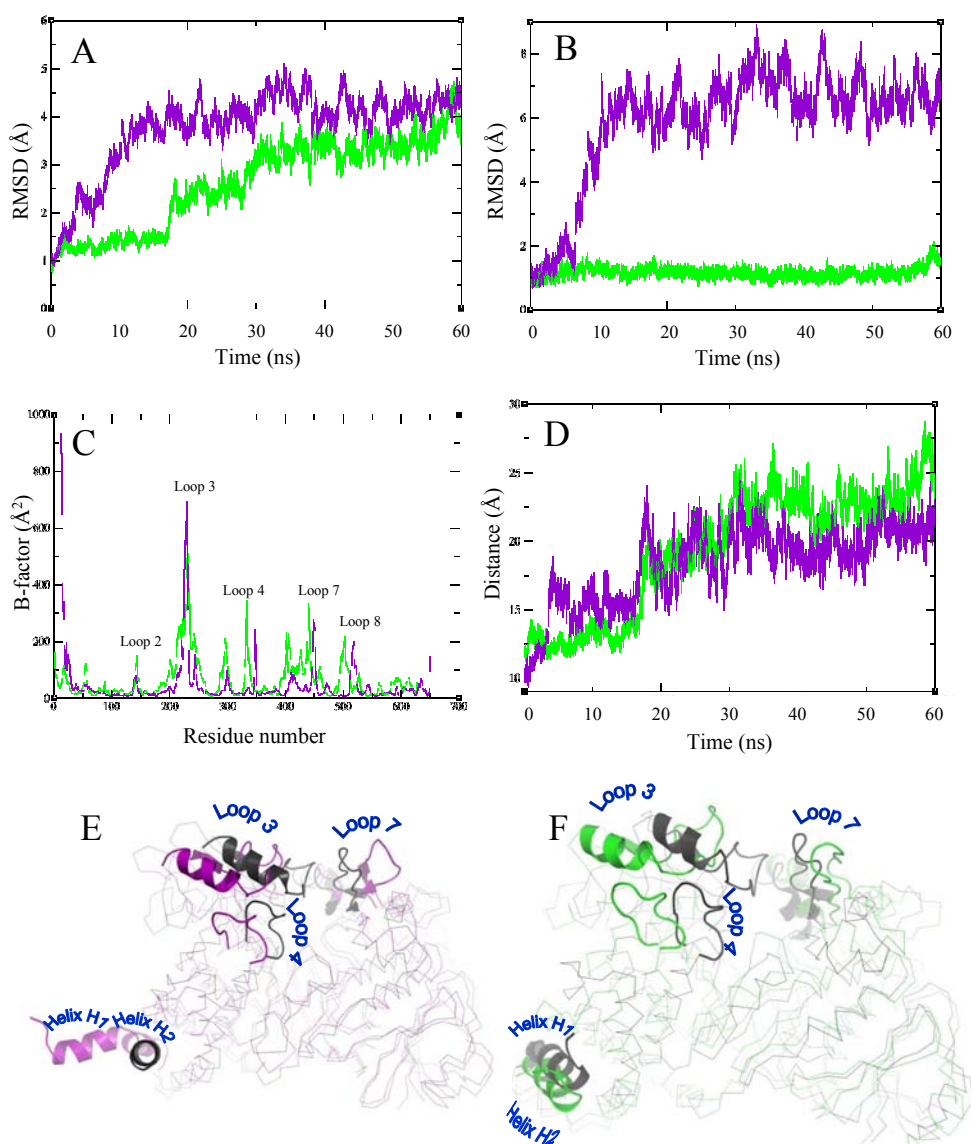
Acknowledgments

F. Guérin is supported by a PhD grant from the Pôles de Recherche et d'Enseignement Supérieur de l'Université de Toulouse and the Région Midi-Pyrénées, France. We thank the Computing Center of Region Midi-Pyrénées (CALMIP, Toulouse, France) and the Center for Computing Resources (CRI) of INSA-Toulouse for providing computing resources and support. We are grateful to the staff of synchrotron beamlines ID14-1 and ID29 at European Synchrotron Radiation Facility (Grenoble, France) for providing assistance in using the beamlines. Authors are grateful to Javier Perez (Synchrotron SOLEIL) for the fruitful discussions. We also thank Stéphanie Terme (Wyatt technology France) for her helpful assistance in the SEC-MALLS experiments.

Supplementary informations

*Supplementary Table S1: Interactions of glucosyl moieties from sucrose and turanose
in complex with NpAS and DgAS*

Glucosyl moiety in NpAS- sucrose complex (PDB code: 1JGI) (19)	Glucosyl moiety of NpAS- turanose complex (PDB code: 3UEQ)	Glucosyl moiety of DgAS-turanose complex (PDB code: 3UER)	
		α -anomer	β -anomer
D144(OD2)•••Glc(O4) 2.5 Å	D144(OD2)•••Glc(O4) 2.7 Å	D137(OD2)•••Glc(O4) 3.5 Å	D137(OD2)•••Glc(O4) 3.5 Å
		D137(OD2)•••Glc(O6) 2.7 Å	D137(OD2)•••Glc(O6) 2.7 Å
H187(NE2)•••Glc(O6) 3.9 Å	H187(NE2)•••Glc(O6) 2.8 Å	H180(NE2)•••Glc(O6) 3.0 Å	H180(NE2)•••Glc(O6) 3.0 Å
R284(NH2)•••Glc(O2) 2.8 Å	R284(NH2)•••Glc(O2) 2.8 Å	R282(NH2)•••Glc(O2) 3.2 Å	R282(NH2)•••Glc(O2) 3.2 Å
D286(OD2)•••Glc(O6) 3.1 Å	D286(OD2)•••Glc(O6) 3.0 Å	D284(OD2)•••Glc(O6) 3.1 Å	D284(OD2)•••Glc(O6) 3.1 Å
Q328(NE2)•••Glc(O1) 3.2 Å	E328(OE2)•••Glc(O2) 3.4 Å		
H392(NE2)•••Glc(O2) 2.9 Å	H392(NE2)•••Glc(O2) 2.9 Å	H395(NE2)•••Glc(O2) 2.9 Å	H395(NE2)•••Glc(O2) 3.0 Å
H392(NE2)•••Glc(O3) 2.9 Å	H392(NE2)•••Glc(O3) 2.9 Å	H395(NE2)•••Glc(O3) 2.9 Å	H395(NE2)•••Glc(O3) 2.9 Å
D393(OD1)•••Glc(O3) 2.9 Å	D393(OD1)•••Glc(O3) 2.7 Å	D396(OD1)•••Glc(O3) 2.7 Å	D396(OD1)•••Glc(O3) 2.7 Å
D393(OD2)•••Glc(O2) 2.5 Å	D393(OD2)•••Glc(O2) 2.7 Å	D396(OD2)•••Glc(O2) 2.9 Å	D396(OD2)•••Glc(O2) 2.9 Å
	D393(OD2)•••Glc(O3) 3.4 Å	D396(OD2)•••Glc(O3) 3.4 Å	D396(OD2)•••Glc(O3) 3.4 Å
R509(NH1)•••Glc(O3) 3.4 Å	R509(NH1)•••Glc(O3) 3.4 Å	R520(NH1)•••Glc(O3) 3.3 Å	R520(NH1)•••Glc(O3) 3.3 Å
R509(NH1)•••Glc(O4) 3.0 Å	R509(NH1)•••Glc(O4) 3.0 Å	R520(NH1)•••Glc(O4) 2.9 Å	R520(NH1)•••Glc(O4) 2.9 Å
Water molecules			
Wat1110•••Glc(O3) 2.6 Å	Wat1110•••Glc(O3) 2.8 Å	Wat1110•••Glc(O3) 2.8 Å	Wat1110•••Glc(O3) 2.8 Å



Supplementary Figure S1: Quantitative measurement of DgAS (purple) and NpAS (green) movements during the course of 60 ns MD simulations carried out in explicit water, starting from the X-ray structures (DgAS PDB code: 3UCQ and NpAS PDB code: 1G5A)

Panel A. Root mean squared deviation of overall AS backbone atoms as a function of time.

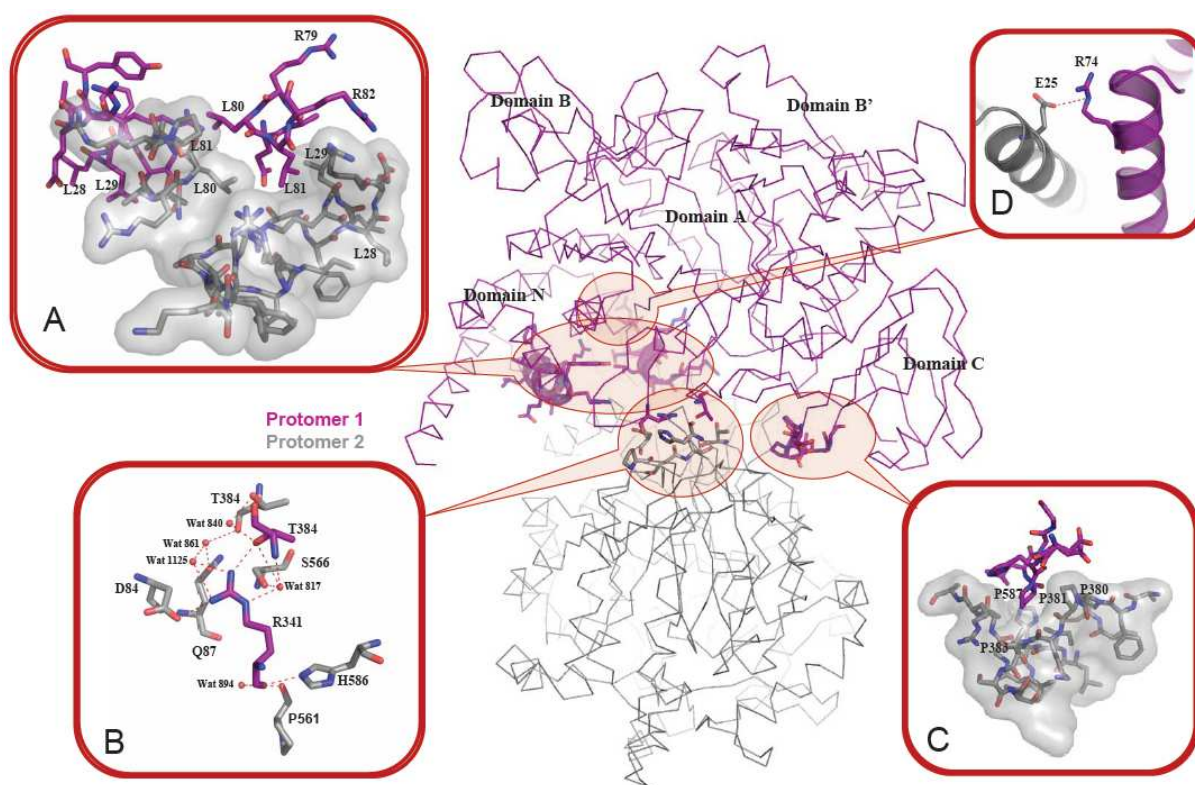
Panel B. Time variation of the RMSD of backbone atoms of domain N.

Panel C. Calculated B-factors of AS residues from MD simulations.

Panel D. Time variation of the distance between loop 3 and loop 7 (belonging respectively to domain B and B').

Panel E. Superimposition of DgAS structure at the beginning (black) and at the end (purple) of the MD simulation.

Panel F. Superimposition of NpAS structure at the beginning (black) and at the end (green) of the molecular dynamic simulation.



Supplementary Figure S2: View of the main interactions involved in the dimer interface of DgAS

Panel A. A bundle of leucine residues in domain N.

Panel B. A hydrogen-bonding interaction network with R341 residue.

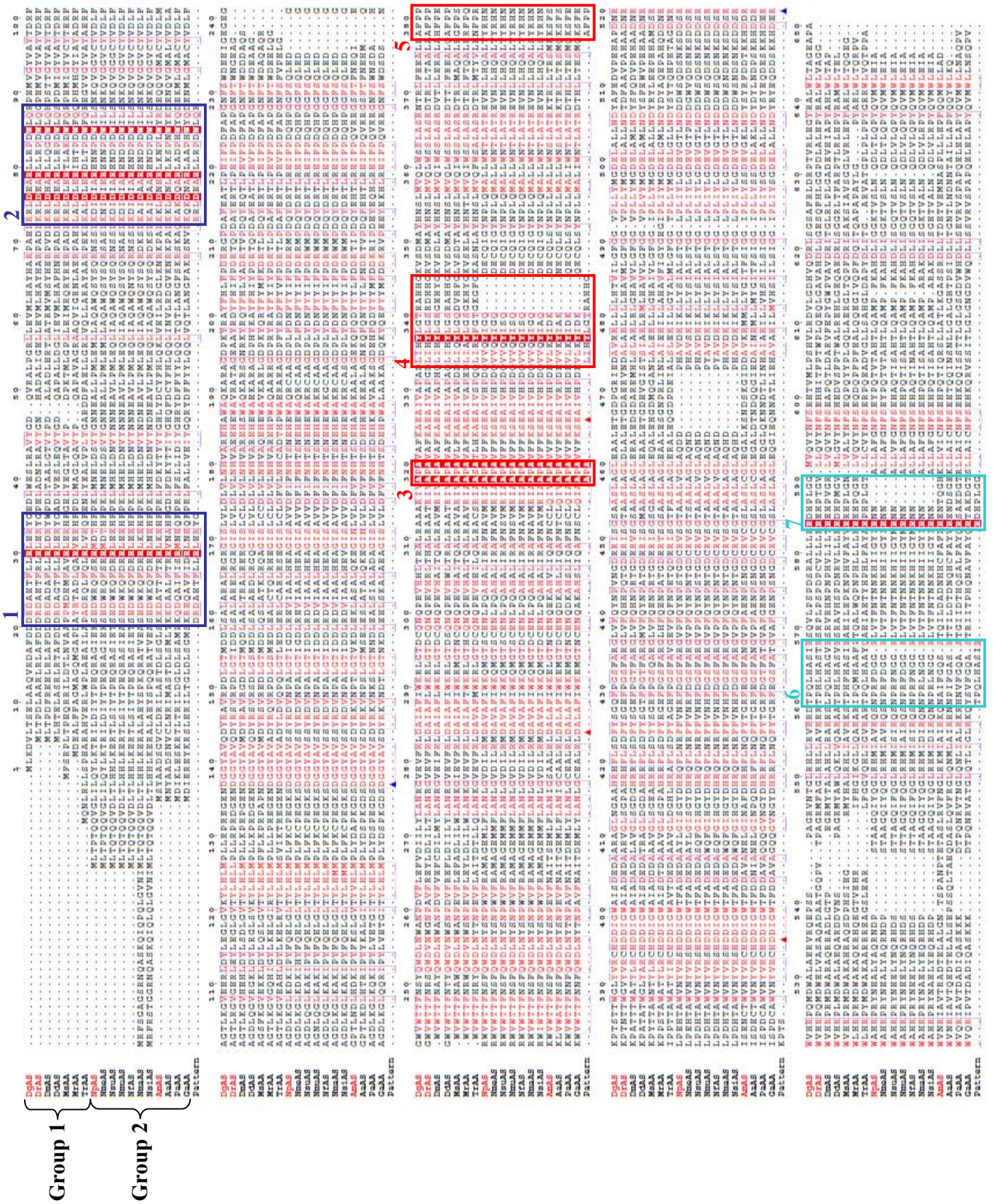
Panel C. A hydrophobic bundle of proline residues from domains C and A.

Of note, this last interaction is found at two locations of the dimerization interface but only one is depicted for clarity purpose. Salt bridge interaction between E25 and R74 is shown in **Panel D**.

In addition to the interactions depicted on the figure, some polar interactions not shown for sake of clarity are mediated by either one or several water molecules (Wat 1021 940 1078 965 894 817 817 840). They are involved in interactions with residues from both protomers: R30-E77, R33-D84, Y34-L81, L80-L80, R341-P561, R341-S566, T384-T384 and R22-R74, T26-E77, L29-A78, Q87-R341, Q87-T384, R341-T384, respectively.

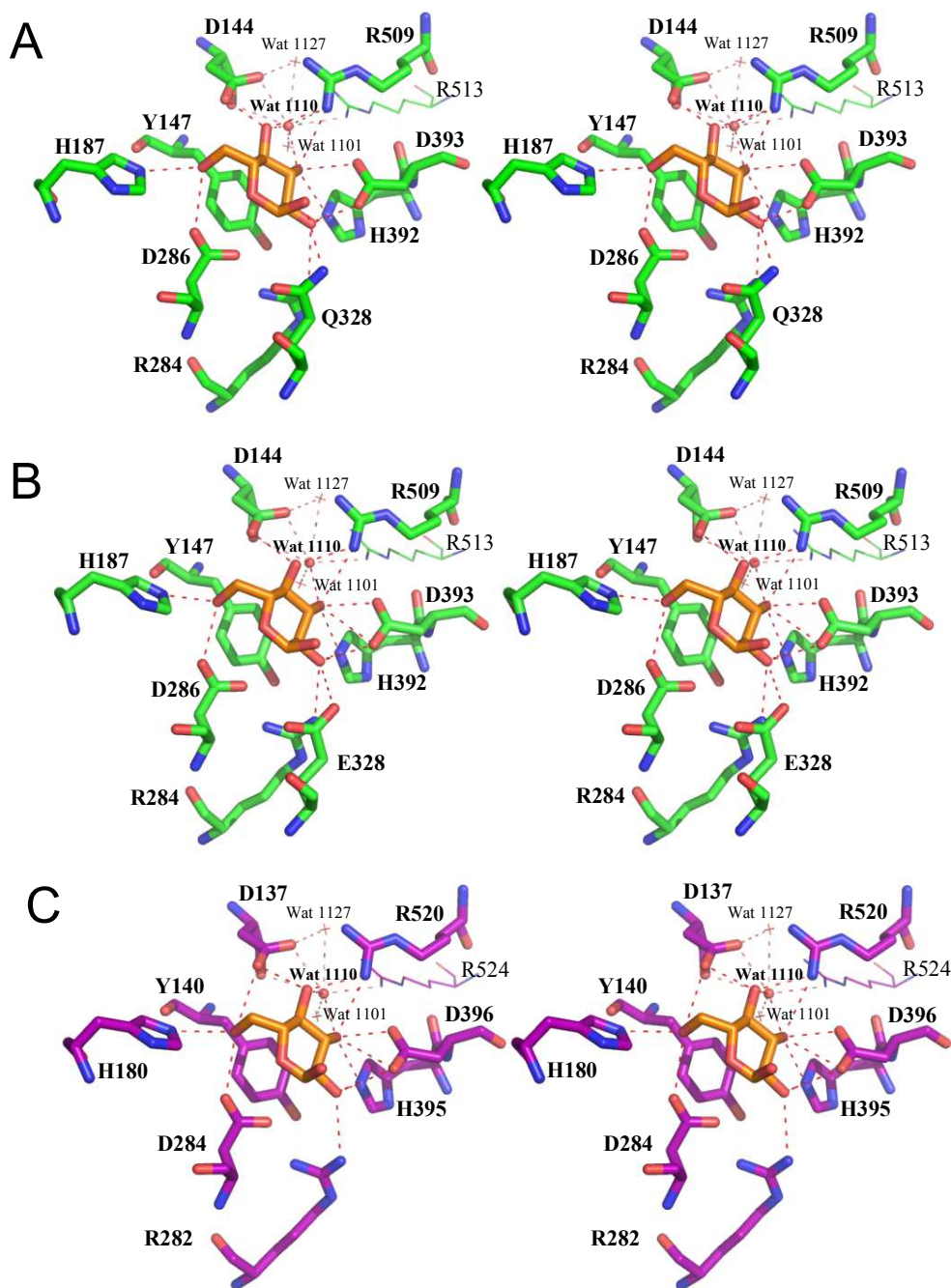
Supplementary Figure S3: Sequence alignment of amylosucrases

Amylosucrase from *Deinococcus geothermalis* (DgAS), *Deinococcus radiodurans* (DrAS), *Deinococcus maricopensis* (DmAS), *Deinococcus deserti* (DdAS), α -amylases from *Meiothermus silvanus* (MsAA), *Meiothermus ruber* (MrAA), *Truepera radiovictrix* (TrAA), amylosucrases from *Neisseria polysaccharea* (NpAS), *Neisseria meningitidis* (NmeAS), *Neisseria subflava* (NsuAS), *Neisseria mucosa* (NmuAS), *Neisseria flavescens* (NfAS), *Neisseria macacae* (NmaAS), *Neisseria sicca* (NsiAS), *Alteromonas macleodii* (AmAS), *Alteromonas addita* (AaAS), and α -amylases from *Pseudoalteromonas atlantica* (PaAA), *Glaciecola agarilytica* (GaAA) and the dimerization site pattern (pattern). Names in red are characterized amylosucrases. Catalytic triade is shown with red triangles and residues involved in the active site pocket salt bridge by blue triangles. Based on this sequence alignment, two groups can be distinguished: Group 1 (sequences 1 to 7) containing the 7 conserved regions related to dimerization of DgAS, Group 2 (sequences 8 to 18) which does not show the dimerization pattern, and thus might be considered as putative monomers like NpAS.



Group 1

Group 2



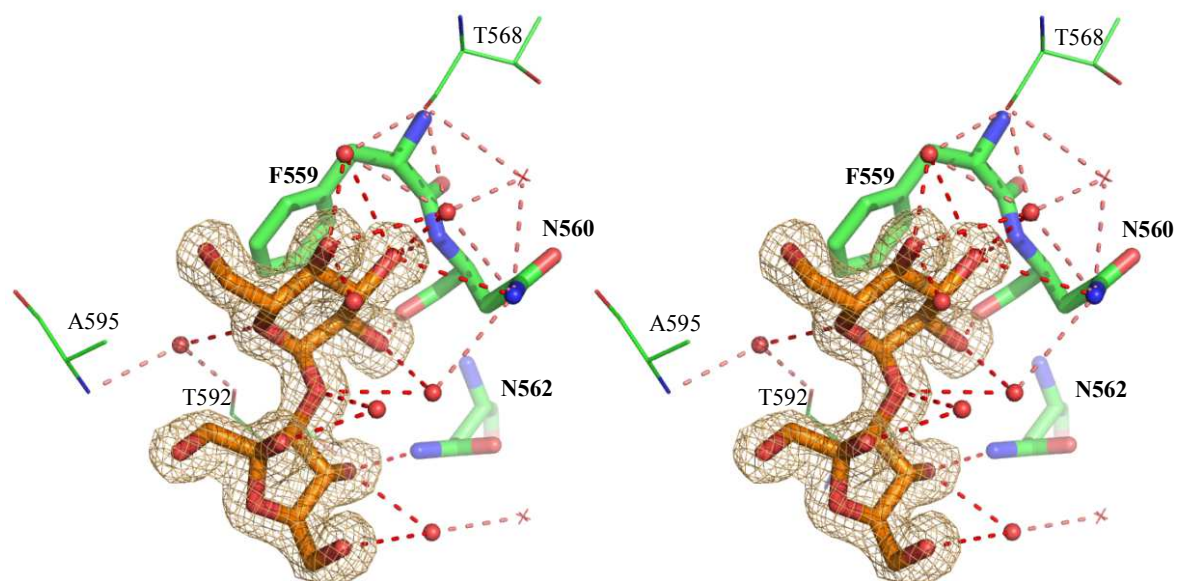
Supplementary Figure S4: Stereoview of the glucosyl moiety of sucrose or turanose

Panel A. Glucosyl moiety conformation in NpAS sucrose 1JGI complex (green).

Panel B. Glucosyl moiety conformation in NpAS turanose 3UEQ complex (green).

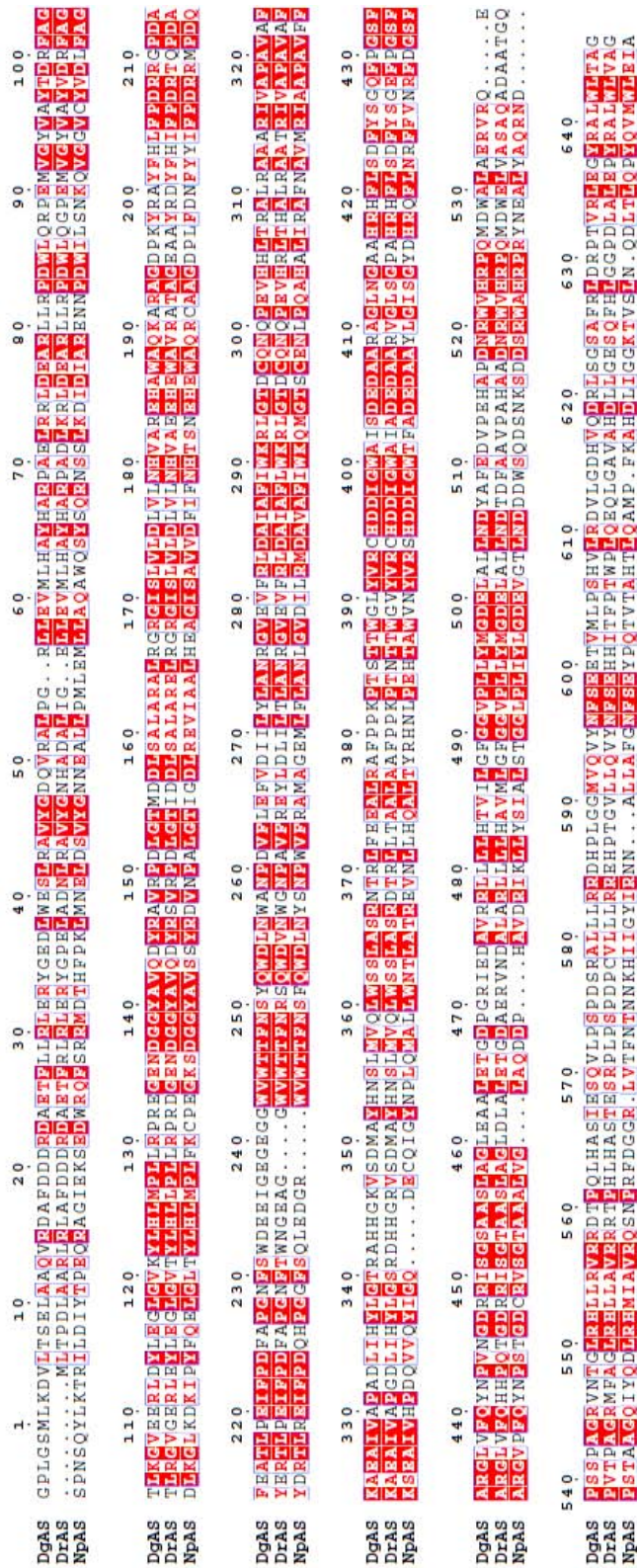
Panel C. Glucosyl moiety conformation in DgAS turanose 3UER complex (purple).

Hydrogen-bonding interactions are shown as red dotted lines and water molecules as red spheres.



Supplementary Figure S5: Stereoview of TB2 turanose binding site at the surface of the NpAS domain C

PDB code: 3UEQ with a sigmaA weighted $2F_o - F_c$ electron density map contoured at 1.0σ around the ligand. A stacking interaction between F559 and the glucosyl moiety from turanose is observed together with weak interactions between N562 and N560 from NpAS with O4' and O2/O3 from turanose, respectively.



Supplementary Figure S6: Sequence alignment of DgAS, DrAS and NpAS

DgAS: Amylosucrase from *Deinococcus geothermalis*
 DrAS: Amylosucrase from *Deinococcus radiodurans*
 NpAS: Amylosucrase from *Neisseria polysaccharea*

References

1. Henrissat, B. (1991) A classification of glycosyl hydrolases based on amino acid sequence similarities. *Biochem. J.* **280**, 309–316
2. Henrissat, B., and Davies, G. (1997) Structural and sequence-based classification of glycoside hydrolases. *Curr. Opin. Struct. Biol.* **7**, 637–644
3. 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**, D233–238
4. Potocki de Montalk, G., Remaud-Siméon, M., Willemot, R. M., Sarçabal, P., Planchot, V., and Monsan, P. (2000) Amylosucrase from *Neisseria polysaccharea*. Novel catalytic properties. *FEBS Lett.* **471**, 219–223
5. Okada, G., and Hehre, E. J. (1974) New studies on amylosucrase, a bacterial α -D-glucosylase that directly converts sucrose to a glycogen-like α -glucan. *J. Biol. Chem.* **249**, 126–135
6. Goulter, K. C., Hashimi, S. M., and Birch, R. G. (2012) Microbial sucrose isomerases. Producing organisms, genes, and enzymes. *Enzyme Microb. Technol.* **50**, 57–64
7. Ravaud, S., Robert, X., Watzlawick, H., Haser, R., Mattes, R., and Aghajari, N. (2007) Trehalulose synthase native and carbohydrate-complexed structures provide insights into sucrose isomerization. *J. Biol. Chem.* **282**, 28126–28136
8. Ravaud, S., Robert, X., Watzlawick, H., Haser, R., Mattes, R., and Aghajari, N. (2009) Structural determinants of product specificity of sucrose isomerases. *FEBS Lett.* **583**, 1964–1968
9. Thompson, J., and Pikis, A. (2012) Metabolism of sugars by genetically diverse species of oral *Leptotrichia*. *Mol. Oral Microbiol.* **27**, 34–44
10. Seo, D. H., Jung, J.-H., Jung, Ha, S.-J., Song, M.-C., Cha, J., Yoo, S. H., Kim, T.-J., Baek, N. I., and Park, C.-S. (2009) Highly selective biotransformation of arbutin to arbutin- α -glucoside using amylosucrase from *Deinococcus geothermalis* DSM 11300. *J. Mol. Catal. B* **60**, 113–118
11. 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
12. Chang, A., Singh, S., Phillips, G. N., Jr., and Thorson, J. S. (2011) Glycosyltransferase structural biology and its role in the design of catalysts for glycosylation. *Curr. Opin. Biotechnol.* **22**, 800–808
13. Champion, E., André, I., Moulis, C., Boutet, J., Descroix, K., Morel, S., Monsan, P., Mulard, L. A., and Remaud-Siméon, M. (2009) Design of α -transglucosidases of controlled specificity for programmed chemoenzymatic synthesis of antigenic oligosaccharides. *J. Am. Chem. Soc.* **131**, 7379–7389
14. Pizzut-Serin, S., Potocki-Véronèse, G., van der Veen, B. A., Albenne, C., Monsan, P., and Remaud-Siméon, M. (2005) Characterization of a novel amylosucrase from *Deinococcus radiodurans*. *FEBS Lett.* **579**, 1405–1410
15. Emond, S., Mondeil, S., Jaziri, K., André, I., Monsan, P., Remaud-Siméon, M., and Potocki-Véronèse, G. (2008) Cloning, purification and characterization of a thermostable amylosucrase from *Deinococcus geothermalis*. *FEMS Microbiol. Lett.* **285**, 25–32
16. Ha, S. J., Seo, D. H., Jung, J. H., Cha, J., Kim, T. J., Kim, Y. W., and Park, C. S. (2009) Molecular cloning and functional expression of a new amylosucrase from *Alteromonas macleodii*. *Biosci. Biotechnol. Biochem.* **73**, 1505–1512

17. Skov, L. K., Mirza, O., Henriksen, A., Potocki de Montalk, G., Remaud-Siméon, M., Sarçabal, P., Willemot, R. M., Monsan, P., and Gajhede, M. (2000) Crystallization and preliminary X-ray studies of recombinant amylosucrase from *Neisseria polysaccharea*. *Acta Crystallogr. D* **56**, 203–205
18. Skov, L. K., Mirza, O., Henriksen, A., De Montalk, G. P., Remaud-Siméon, 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
19. Skov, L. K., Mirza, O., Sprogøe, D., Dar, I., Remaud-Simeon, M., Albenne, C., Monsan, P., and Gajhede, M. (2002) Oligosaccharide and sucrose complexes of amylosucrase. Structural implications for the polymerase activity. *J. Biol. Chem.* **277**, 47741–47747
20. Jensen, M. H., Mirza, O., Albenne, C., Remaud-Siméon, M., Monsan, P., Gajhede, M., and Skov, L. K. (2004) Crystal structure of the covalent intermediate of amylosucrase from *Neisseria polysaccharea*. *Biochemistry* **43**, 3104–3110
21. Emond, S., André, I., Jaziri, K., Potocki-Véronèse, G., Mondon, P., Bouayadi, K., Kharrat, H., Monsan, P., and Remaud-Siméon, M. (2008) Combinatorial engineering to enhance thermostability of amylosucrase. *Protein Sci.* **17**, 967–976
22. Jung, J. H., Seo, D. H., Ha, S. J., Song, M. C., Cha, J., Yoo, S. H., Kim, T. J., Baek, N. I., Baik, M. Y., and Park, C. S. (2009) Enzymatic synthesis of salicin glycosides through transglycosylation catalyzed by amylosucrases from *Deinococcus geothermalis* and *Neisseria polysaccharea*. *Carbohydr. Res.* **344**, 1612–1619
23. De Montalk, G. P., Remaud-Siméon, M., Willemot, R. M., Planchot, V., and Monsan, P. (1999) Sequence analysis of the gene encoding amylosucrase from *Neisseria polysaccharea* and characterization of the recombinant enzyme. *J. Bacteriol.* **181**, 375–381
24. Battye, T. G., Kontogiannis, L., Johnson, O., Powell, H. R., and Leslie, A. G. (2011) iMOSFLM, a new graphical interface for diffraction-image processing with MOSFLM. *Acta Crystallogr. D* **67**, 271–281
25. Evans, P. (2006) Scaling and assessment of data quality. *Acta Crystallogr. D* **62**, 72–82
26. Collaborative Computational Project, Number 4 (1994) The CCP4 suite. Programs for protein crystallography. *Acta Crystallogr. D* **50**, 760–763
27. Potterton, E., Briggs, P., Turkenburg, M., and Dodson, E. (2003) A graphical user interface to the CCP4 program suite. *Acta Crystallogr. D* **59**, 1131–1137
28. McCoy, A. J. (2007) Solving structures of protein complexes by molecular replacement with Phaser. *Acta Crystallogr. D* **63**, 32–41
29. Murshudov, G. N., Vagin, A. A., and Dodson, E. J. (1997) Refinement of macromolecular structures by the maximum-likelihood method. *Acta Crystallogr. D* **53**, 240–255
30. Emsley, P., and Cowtan, K. (2004) COOT, model-building tools for molecular graphics. *Acta Crystallogr. D* **60**, 2126–2132
31. Duan, Y., Wu, C., Chowdhury, S., Lee, M. C., Xiong, G., Zhang, W., Yang, R., Cieplak, P., Luo, R., Lee, T., Caldwell, J., Wang, J., and Kollman, P. (2003) A point-charge force field for molecular mechanics simulations of proteins based on condensed-phase quantum mechanical calculations. *J. Comput. Chem.* **24**, 1999–2012
32. Lee, M. C., and Duan, Y. (2004) Distinguish protein decoys by using a scoring function based on a new AMBER force field, short molecular dynamics simulations, and the generalized born solvent model. *Proteins* **55**, 620–634

33. Case, D. A., Darden, T. E., Cheatham, I. T. E., Simmerling, C. L., Wang, J., Duke, R. E., Luo, R., Merz, K. M., Pearlman, D. A., Crowley, M., Walker, R. C., Zhang, W., Wang, B., Hayik, S., Roitberg, A., Seabra, G., Wong, K. F., Paesani, F., Wu, X., Brozell, S., Tsui, V., Gohlke, H., Yang, L., Tan, C., Mongan, J., Hornak, V., Cui, G., Beroza, P., Mathews, D. H., Schafmeister, C., Ross, W. S., and Kollman, P. A. (2006) *AMBER 9*, University of California, San Francisco, CA
34. Pastor, R. W., Brooks, B. R., and Szabo, A. (1988) An analysis of the accuracy of Langevin and molecular dynamics algorithms. *Mol. Phys.* **65**, 1409–1419
35. Berendsen, H. J. C., Postma, J. P. M., van Gunsteren, W. F., DiNola, A., and Haak, J. R. (1984) Molecular dynamics with coupling to an external bath. *J. Chem. Phys.* **81**, 3684–3690
36. Essmann, U., Perera, L., Berkowitz, M. L., Darden, T., Lee, H., and Pedersen, L. G. (1995) A smooth particle mesh Ewald method. *J. Chem. Phys.* **103**, 8577–8593
37. Ryckaert, J.-P., Ciccotti, G., and Berendsen, H. J. C. (1977) Numerical integration of the Cartesian equations of motion of a system with constraints: molecular dynamics of n-alkanes. *J. Comput. Phys.* **23**, 327–341
38. Larkin, M. A., Blackshields, G., Brown, N. P., Chenna, R., McGettigan, P. A., McWilliam, H., Valentin, F., Wallace, I. M., Wilm, A., Lopez, R., Thompson, J. D., Gibson, T. J., and Higgins, D. G. (2007) Clustal W and Clustal X version 2.0. *Bioinformatics* **23**, 2947–2948
39. Gouet, P., Robert, X., and Courcelle, E. (2003) ESPript/ENDscript. Extracting and rendering sequence and 3D information from atomic structures of proteins. *Nucleic Acids Res.* **31**, 3320–3323
40. Guex, N., and Peitsch, M. C. (1997) SWISS-MODEL and the Swiss-Pdb-Viewer. An environment for comparative protein modeling. *Electrophoresis* **18**, 2714–2723
41. Krissinel, E., and Henrick, K. (2004) Secondary-structure matching (SSM), a new tool for fast protein structure alignment in three dimensions. *Acta Crystallogr. D* **60**, 2256–2268
42. Sarçabal, P., Remaud-Siméon, M., Willemot, R., Potocki de Montalk, G., Svensson, B., and Monsan, P. (2000) Identification of key amino acid residues in *Neisseria polysaccharea* amylosucrase. *FEBS Lett.* **474**, 33–37
43. Krissinel, E., and Henrick, K. (2007) Inference of macromolecular assemblies from crystalline state. *J. Mol. Biol.* **372**, 774–797
44. Fritzsche, H. B., Schwede, T., and Schulz, G. E. (2003) Covalent and three-dimensional structure of the cyclodextrinase from *Flavobacterium* sp. No.92. *Eur. J. Biochem.* **270**, 2332–2341
45. Kim, J. S., Cha, S. S., Kim, H. J., Kim, T. J., Ha, N. C., Oh, S. T., Cho, H. S., Cho, M. J., Kim, M. J., Lee, H. S., Kim, J. W., Choi, K. Y., Park, K. H., and Oh, B. H. (1999) Crystal structure of a maltogenic amylase provides insights into a catalytic versatility. *J. Biol. Chem.* **274**, 26279–26286
46. Hondoh, H., Kuriki, T., and Matsuura, Y. (2003) Three-dimensional structure and substrate binding of *Bacillus stearothermophilus* neopullulanase. *J. Mol. Biol.* **326**, 177–188
47. Jaenicke, R., and Böhm, G. (1998) The stability of proteins in extreme environments. *Curr. Opin. Struct. Biol.* **8**, 738–748
48. Prakash, O., and Jaiswal, N. (2010) α -Amylase, an ideal representative of thermostable enzymes. *Appl. Biochem. Biotechnol.* **160**, 2401–2414

49. Zhang, D., Li, N., Swaminathan, K., and Zhang, L. H. (2003) A motif rich in charged residues determines product specificity in isomaltulose synthase. *FEBS Lett.* **534**, 151–155
50. Lichtenthaler, F. W., and Rönninger, S. (1990) Studies on ketoses, 4. -D-glucopyranosyl-D-fructoses: distribution of furanoid and pyranoid tautomers in water, dimethyl sulphoxide, and pyridine. *J. Chem. Soc. Perkin Trans.* **2**, 1489–1497

*Chapitre 3 : Mutants de
l'amylosaccharase de *Neisseria
polysaccharea* adaptés à la
glucosylation d'accepteurs non
naturels*

Au cours de précédents travaux d'ingénierie réalisés par Champion et coll. (10), l'amylosaccharase de *Neisseria polysaccharea* (ASNp) a été adaptée à la reconnaissance de précurseurs osidiques non naturels intervenant dans la synthèse d'oligosaccharides complexes. En particulier, le mutant F290K a révélé des propriétés remarquables pour la synthèse de l' α -D-glucopyranosyl-(1 \rightarrow 4)- α -allyl-N-acétyl-glucosamine qui est améliorée d'un facteur \sim 130 par rapport à l'enzyme sauvage. Cette nouvelle transglucosidase a été utilisée dans le cadre du projet ANR OptiGluc afin de proposer de nouvelles voies de synthèse chimio-enzymatique d'antigènes saccharidiques.

Toutefois, les performances catalytiques de ce mono-mutant n'étant pas optimales pour avoir un procédé de synthèse chimio-enzymatique efficace, nous avons cherché à améliorer d'avantage la réaction de glucosylation de l'accepteur cible (α -allyl-N-acétyl-glucosamine nommé D' par la suite). Ainsi, la première partie de nos travaux a consisté à construire des banques de double-mutants de l'ASNp sur les positions 228-290, 289-290, 228-229. Le développement d'une technique de criblage basée sur la glucosylation de l'accepteurs D' a permis d'isoler trois double-mutants : A289P-F290L, A289P-F290C et A289P-F290I dont l'efficacité catalytique pour la monoglucosylation régiospécifique de l'accepteur D' a été significativement améliorée d'un facteur \sim 400.

La seconde partie de ce travail a consisté à comprendre les origines moléculaires de ces améliorations d'efficacité catalytique. Pour cela, nous avons entrepris différentes études biophysiques, à commencer par la résolution de la structure cristallographique des différents mutants d'intérêt. Ainsi, les structures des trois doubles mutants A289P-F290L, A289P-F290C et A289P-F290I ont été déterminées à 2,4 ; 2,2 et 2,5 Å de résolution respectivement. La cristallisation du monomutant F290K, précédemment isolé par Champion et coll. (10), a été réalisée avec la protéine sous forme inactive (par introduction de la mutation E328Q, résidu catalytique jouant le rôle acide/base) et en complexe avec le saccharose. Sa structure a été résolue à 2,3 Å de résolution.

De façon générale, les changements opérés par ces mutations sur la structure se sont avérés mineurs, ne permettant pas d'expliquer les gains observés en efficacité catalytique. Afin de caractériser plus en détail la stabilité de ces mutants, nous avons déterminé les températures de dénaturation, en présence ou en absence de saccharose, par fluorimétrie différentielle à balayage ce qui a permis de constater des pertes significatives de T_m des différents mutants par rapport à l'enzyme sauvage allant jusqu'à 6 °C.

Enfin, une étude par dynamique moléculaire a permis de mettre en évidence des changements considérables dans la flexibilité de ces néo-enzymes par rapport à l'ASNp sauvage, notamment dans des régions clefs comme les boucles 3 et 7 appartenant aux domaines B et B', respectivement, et recouvrant le site actif.

L'ingénierie de neo-enzymes couplée à une étude des relations structure-dynamique-activité, nous a permis de mettre en évidence, dans le cas de l'ASNp, que les mutations introduites, bien qu'ayant un effet spectaculaire sur l'efficacité catalytique, n'affectent la structure que de manière très localisée, jouant essentiellement sur le volume de la poche catalytique. En revanche, ces mutations introduiraient une déstabilisation importante des mutants améliorés de l'ASNp qui est également accompagnée d'une altération de la flexibilité des mutants par rapport à l'ASNp sauvage.

Ces travaux font l'objet d'une publication en cours de préparation pour une soumission à *The Journal of Biological Chemistry*.

Applying pairwise combinations of amino acid mutations for sorting out highly efficient glucosylation tools for chemo-enzymatic synthesis of bacterial oligosaccharides

F. Guérin[§], E. champion[§], C. Moulis, T.H. Tran, S. Barbe, S. Morel, K. Descroix, Pierre Monsan, L. Mourey, L.A. Mulard, S. Tranier, M. Remaud-Siméon and I. André

[§] These authors have equally contributed to the work.

Summary

Iterative saturation mutagenesis and combinatorial active site saturation focused on vicinal amino acids were used to alter the acceptor specificity of amylosucrase, a sucrose-utilizing α -transglucosidase, and sort out improved variants. From the screening of three semi-rational sub-libraries accounting in total for 20,000 variants, we report here the isolation of three double-mutants displaying a spectacular catalytic efficiency enhancement for α -allyl-*N*-acetyl-2-deoxy- α -D-glucopyranoside glucosylation substrate biotransformation. Such levels of activity improvement have never been reported before for this class of carbohydrate-active enzymes. X-ray structural analysis of the best performing enzymes followed by Molecular Dynamics (MD) simulations showed both local rigidity of the -1 subsite and flexibility of loops involved in active site topology which both account for the enhanced catalytic performances of the mutants. The study well illustrates the importance when designing enzyme libraries of taking into account the local conformation of catalytic residues as well as protein dynamics during the catalytic process.

Introduction

Shigella, the causal agent of shigellosis, or bacillary dysentery, represents a major burden worldwide (1). In countries where shigellosis is endemic, at least fourteen distinct *Shigella flexneri* serotypes have been isolated in addition to *Shigella sonnei*, emphasizing the need for a vaccine with broad coverage (2). The bacterial surface polysaccharides are seen as major targets of the host's protective immunity against infection, and as such as key components of *Shigella* vaccines. This observation led to the development of, among others, a number of *S. flexneri* carbohydrate conjugate vaccine candidates, either LPS-based (3,4), or synthetic oligosaccharide-based (5-7). In view of the need for a multivalent *Shigella* vaccine,

the development of the latter approach requires a straightforward access to well-defined O-antigen (O-Ag) fragments of the prevalent *S. flexneri*. Interestingly, most *S. flexneri* surface polysaccharides share the same linear backbone, but differ in terms of α -D-glucosyl and/or O-acetyl decorations (2,8). As an extension to the pure chemical synthesis (5), such structural diversity of *S. flexneri* O-Ags has led us to investigate, in recent years, the construction of a synthetic toolbox (9-11), involving both chemical and enzymatic processes, to access chemically defined oligosaccharide fragments of selected *S. flexneri* O-Ags. In spite of accomplished advancements, there is still a crucial need for appropriate biocatalysts with requisite efficiency, specificity, and stability to enable critical steps in chemical synthesis. Natural availability of “ideal enzymes” able to perform the desired transformation is often critical. However, with the progress in biomolecular and bioinformatics technologies, protein engineering techniques have contributed to extend the repertoire of available biocatalysts that could be useful in organic synthesis (12,13).

As a direct continuation of the strategy earlier reported (10), the current work aims at developing novel chemo-enzymatic routes to oligosaccharide fragments of some *S. flexneri* O-Ags and overcome the limitations of the chemical 1,2-*cis* glucosylation developed so far. Previous work already enabled the design of α -transglucosidases with altered specificity towards acceptor substrates that were used for the synthesis of glucosylated building blocks properly protected for their subsequent chemical conversion into fragments of the *S. flexneri* 3a and 1b O-Ags. The use of α -transglucosidases, named glucansucrases (14), was favoured over Leloir-type glucosyltransferases utilizing nucleotide-activated sugar substrates as glucosyl unit donor. Glucansucrases, which belong to Glycoside-Hydrolase (GH) families 13 and 70 (15,16), naturally proceed by successive transfers of α -D-glucopyranosyl units from cheap and abundant sucrose to the synthesis of α -glucans linked through distinct osidic linkages depending on the enzyme regiospecificity (Figure 33). One major advantage of these enzymes is their natural promiscuity for a variety of non-natural acceptors that opens access to gluco-derivatives (11). Focusing on *N. polysaccharea* amylosucrase (NpAS) which belongs to GH13 family, a structurally guided engineering approach enabled us to identify, from a small library of only 133 single-mutants, new catalysts such as the F290K mutant that gained in specificity towards a non-natural acceptor, the allyl-2-acetamido-2-deoxy- α -D-glucopyranoside (**D'**). The glucosylation products were shown to serve as a precursor in a novel chemo-enzymatic synthesis of complex oligosaccharides (10). However, the observed

improvement of enzyme specificity was also accompanied by a decrease of catalytic efficiency compared to that of the wild type enzyme acting on its natural substrate.

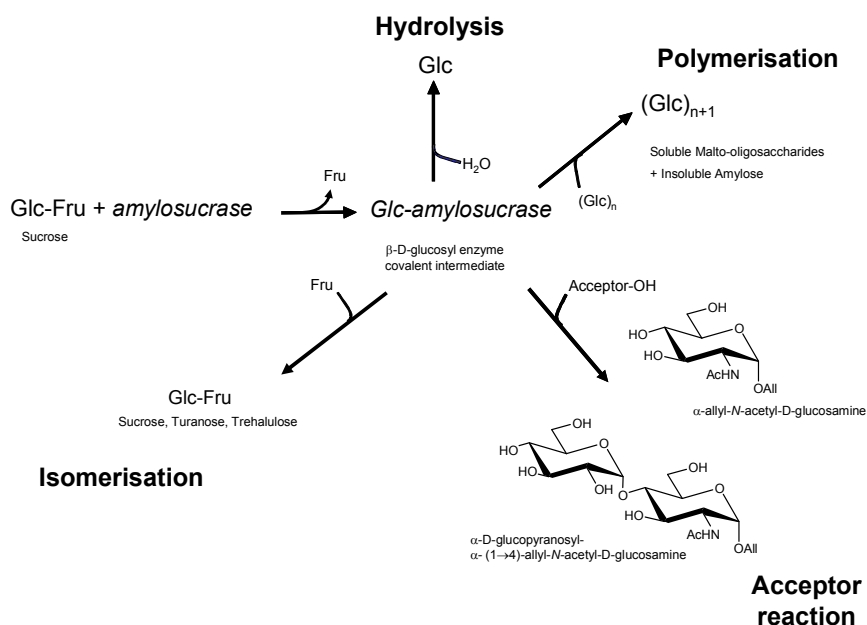


Figure 33: Reactions catalyzed by amylosucrases from sole sucrose

Glc: Glucose; Fru: Fructose; Sucrose: α -D-glucopyranosyl-1,2- β -D-fructofuranoside; Turanose: α -D-glucopyranosyl-1,3- β -D-fructose; Trehalulose: α -D-glucopyranosyl-1,1- β -D-fructose, all: allyl-.

Advances in enzyme engineering have demonstrated the importance of the quality of the variant library and, in particular, the efficiency of iterative saturation mutagenesis and combinatorial active site mutagenesis (17,18) to change enzyme specificity and sort out variants with desired properties from libraries of reasonable size. Here, we have investigated the value of such an approach for carbohydrate-active enzyme engineering. In particular, we focused part of our library design onto combinatorial tandem mutations targeting adjacent amino acids. Indeed, simultaneous mutations onto vicinal amino acid residues are not frequently observed in nature. In contrast, they can be easily performed *in vitro* once key positions are identified and could be of interest for a fine tuning of unnatural substrate. The ultimate goal of this contribution remains the enzymatic synthesis of the α -D-Glcp-(1 \rightarrow 4)- α -D-GlcpNAc-OAll (**ED'**), a lightly protected disaccharide not found in nature, that was shown to serve in a chemo-enzymatic synthesis of *S. flexneri* 1b specific oligosaccharides (10). Key positions for **D'** recognition had been previously sorted out from a single-mutant library targeting 7 positions (10). Here we examined the effect of the close environment of these

amino acids onto the substrate specificity and enzyme efficiency. To achieve this, combinatorial libraries of tandem mutations were generated and screened. Remarkable improvement of mutant properties was obtained from only one round of screening. An X-ray structural analysis of the best performing enzymes completed by Molecular Dynamics (MD) simulations was carried out to supply new insights into structural and dynamical determinants involved in the enhanced catalytic performances of mutants. The efficiency of the engineering approach used to rapidly sort out improved mutants and enrich our glucosylation toolbox with efficient catalysts is discussed.

Experimental procedures

Bacterial strains, plasmids and chemicals. Plasmid pGST-AS G537D, derived from the pGEX-6P-3 (GE Healthcare Biosciences) and containing the *N. polysaccharea* amylosucrase encoding gene (19) was used for the amylosucrase libraries construction. *E. coli* TOP 10 electrocompetent cells (Invitrogen, Carlsbad, USA) were used as host for each plasmid library transformation, gene expression, and production of the selected mutants. Fusion DNA-polymerase was purchased from Finnzymes (Espoo, Finland), and *DpnI* restriction enzyme from New England Biolabs (Beverly, MA, USA). Oligonucleotides were synthesized by Eurogenetec (Liege, Belgium). DNA extraction (QIASpin) and purification (QIAQuick) columns were purchased from Qiagen (Chatsworth, CA). DNA sequencing was performed by Beckman Coulter Genomics (Grenoble, France). All positive clones for **D** glucosylation were sequenced on the mutated region (~600 bp) using the primer pGEX_int: CCAACGAACACGAATGGGC.

Ampicillin (Amp), lysozyme, and isopropyl- β -D-thiogalactopyranoside (IPTG) were purchased from Euromedex (Souffelweyersheim, France); Bromothymol Blue sodium salt, sucrose and *N*-acetyl-D-glucosamine (D-GlcpNAc) from Sigma-Aldrich (Saint-Louis, MO, USA). **D'** was synthesized using known protocols (20). Reference disaccharide α -D-Glcp-(1 \rightarrow 4)-D-GlcpNAc was enzymatically synthesized and characterized at LISBP, Université de Toulouse, France (10).

Construction, expression and screening of mutant libraries. Amylosucrase libraries focused on positions 289-290 or 228-229 (libraries 2 and 3), were constructed by site-saturation mutagenesis using pGST-AS G537D as vector template. It was previously checked

that G537D mutation had no impact on the native enzyme catalytic properties (data not shown). Three partial overlapping primer pairs surrounding double codons were designed. Each of these codons was replaced with degenerate NNS or NNW sequences, where N= A, C, G or T; S= C or G and W=A or T (Supplementary Table S3). Such degenerate primers were designed to generate 32 codons encoding the 20 possible amino acids. PCR amplifications were carried out on the whole plasmid with 1 U of Phusion DNA-polymerase for 30 cycles (98 °C, 10 s; 75 °C, 20 s; 57 °C, 15 s; 72 °C, 5 min). The DNA was digested with *DpnI* to eliminate methylated parental template and purified using Qiaquick spin column, following manufacturer's recommendations. *E. coli* TOP10 was transformed by electroporation with 4 µL of each plasmid library using standard procedures.

Library **1** combining mutations at positions 228 and 290 was constructed by saturation mutagenesis using a two-step procedure described in (21).

After *E. coli* transformation, clones producing active amylosucrases were detected using a pH-based high throughput screening assay on solid LB medium containing sucrose (21). Freshly transformed cells were plated on membranes (Durapore membrane filters, 0,22 µm GV-Millipore, Ireland) previously soaked in physiological water and placed onto square plates (Corning, USA) containing LB agar supplemented with ampicillin (100 µg.mL⁻¹) and glycerol (1 % w/v). After overnight growth at 37 °C, each membrane was transferred onto another square plate containing inducing medium (LB agar, ampicillin and IPTG (1 mM), sucrose (50 g.L⁻¹) and stained blue by addition of 50 mM Tris-HCl pH 7.5 with Bromothymol blue indicator (0.1 g.L⁻¹ in 1 % ethanol). After 24 h of incubation at 30 °C, active clones (green and yellow) were picked and cultured in 96-well microplates containing LB medium, ampicillin and glycerol (12 % w/v), before storage (-80 °C) or further investigation. The high-throughput screening work was carried out at the Laboratoire d'Ingénierie des Systèmes Biologiques et des Procédés (Toulouse, France) with the equipments of the ICEO facility, dedicated to the screening and screening of new and original enzymes.

Variants able to glucosylate D-GlcNAc

(D)

 were screened using the HPLC procedure previously described (10,21). Briefly, microplates were duplicated into 96-Deep Well plates containing LB medium supplemented with ampicillin and IPTG (1 mM) to induce GST-AS expression. Cultures were then grown for 24 h at 30 °C under agitation (200 rpm), centrifuged (20 min, 3,700 g, 4 °C) and the supernatant was removed. The cell pellet was resuspended in 200 µL of lysozyme (0.5 mg.mL⁻¹), followed by freezing at -80 °C for 8 to 12 h. After thawing at room temperature, 100 µL of sucrose and 100 µL of acceptor (each at a final

concentration of 73 mM) were added to each well. Enzymatic reaction was incubated at 30 °C during 24 h under agitation. The DeepWell plates were then centrifuged (20 min, 3,700 g, 4 °C) and 300 µL of the supernatant was transferred to a filter micro-plate (Glass fiber membrane, PS, 0.25 mm pore- Corning, USA) to be clarified. Supernatant filtration was carried out by centrifugation of the filter microplate (5 min, 2,000 g, 4 °C) into a microplate for HPLC screening. Acceptor reaction products were analyzed using a C18-AQ column (Bischoff C18, 125x4 mm, 3 µm) kept at room temperature and eluted with 0.6 mL.min⁻¹ of ultra-pure water to detect **ED** formation (analysis time: 7 min). Complementary HPLC analyses with a Biorad HPX-87K Carbohydrate Analysis column (maintained at 65 °C, and eluted at a flow rate of 0.6 mL.min⁻¹ with ultra-pure water) allowed to measure sucrose consumption by RI detection and concomitant **ED** formation by UV220nm. The mutant ability to synthesize the desired disaccharide **ED** was estimated as follows: % Glucosyl units transferred onto acceptor derivatives = [Glucosyl units transferred onto acceptor derivatives]/[Glucosyl units transferred from initial sucrose].

Production, purification and characterization of improved variants. Prior to production and purification of double-mutants, their entire genes were sequenced by Cogenics (Grenoble, France) showing no other mutations. The A289P-F290C, A289P-F290I and A289P-F290L double-mutants were produced, and purified to the GST-AS fusion protein stage (96 kDa) as previously described (19). The protein content was determined using a Nanodrop ND-1000 spectrophotometer. Specific activities were determined at 30 °C in 50 mM Tris-HCl buffer, pH 7.0 using 250 mM sucrose alone, or with 250 mM sucrose supplemented with 250 mM of **D'** acceptor. The concentration of reducing sugars was determined using the dinitrosalicylic acid assay (22). One unit of amylosucrase activity corresponds to the amount of enzyme that catalyzes the release of 1 µmol of reducing sugars per minute in the assay conditions.

Comparison of products synthesized by wild-type amylosucrase and variants. Reactions were performed at 30 °C, in the presence of 250 mM sucrose alone or supplemented with 250 mM **D'** acceptor and 1 U/mL of wtAS or variants. The reactions were stopped after 24 h by heating at 95 °C for 5 min. The soluble part of the reaction mixture was submitted to HPAEC (high-performance anion-exchange chromatography with pulsed amperometric detection) analysis. To quantify the concentration of monosaccharides and

disaccharides, the soluble fraction was diluted in water and separated on a 4 x 250 mm Dionex Carbo-pack PA100 column. A gradient of sodium acetate (from 6 to 300 mM in 28 min) in 150 mM NaOH was applied at 1 mL.min⁻¹ flow rate. Detection was performed using a Dionex ED40 module with a gold working electrode and an Ag/AgCl pH reference. Note that acceptor **D'**, and their derivatives, are not oxidable products and thus are not detectable by HPAEC. Therefore, **D'** and their glucosylation products, were quantified by HPLC with a Biorad HPX-87K (see above). Concentration of sucrose, glucose, fructose, turanose, trehalulose and maltose was determined by HPAEC.

In parallel, the reaction mixture containing soluble and insoluble malto-oligosaccharides was solubilized in 1 M aq KOH at a final total sugar concentration of 10 g.L⁻¹ and analyzed by HPAEC using a Dionex Carbo-Pack PA100 column at 30 °C. Mobile phase (150 mM aq NaOH) was set at 1 mL.min⁻¹ flow rate with a sodium acetate gradient (6 to 500 Mm over 120 min).

Determination of kinetic parameters. Catalytic efficiency (k_{cat}/K_m) of wtAS and AS variants was determined with (i) sucrose alone and with (ii) both sucrose and acceptor **D'** as variable substrates. Enzyme assays were carried out in a total volume of 2 mL containing respectively (i) 0.073 mg of A289P-F290C, 0.153 mg of A289P-F290I or 0.092 mg A289P-F290L in presence of sucrose concentration varying from 0 to 500 mM sucrose or (ii) 0.037 mg A289P-F290C, 0.076 mg of A289P-F290I and 0.046 mg of A289P-F290L in presence of 250 mM sucrose and **D'** concentrations varying from 0 to 250 mM. At regular time intervals, aliquots (200 µL) were removed, heated (95 °C, 2 min) and centrifuged (18,000 g, 5 min). Samples were filtered on a 0.22 µm membrane and analyzed using HPLC material previously described. The initial rate of sucrose consumed, corresponding to the initial rate of fructose released, was expressed in µmole of fructose released per minute and per gram of enzyme. For reaction with sucrose only, the kinetic parameters k_{cat} and K_m were determined using the Eadie-Hofstee plot. For reaction with sucrose and acceptor, saturation was not achieved for the double-mutants and the efficiency k_{cat}/K_m (**D'**) was calculated by linear regression of the initial rate of formation of desired disaccharide (**ED'**) versus substrate concentration.

Crystallization. A289P-F290L, A289P-F290C and A289P-F290I mutants were crystallized using conditions previously described by Skov *et al.* (23,24). Crystals of F290K-

E328Q mutant were obtained using the same conditions described above and in presence of 20 mM sucrose.

Data collection and structure determination. X-ray experiments were carried out at 100 K. Prior to flash cooling, native crystals of NpAS mutants were soaked for a few seconds in the reservoir solution supplemented with 20 % (v/v) glycerol to avoid ice formation, except for the co-crystal of F290K-E328Q-sucrose which is intrinsically cryoprotected by sucrose. A289P-F290L, A289P-F290C and A289P-F290I diffraction datasets were collected to a maximum resolution of 2.40 Å, 2.20 Å and 2.50 Å, respectively on beamline ID14-1 (for A289P-F290C and A289P-F290I mutants) and ID14-2 (for A289P-F290L mutant) at the European Synchrotron Radiation Facility (ESRF, Grenoble, France). The F290K-E328Q-sucrose complex dataset was collected to 2.30 Å on the ESRF beamline ID23-1. Diffracted intensities were integrated using iMOSFLM (25) and scaled with SCALA (26) from the CCP4 software suite (27,28) and 10 % (for A289P-F290L, A289P-F290C and A289P-F290I mutants) or 5 % (for F290K-E328Q mutant in complex with sucrose) of the scaled amplitudes were randomly selected and excluded from the refinement procedure. Crystals of all amylosucrase mutants in complex or not with sucrose belong to the P₂₁2₁2 spacegroup with one molecule per asymmetric unit giving a Matthews coefficient of 2.3 Å³/Da. Data collection statistics are given in Table 4. Structures of amylosucrase mutants were **straight** refined from their native structures (PDB code: 1G5A) using refmac5 (29).

Building and refinement. Structure refinement was performed using refmac5 from the CCP4 GUI (27,28) and models were manually reconstructed in SigmaA weighted electron density maps using COOT (30). Water molecules were automatically assigned and ligand molecules were manually fitted onto residual maps. The final models of all mutants contain 628 residues out of the 632 theoretical residues with 4 missing residues at the N-terminal extremity. Refinement statistics are given in Table 4.

Table 4: Data collection and refinement statistics

	A289P-F290L	A289P-F290C	A289P-290I	F290K-E328Q
	<i>Apo</i>	<i>Apo</i>	<i>Apo</i>	Sucrose
Data collection				
Space group	P2 ₁ 2 ₁ 2	P2 ₁ 2 ₁ 2	P2 ₁ 2 ₁ 2	P2 ₁ 2 ₁ 2
a, b, c (Å)	96.0, 116.3, 60.2	96.0, 116.0, 60.5	96.5, 115.8, 60.8	95.1, 114.7, 54.6
α, β, γ (°)	90.0, 90.0, 90.0	90.0, 90.0, 90.0	90.0, 90.0, 90.0	90.0, 90.0, 90.0
Resolution (Å)	38.78-2.40 (2.53-2.40)*	38.66-2.20 (2.32-2.20)	37.06-2.50 (2.64-2.50)	57.35-2.30 (2.42-2.30)
R_{sym}	0.098 (0.337)	0.101 (0.306)	0.096 (0.310)	0.142 (0.418)
I/σI	6.8 (2.0)	7.2 (2.5)	7.1 (2.4)	5.1 (1.9)
Completeness (%)	97.9 (94.8)	99.8 (99.1)	98.3 (99.4)	98.9 (96.7)
Redundancy	3.0 (2.9)	4.7 (4.0)	2.4 (2.4)	3.3 (3.2)
Nb of molecule / AU	1			
Matthews coefficient (Å³/Da)	2.3	2.3	2.3	2.3
Refinement				
Resolution (Å)	12.00-2.40	14.00-2.20	12.00-2.50	12.00-2.30
No. of unique reflections	26,360 (3,664)	35,028 (5,007)	23,776 (3,464)	26,993 (3,768)
R_{work} / R_{free} (%)	15.84/24.16	14.77/21.67	15.49/22.65	15.43/22.29
Total number of atoms	5,427	5,679	5,419	5,367
Nb of protein residues	628	628	628	628
Nb of ligand molecules	1 Tris, 18 glycerols, 1 PEG	1 Tris, 3 glycerols, 1 PEG	1 Tris, 17 glycerols, 2 PEG	2 sucroses, 1 Cl ⁻
Nb of water molecules	258	571	335	283
B-factors (Å²)				
Protein	22.3	16.0	22.6	17.5
Ligand	33.9	29.1	35.3	18.5
Water	23.1	23.1	25.4	18.8
Rmsd				
Bond lengths (Å)	0.021	0.018	0.019	0.018
Bond angles (°)	2.2	1.8	2.1	2.0

* Values in parentheses are for the outer resolution shell

Thermostability of amylosucrases. The melting point (T_m) of amylosucrase mutants was assayed by differential scanning fluorimetry (DSF). A mix of enzyme (2 μ M), Sypro-orange (5 X) (Invitrogen, Paisley, UK) and 150 mM HEPES, pH 7.3, 150 mM NaCl, 1 mM DTT, 1 mM EDTA in absence or in presence of sucrose 100, 300 or 600 mM were incubated using a temperature gradient from 20 to 80 °C with a 0.3 °C increment. The thermal transition was monitored using a RTQ-PCR CFX96 Real-Time System (Biorad, Marnes-la-Coquette, France). T_m was given by the inflexion point of the curve RFU=f(T), with RFU standing for Relative Fluorescence Unit.

Molecular dynamics (MD) simulations. All MD simulations were carried out using the AMBER 9 suite of programs (31). The molecular all-atom ff03 (32,33) and carbohydrate GLYCAM06 (34) force fields were used for the proteins and the sucrose, respectively. The starting models were derived from the high-resolution crystal structures of NpAS (PDB code: 1G5A) (24) and its mutants in *apo*-form or in complex with the sucrose substrate. To obtain a neutral charge of the simulated systems, a number of counter-ions were included. Each enzyme or enzyme/sucrose complex together with the counter-ions was solvated with TIP3P water molecules, using the rectangular parallelepiped box with a minimum distance of 0.12 nm between the solute and the simulation box edge. Preparation of simulations consisted of initial energy minimization steps (steepest descent and conjugate gradient methods). The minimization steps were then followed by a slow heating to 303 K under constant volume over a period of 100 ps. At the final required temperature (303 K), the system was equilibrated under constant volume condition over 10 ps and then it was turned on constant pressure (1 bar) condition over 90 ps. Atomic positions of the protein backbone were first restrained using a harmonic potential during the minimization schedule. The force constant was then progressively diminished until a final unrestrained minimization step. Harmonic constraints applied on sucrose were maintained until the end of the heating phase and thereafter they were gradually removed along the equilibration in the isothermal-isobaric ensemble. The final production phase of simulations was then carried out for a total of 20 ns at constant temperature (303 K) and pressure (1 bar) conditions. The simulation temperature was controlled using Langevin (35) thermostat and Berendsen barostat (36) with a collision frequency 2 ps and pressure relaxation time as 2 ps. Long-range electrostatic forces were handled by using the particle-mesh Ewald method (37). The time step of the simulations was

2.0 fs and the SHAKE algorithm was used to constrain the lengths of all chemical bonds involving hydrogen atoms to their equilibrium values (38). The resulting trajectories were analyzed using the Ptraj module of the AMBER 9 package. The RMSD was calculated for the protein backbone atoms using least squares fitting. Atomic positional fluctuations (Δr_i^2) of protein backbone were calculated. A mass-weighted average value was then calculated for each residue. These parameters are related to the B-factors through the following relationship:

$$B_i = \frac{8\pi^2}{3} \langle \Delta r_i^2 \rangle$$

The simulated B-factors were calculated using the coordinates of the 20 ns trajectories.

Results and discussion

Generation of libraries harbouring tandem mutations

Three saturation libraries of double-mutants were generated. Positions 228, 289 and 290, which were previously identified as being favourable for **D'** recognition, were selected (10). In Libraries 1 and 2, the second round of randomization was focused on positions 228-290 and 289-290, respectively. We also opted for the construction of a library focused onto amino acids 228 and 229 to target two other adjacent positions in the active site. Position 229 was preferred to position 227 due to the presence of a bulky and hydrophobic Phe at position 227 that is likely to disturb **D'** accommodation. We used a NNS or NNW codon degeneracy in order to generate the 20 possible amino acids at each position. The three libraries were first screened using a previously developed colorimetric pH-based assay on solid medium allowing rapid determination of library size and sucrose-active clone ratio (21). Altogether, 1,140 clones were found able to utilize sucrose as a glucosyl donor, what corresponds to an active clone ratio comprised between 2 and 29 %, depending on the sub-library (Table 5). Notably, with 29 % of positive clones, library 2 combined two positions that are quite tolerant to mutations. Among active clones, 288 clones were considered as very active due to a strong pH change that induced an intense colour shift from blue to yellow, whereas 852 were found moderately active (green colour). Overall, the screening efforts based on sucrose utilization allowed downsizing by ~94 % the number of mutants to be tested for their ability to glucosylate commercial *N*-acetyl-glucosamine (D-GlcpNAc, **D**) acceptor, the detection of

glucosylation product α -D-Glcp-(1→4)-D-GlcpNAc (**ED**) being carried out by HPLC. Using this procedure, 30 mutants were found to synthesize the **ED** compound with a glucosylation yield above 20 %. Among them, 10 single-mutants had already been characterized in previous work (10). Therefore, we focused our work on the 20 remaining active amylosucrase double-mutants which are listed in Table 6. Among them, only two were found twice (I228V-F290V and A289P-F290C). Compared with the corresponding single-mutants (10), most of the double-mutants produced higher yields of **ED** disaccharide, thus demonstrating the relevance of iterative saturation mutagenesis strategy. Moreover, results obtained from libraries 2 and 3 also show that the double-mutation cycle approach focusing onto vicinal residues is efficient to rapidly identify beneficial coupling interactions that might correspond to optimal compensations of structural rearrangements and isolate superior mutants compared to single ones. Based on our screening criteria, the most promising double-mutants (A289P-F290L, A289P-F290C and A289P-F290I) emerged from library 2.

Table 5: Results of the pH screening of amylosucrase sub-libraries for sucrose utilization

	Library 1	Library 2	Library 3
	228-290	289-290	228-229
Number of clones	~ 8,000	~ 2,000	~ 8,000
Number of active clones isolated (%)	384 (5 %)	576 (29 %)	180 (2 %)
Number of active clones (yellow)	96	96	96
Number of moderately active clones (green)	288	480	84

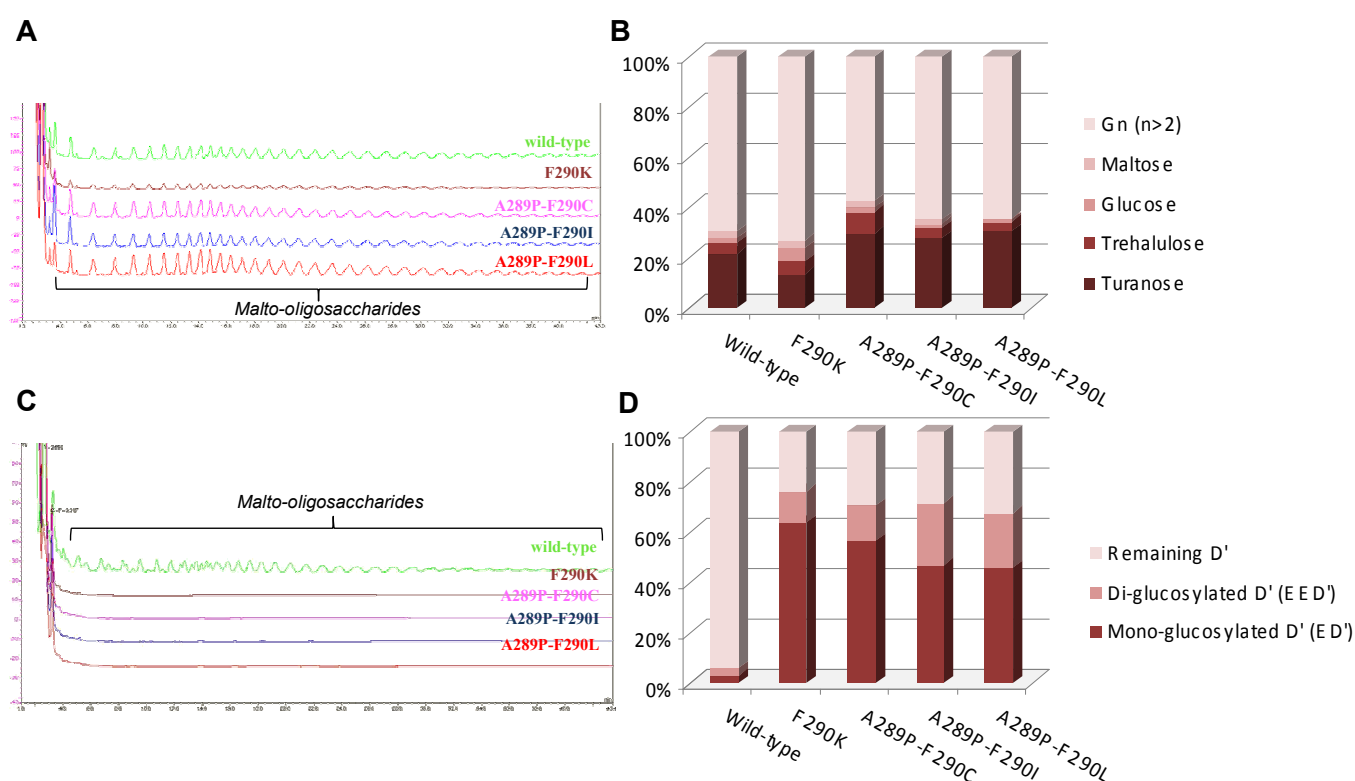
Table 6: Improved not purified amylosucrase mutants identified for α -D-Glcp-(1 \rightarrow 4)-D-GlcpNAc (ED) synthesis

Library	Amino acid mutation(s)	Mutation codon(s)	Occurrence	Glucosylation Yield to synthesize ED (%)
	Wild-type amylosucrase		1	2
1	I228V-F290V	GTC GTT	2	90
1	I228V-F290K	GTC AAA	1	89
1	I228A-F290H	GCC CAT	1	88
1	I228T-F290K	ACC AAA	1	72
1	I228V-F290H	GTC CAT	1	61
1	I228T-F290H	ACC CAT	1	56
1	I228K-F290W	AAG TGG	1	50
1	I228K-F290L	AAG CTT	1	47
1	I228V-F290R	GTC CGT	1	25
1	I228P-F290C	CCC TGT	1	22
2	A289P-F290L	CCC TTA	1	100
2	A289P-F290C	CCG TGT	2	100
2	A289P-F290I	CCC ATT	1	100
2	A289T-F290H	ACG CAT	1	40
2	A289H-F290S	CAC TCT	1	40
3	I228C-F229L	TGC CTC	1	78
3	I228V-F229L	GTG TTG	1	56
3	I228L-F229L	TTG CTC	1	35
3	I228V-F229M	GTG ATG	1	29
3	I228M-F229M	ATG ATG	1	28

Biochemical characterization of improved amylosucrase variants

The three most active double-mutants A289P-F290L, A289P-F290C and A289P-F290I were produced and purified to homogeneity to evaluate their steady-state kinetic properties and the product reaction profile obtained with and without the presence of **D'**. From sole sucrose (250 mM), all three double-mutants synthesized malto-oligosaccharides (up to DP 20) in the same way as the wild-type enzyme and the previously characterized F290K single-mutant (Figure 34A) (10). In comparison to these enzymes, double-mutants also produced a higher amount of sucrose isomers (trehalulose and mainly turanose) (Figure 34B).

When **D'** was added to the reaction mixture, double-mutants exhibited a strong propensity to catalyze **D'** glucosylation, similar to that previously described for F290K mutant. Indeed, malto-oligosaccharide production was fully suppressed (Figure 34C), more than 90 % of the glucosyl residues coming from sucrose being transferred onto **D'** to give the mono-glucosylated product **ED'** and a di-glucosylated molecule, the α -D-Glcp-(1→4)- α -D-Glcp-(1→4)- α -D-GlcpNAc-OAll (**EED'**) as characterized based on MS and NMR analysis (Figure 34D). Notably, A289P-F290L and A289P-F290I double-mutants produced a higher amount of trisaccharide **EED'** than A289P-F290C mutant for which the profile of products is close to that observed for F290K single-mutant (Table 7).



*Figure 34: Comparison of malto-oligosaccharide synthesis by wild-type amylosucrase (wtAS), F290K, A289P-F290C, A289P-F290I and A289P-F290L mutants from 250 mM sucrose supplemented or not with 250 mM allyl-2-acetamido-2-deoxy- α -D-glucopyranoside (**D'**)*

Panel A. Superposition of the HPAEC-PAD profiles obtained at the end of the reaction ($t_f=24$ h) from sole sucrose.

Panel B. Yields of glucosyl units incorporated into the various products synthesized in the total reaction medium from sole sucrose.

Panel C. Superposition of the HPAEC-PAD profiles obtained at the end of the reaction ($t_f=24$ h) from sucrose supplemented with **D'**.

Panel D. Yields of mono- and di-glucosylated **D'**, namely **ED'** and **EED'** respectively, from sucrose supplemented with **D'**. At final time, > 90 % of sucrose was consumed.

In order to determine the effect of amino acid substitution on catalytic properties, we performed steady-state kinetics. Upon varying sucrose concentration from 0 to 250 mM, all three double-mutants showed a standard saturation kinetic behaviour, whereas the previously identified F290K mutant was unable to be saturated. As shown in Table 7, the catalytic efficiency values for all three double-mutants are increased by 100 to 350-fold in comparison to the F290K mutant and are higher than that of wild-type amylosucrase by 2.5 to 8-fold. The K_m values for sucrose diminished for all three double-mutants. The k_{cat} values also increased for A289P-F290C and A289P-F290L double-mutants with the exception of A289P-F290I, which remained close to the wild-type value.

Table 7. Kinetic parameters determination of wild-type amylosucrase and improved mutants for sucrose donor (error range is within 5 to 10 %)

	k_{cat} (s^{-1})	K_m (mM)	$\frac{k_{cat}}{K_m}$ ($s^{-1} mM^{-1}$)	V_{250mM}^c ($\mu mol/min/g$)	Behaviour (at high concentration)
Wild-type amylosucrase^a	1.3	50.2	0.0261	900	saturation
F290K^b	<i>nd</i>	<i>nd</i>	0.0006	110	linear
A289P-F290L	2.1	10.1	0.2094	1717	saturation
A289P-F290C	2.8	18.9	0.1502	2186	saturation
A289P-F290I	1.1	17.2	0.0655	946	saturation

nd: Not determined

^a Data from (39)

^b Data from (10)

^c Initial rate of sucrose consumption was determined at sucrose concentration of 250 mM

The kinetic parameters were also determined from reactions carried out in the presence of a fixed concentration of sucrose donor (250 mM) and varying **D'** acceptor concentrations. The catalytic efficiency values of the double-mutants are increased by up to 395-fold by comparison to wild-type amylosucrase and by 3-fold respect to the F290K single-mutant (Table 8). Such an improvement in catalytic properties toward unnatural substrates has never been reported before for this class of enzymes. This shows that amylosucrase mutants are perfectly designed to glucosylate unnatural **D'** compound. The double-mutants are more efficient for **D'** glucosylation than the wild-type enzyme acting on its natural substrate. This

demonstrates that changing vicinal positions was extremely valuable to isolate remarkably improved mutants in the second round of our iterative enzyme optimization process.

Table 8: Kinetic parameters determination for wild-type amylosucrase and improved mutants for allyl-2-acetamido-2-deoxy- α -D-glucopyranoside acceptor (D') using a fixed 250 mM sucrose concentration (error range is within 5 to 10 %)

	$k_{cat}/K_m \text{ acc}$ ($s^{-1} \text{ mM}^{-1}$)	$V_{250\text{mM}}^b$ ($\mu\text{mol}/\text{min}/\text{g}$)
Wild-type amylosucrase^a	0.002	290
F290K^a	0.265	6,690
A289P-F290L	0.790	67,300
A289P-F290C	0.690	50,000
A289P-F290I	0.700	51,400

^a Data from (10)

^b Initial rate of acceptor consumption was determined at a concentration of 250 mM for both sucrose and α -D-GlcPNAc-Oall.

Impact of mutations and ligand on protein stability

The thermal stability of amylosucrases was assessed using a fluorescence-based approach, namely the differential scanning fluorimetry (DSF), in the absence and in the presence of a bound ligand. Purified proteins were subjected to gradually increasing temperatures to determine the melting temperature (T_m). The denaturation process was monitored by measuring the fluorescence from a dye (SYPRO orange) whose emission properties changed upon interaction with unfolded protein (40).

The average T_m values determined for the wild-type *N. polysaccharea* amylosucrase and the A289P-F290L, A289P-F290C and A289P-F290I mutants are listed in Table 9. Measurements were highly reproducible (Supplementary Figure S6A). For comparison purpose, the T_m of earlier reported F290K mutant (10) as well as that of A289P single-mutant were measured. We also determined the T_m values in the presence of sucrose substrate at varying concentrations. All enzymes under investigation were first inactivated by introduction of the E328Q mutation to prevent enzymatic cleavage of sucrose. The low impact of E328Q mutation introduction on the T_m value of the enzyme was checked for all mutants (Table 9). The T_m values for E328Q-AS in presence of 100 mM, 300 mM and 600 mM sucrose are provided in Table 9 (Supplementary Figure S6B).

Table 9: Average T_m of amylosucrases determined by differential scanning fluorimetry

	T_m (°C) of active enzymes	T_m (°C) for inactive enzyme (containing the E328Q mutation)			
		0 mM sucrose	100 mM sucrose	300 mM sucrose	600 mM sucrose
Wild-type amylosucrase	50.4 ± 0.2	49.0 ± 0.2	49.8 ± 0.2	50.3 ± 0.2	52.2 ± 0.4
F290K	45.1 ± 0.2	44.3 ± 0.1	45.5 ± 0.3	45.6 ± 0.2	46.6 ± 0.4
A289P	47.7 ± 0.1	nd	nd	nd	nd
A289P-F290L	45.5 ± 0.2	45.4 ± 0.6	47.0 ± 0.8	47.0 ± 1.2	48.8 ± 0.6
A289P-F290C	45.7 ± 0.2	45.6 ± 0.3	46.6 ± 0.2	47.4 ± 0.6	48.6 ± 0.8
A289P-F290I	44.5 ± 0.2	44.2 ± 0.3	44.8 ± 0.5	46.0 ± 0.7	45.9 ± 0.8

nd: Not determined

From these data, we can conclude that the mutations clearly diminished the thermal stability of both double-mutants and single-mutants. Comparison between T_m values of the single-mutants at positions 289 and 290 also indicates that the loss of stability can be mainly attributed to the mutation at position 290. Replacement of Phe290 at this position is also likely to be responsible for the improved recognition of **D'**. In addition, all mutants are more stabilized by sucrose substrate than the wild-type enzyme. We have seen that these double-mutants were all more efficient than the wild-type enzyme for catalyzing natural reaction or unnatural glycosylation of **D'**. The increase of catalytic efficiency values could well be due to their highest flexibility (correlated with lower T_m values) and their increased affinity for sucrose substrate (as indicated by lower $K_{m,suc}$, Table 7). Also, the replacement of Ala289 by a Proline residue probably contributes to catalytic efficiency improvements of the double-mutants. Indeed, F290K single-mutant is well designed for the glycosylation of the **D'** acceptor. However, the initial velocity (V_{max}) for transglucosylation is by 10 fold lower than that of the double-mutants. In addition, this mutant is way less efficient than the double-mutants when acting onto sucrose alone. Similarly, sucrose consumption by F290C, F290I and F290L single-mutants was not detectable in the absence of **D'** (data not shown). Consequently, the replacement of Ala289 by a Proline residue appears to contribute significantly to the properties of double-mutants. In order to support these observations by molecular interpretation, X-ray crystallographic analyses were performed.

Structural insight on improved acceptor specificity

The double-mutants were crystallized in conditions very close to those used to crystallize wild-type amylosucrase from *N. polysaccharea* (23,41). The structures of A289P-F290L, A289P-F290C and A289P-F290I mutants were solved at 2.4, 2.2, and 2.5 Å, respectively (Table 10). Our attempts to crystallize F290K mutant failed but inactive F290K-E328Q mutant (modified at amino acid position playing the role of general acid/base) was successfully co-crystallized in the presence of sucrose and its structure solved at 2.3 Å. For each mutant, electron density map over the mutated residues is quite well defined and allowed an unambiguous reconstruction (Figure 35B).

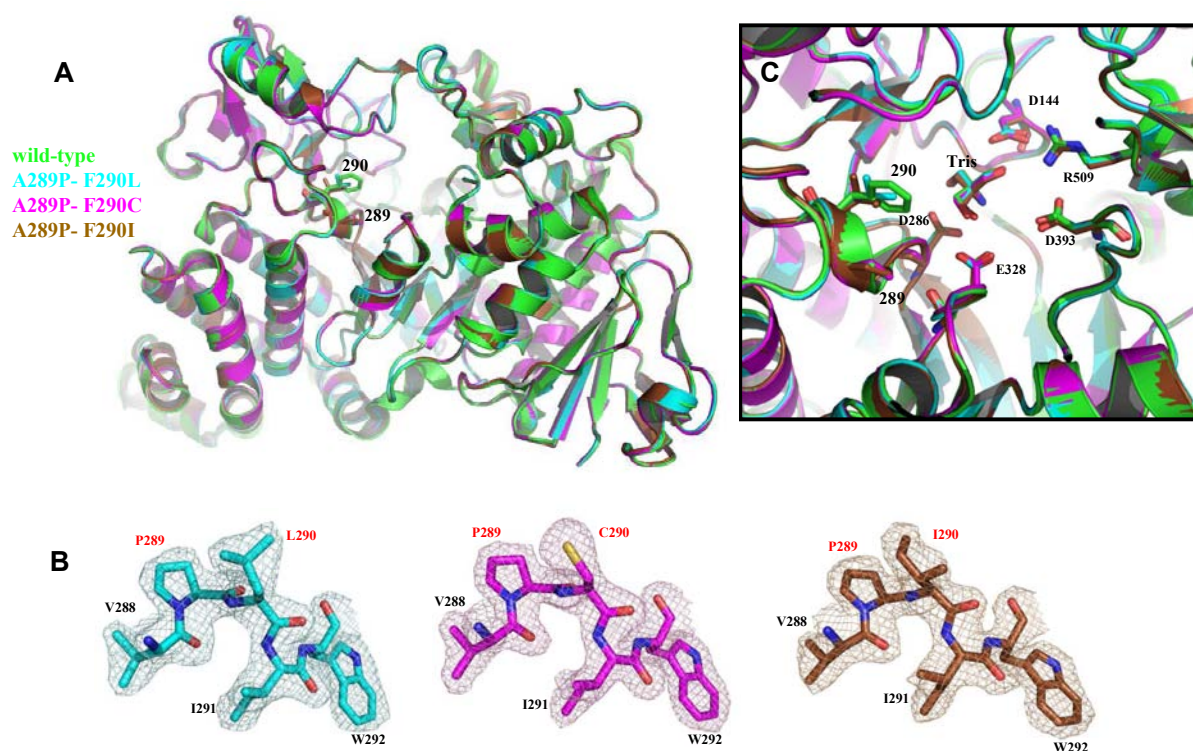


Figure 35: Comparison of the three-dimensional architecture of wild-type NpAS and NpAS mutants

With NpAS in green (PDB code: 1G5A (23,24)) A289P-F290L in cyan, A289P-F290C in magenta, A289P-F290I in brown.

Panel A. Overall organization with mutated residues in stick.

Panel B. Zoom of the mutated regions with a SigmaA weighted $2F_o - F_c$ electron density map contoured at 1.0σ around the residues.

Panel C. Superimposition of active site with mutated residues in positions 289 and 290; catalytic residues D286, E328, D393; salt bridge residues D144, R509 and the Tris ligand.

All mutants show a tertiary structure highly similar to that observed for wild-type amylosucrase with an overall RMSD between the C α atoms of 0.2 Å (Figure 35A). The amylosucrases are organized in five domains, three of which being characteristic of the GH family 13: the non-permuted (β/α)₈-barrel folded catalytic domain A, the domain B which is an extension of domain A between β -strand 3 and α -helix 3, and the C-terminal Greek key. Beside these common features, the three-dimensional architecture also contains two additional domains, namely the N-terminal α -helical domain and a B'-domain between β -strand 7 and α -helix 7 of the catalytic barrel. Positions 289 and 290 are carried by loop 4 (residues 285-303) of the catalytic barrel and closely interacting with loop 3 (residues 224-236) which both together with loop 7 participate, among others, to confer a pocket topology to the catalytic site.

Of note, all double-mutants were crystallized in the presence of a Tris molecule in their active site, like wild-type amylosucrase (Figure 35C). We assume that the presence of a bound ligand might stabilize a “closed” conformation of the enzyme in which loops recovering the catalytic site are stabilized by interactions with the ligand.

The F290K-E328Q mutant in complex with sucrose revealed two binding sites of sucrose which are unambiguously defined in the electron density map. These sites were previously described as SB1 for NpAS active site (PDB code: 1JGI (42)) and SB2 at the surface of domain B (PDB code: 1MW1 (41)) (Supplementary Figure S7A). SB1 was compared to its equivalent in the inactive E328Q NpAS-sucrose complex (Supplementary Figure S7B). Except for a slight reorientation of the Q328 side chain, no significant difference was observed in the sucrose binding mode (a listing of interactions is provided in Supplementary Table S2), which is involved in interactions with the catalytic residues (D286, E328Q, D393), the R509 and D144 amino acids (involved in a salt bridge that delineates the bottom of the catalytic pocket), the stabilizing Histidines (H187 and H392 which are highly conserved among GH13 family), and other residues (Y147, E284, D394 and R446). Interestingly, close inspection of the electron density map revealed a strong positive residual density located near a hydroxyl group of the fructose moiety (O1') and which was automatically assigned as a water molecule having a low B-factor value. Distances between this positive residual density and O1' from sucrose fructosyl moiety, wat817, wat1075, wat1334, the I330 main chain and the K290 side chain are 3.10 Å, 3.10 Å, 3.39 Å, 2.97 Å, 3.10 Å and 4.10 Å, respectively. Such hexacoordination together with the vicinity of K290 led us however to suspect that the positive density might in fact correspond to the presence of a chloride ion (Supplementary

Figure S8). Of note, a water molecule (wat1273) has been observed at this same position in the E328Q-sucrose complex (PDB code: 1JGI). Introduction of a positively charged residue at position 290 (F290K mutation) in amylosucrase active site might have thus contributed to stabilize the binding of the Cl⁻ anion. Chloride is known to be an activator of some α -amylases (43). Nevertheless, the localisation of this anion at +1 subsite, near the K290 mutated residue, is distinct in the case of amylosucrase compared to α -amylases, for which chloride is found to bind at equivalent subsite but not at the same location, close to an arginine residue (in NpAS, such position corresponds to the Y388 residue).

Regarding SB2 sucrose binding, no significant differences were observed between F290K-E328Q-sucrose and wild-type NpAS-sucrose (PDB code: 1MW1) complexes (Supplementary Figure S7C). In both cases, interactions with sucrose molecule involved residues D231, Q437, Y438 and S508 as summarized in Supplementary Table 2.

In all mutants, amino acid substitutions at positions 289 and 290 by less hindering residues did not modify side chain orientation of neighbouring residues, nor induced distortion of the polypeptide main chain (Figure 35C). Mutations of Ala289 by Pro, and Phe290 by Leu, Ile, Cys or Lys amino acid residues, also clearly led to a slight increase of the active site molecular surface that may facilitate accommodation of the *N*-acetyl group of the **D'** acceptor. However, all attempts to co-crystallize the double-mutants with sucrose, the **D'** acceptor or products were unsuccessful. Therefore, we performed molecular docking of sucrose and **ED'** in the active site of the double-mutants to explore further the effect of mutations. As shown in Figure 36A-B, sucrose is bound tightly in the -1 and +1 subsites defined according to the glycoside-hydrolase subsite numbering (44). While the -1 subsite accommodates the glucosyl unit transferred during the reaction, the +1 subsite plays a dual role: it is first occupied by the fructosyl unit from sucrose during the first step of the reaction and it subsequently receives the acceptor molecule (**D'**) during the second step. Amino acid side chains at positions 289 and 290 are not directly involved in interactions with sucrose, however they appear to be determinant in recognition of the *N*-acetyl group in **D'**. Mutations at positions 289 and 290 led to the uncluttering of the active site and the creation of a small lateral pocket which favours accommodation of the *N*-acetyl group of **D'** compared to wild-type amylosucrase (Figure 36C-D). Indeed, given that position 290 is able to tolerate amino acids as variable as Lys, Leu, Cys and Ile, with an increase in **D'** glucosylation rate, one could assume that electrostatic contribution alone is not at the origin of the improvement. In comparison to F290K mutant, the introduction of a Pro amino acid at position 289 is likely to

locally increase rigidity of loop 4 and alter dynamics of amylosucrase while modifying the active site topology. This hypothesis was investigated further by molecular dynamics simulations.

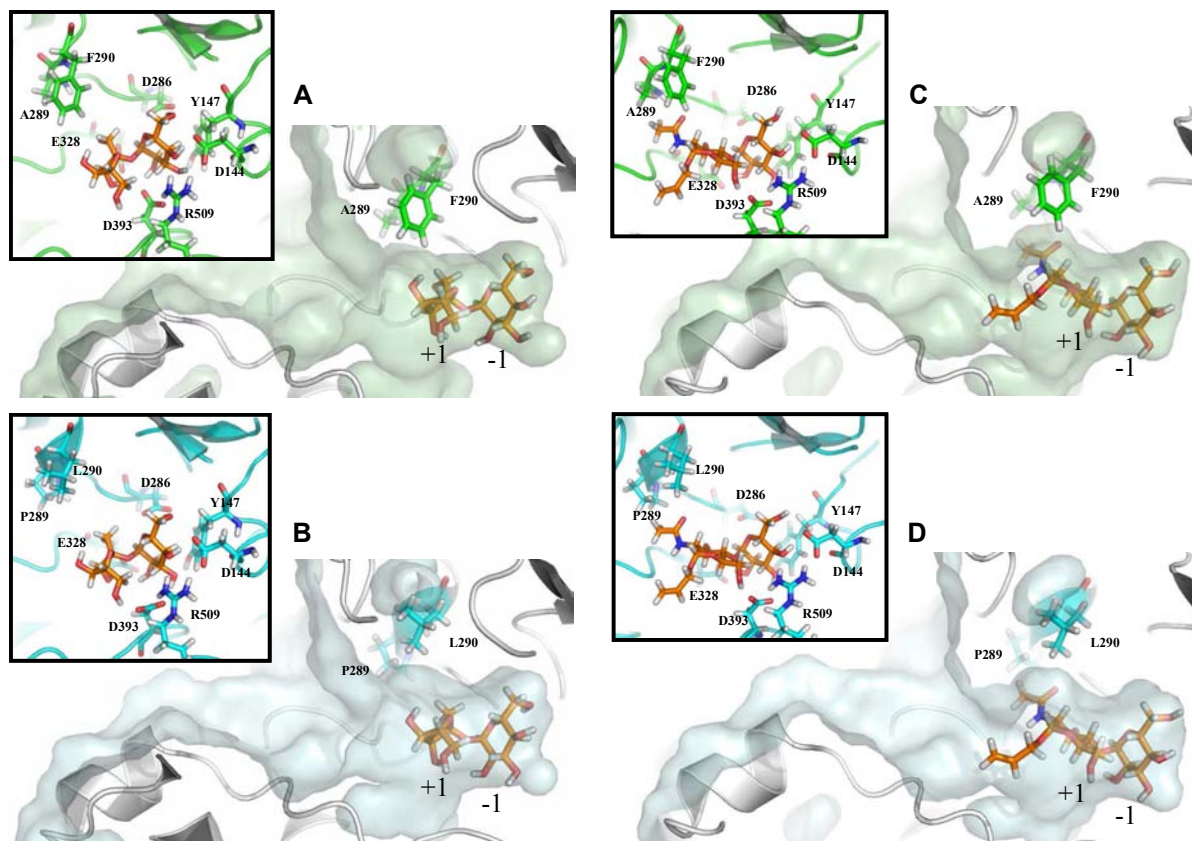


Figure 36: Molecular docking of ligands (sucrose and ED') in the active site of wild-type amylosucrase in green and one double-mutant chosen for illustration, A289P-F290L

A zoom of the active site is shown.

Panel A. Docking of sucrose in the active site of wild-type amylosucrase, in green (PDB code: 1G5A (23,24)).

Panel B. Docking of sucrose in the active site of A289P-F290L mutant, in cyan.

Panel C. Docking of ED' in the active site of wild-type amylosucrase.

Panel D. Docking of ED' in the active site of A289P-F290L mutant.

Role of mutations on amylosucrase flexibility

The wild-type amylosucrase, single-mutant F290K and double-mutants (A289P-F290L, A289P-F290C and A289P-F290I) were subjected to large scale Molecular Dynamics (MD) simulations (20 ns) in explicit water under constant temperature and pressure conditions (Supplementary Figure S9A). The obtained profiles for wild-type amylosucrase and double-mutants indicate that the backbone RMSD gradually increased within the first 5 ns of the simulation and then remains stable until the end of the simulation around an average value of

1.7 Å. Simulations of substrate-free and sucrose-bound F290K amylosucrase were also compared. Whereas a significant increase of the C_{α} RMSD is observed between 5 and 15 ns for the uncomplexed form, the profile corresponding to the complexed form exhibits relatively stable fluctuations (RMSD \sim 1.3 Å), with no major conformational change compared to wild-type amylosucrase (Supplementary Figure S9B).

In addition, the complexed form of F290K has lower backbone RMSD values than the free form, suggesting that the stability of F290K structure is strongly enhanced upon sucrose binding what is in agreement with the T_m values previously determined. Of note, no dissociation of the sucrose from amylosucrase was observed within the 20 ns of the simulation.

For all unliganded amylosucrases, analysis of simulated B-factors during the course of MD simulation (Supplementary Figure S9B) showed a particularly high mobility of loop 3 (residues 224-236), loop 4 (residues 285-303), loop 7 (residues 433-449) and loop 8 (residues 487-532). The amplitude of the conformational changes varied depending on the enzyme. These loops play a key role in conferring the active site a pocket topology in the X-ray structure (PDB code: 1G5A). While loops 2 and 8 close up the bottom of the pocket through a salt bridge formed by D144 and R509, loops 3, 7 and 8 cover up the oligosaccharide binding site, shielding it from solvent, and promoting thus transglucosylation reaction over sucrose hydrolysis. Overall, MD simulations indicate that domains B (residues 1-90) and B' (residues 396-460) exhibit the highest B-factors (Supplementary Figure S9B), although some differences are observed depending on mutants. Introduction of a Proline and Leucine or Lysine amino acid at positions 289 and 290, respectively, led to a drastic flexibility increase of loops 3, 4 and 7 compared to parental wild-type amylosucrase. However, when inserting mutations A289P-F290C and A289P-F290I, less pronounced effect was observed on the flexibility of loops 3, 4 and 7 compared to wild-type amylosucrase. Of all mutants (in unliganded form), the single-mutant F290K is the one for which both domains B and domain B' are highly flexible, suggesting therefore that introduction of Pro289 introduced some rigidity enabling catalytic triad residues to be maintained in productive conformation in spite of the proximity of positions 289 and 290 with nucleophile Asp286 which are all carried by flexible loop 4. In addition, A289P mutation enabled to reinforce sucrose recognition what is in agreement with the lowest K_m for sucrose determined for double-mutants.

In addition, an opening of the loops was observed for all amylosucrase mutants, exposing the oligosaccharide binding subsites and reshaping it into a cleft topology.

Interestingly, A289P-F290L mutant and wild-type amylosucrase also showed a disruption of the D144...R509 salt bridge after 12 ns of the simulation, displaying putative negative subsites observed in α -amylases which were not accessible in amylosucrase X-ray structure. In simultaneous way, loops 3 and 7 moved apart by ~ 12 Å (Figure 37), emphasizing the significant flexibility of amylosucrases in the absence of ligand. For comparison purpose, simulations were also run with the F290K mutant in complex with sucrose (Supplementary Figure S9B-C). When bound to sucrose, the overall flexibility of the enzyme is significantly reduced. Interestingly, external loops 3 and 7 appear strongly stabilized in the presence of sucrose, suggesting a “rigidification” of these loops composing the “roof” of the active site pocket (Supplementary Figure S10). On the other hand, only minor influence of sucrose binding was detected on dynamics of the internal core. Taken together, results suggest that complexed amylosucrase is more compact and stable when sucrose substrate is bound at the active site. This could explain why X-ray structures of amylosucrase from *N. polysaccharea* were always solved in “closed” conformation with a ligand molecule (Tris, glucosyl moiety, sucrose or maltoheptaose) bound in the active site.

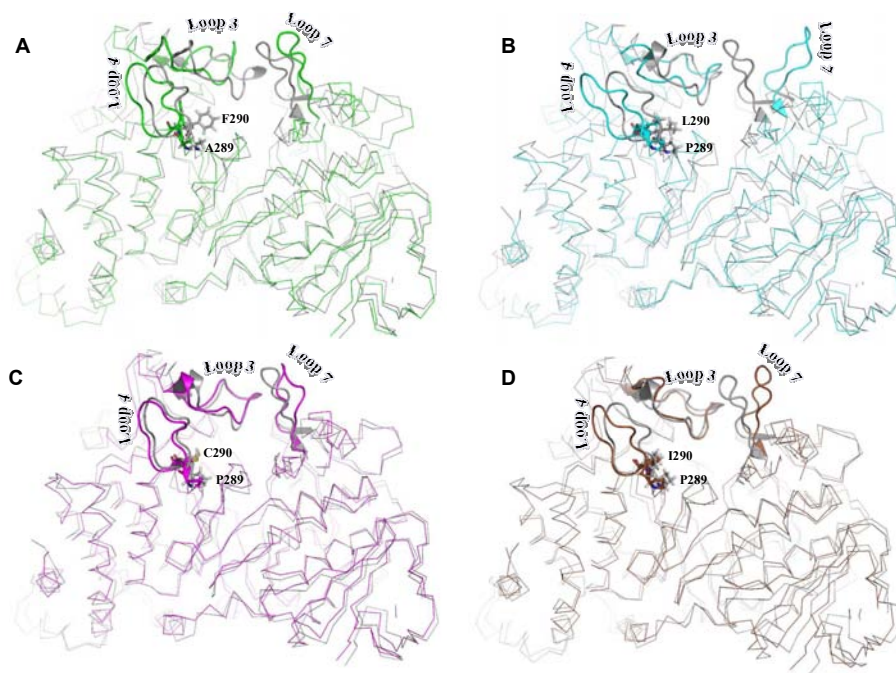


Figure 37: Superimposition of wild-type NpAS and double-mutant structures at the beginning (0 ns, in grey) and the extreme conformation identified along 20 ns MD simulation

- Panel A.** Wild-type amylosucrase (green).
- Panel B.** A289P-F290L mutant (cyan).
- Panel C.** A289P-F290C mutant (magenta).
- Panel D.** A289P-F290I mutant (brown)

The hypothesis that we propose is that only in the presence of a bound molecule in the catalytic site, loops 3, 4, 7 and 8 of the $(\beta/\alpha)_8$ barrel adopt a “closed” conformation conferring a pocket topology to the active site and favouring intermolecular interactions with the ligand. By promoting sucrose and **D'** recognition, the introduction of mutations at positions 289 and 290 from loop 4 is likely to alter significantly the conformational rearrangements involved in the “closure” of amylosucrase to adopt a catalytically productive state.

Conclusions

By recombining pairwise mutations at amino acid positions delineating the active site cavity of amylosucrase in a second round of iterative saturation mutagenesis, we improved by up to 8-fold the catalytic efficiency of best performing mutants towards sucrose substrate compared to wild-type parental enzyme, and by almost 400-fold the catalytic efficiency toward allyl-2-acetamido-2-deoxy- α -D-glucopyranoside, known as the **D'** acceptor. Such levels of activity enhancement have never been reported before for this class of enzymes. These results demonstrate the importance of the experimental design during the construction of the combinatorial library. It emphasizes the high malleability of α -transglucosidases and also further demonstrates the importance of exploring in depth the active site topology in particular in generating vicinal mutations when positive mutants have been targeted in the first run of active site plasticity exploration. This study also demonstrates how important is the understanding of inter-relationships between protein flexibility, function and stability to improve the design of biocatalysts. Indeed, results presented herein suggest that introduction of mutations at positions 289 and 290, which are located at the rim of the substrate pocket, only induced minor structural changes in amylosucrase crystal structures. More significant modifications induced by mutations were observed in MD simulations that revealed an alteration of the flexibility of some key structural regions that control active site topology, in particular loops 3, 4 and 7. Such conformational changes might thus favour recognition of substrates in catalytically active state, as shown by the enhanced catalytic efficiency of amylosucrase double-mutants for **D'** glucosylation. Although the wealth of biochemical and structural data provides invaluable insight into the mechanism of amylosucrases, the detailed aspects of the dynamic processes and the impact on the enzymatic mechanisms, *i.e.* substrate binding, product release, substrate-induced conformational changes and transesterification/hydrolysis ratio control, remain unstudied. Prior molecular modelling

studies suggested that loop 7 of amylosucrase could play a “conformational gating” role (45). By opening up, this loop induces an enlargement of the active site cleft that may channel a second sucrose molecule toward the catalytic site and also assist elongation of glycogen branches (46). The information provided herein about the molecular motion/dynamics of amylosucrase, bound to substrate or in its unliganded form, will thus help to better understand amylosucrase mechanism and enhanced specificity of mutants.

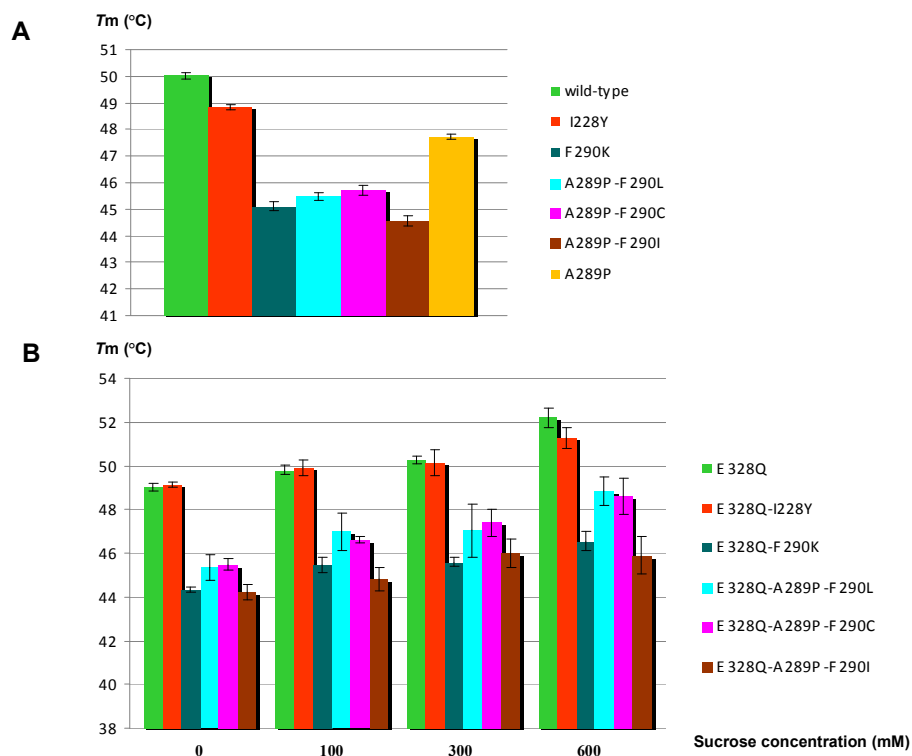
Supplementary informations

Supplementary Table S2: Binding interactions of sucrose with E328Q and F290K-E328Q mutants of amylosucrase

Sucrose Binding site 1 (SB1)		Sucrose Binding site 2 (SB2)	
E328Q-sucrose complex (PDB: 1JGI) (42)	F290K-E328Q-sucrose complex	E328Q-sucrose complex (PDB: 1MW1) (41)	F290K-E328Q-sucrose complex
D144(OD2)•••Glc(O4) 2.5 Å	D144(OD2)•••Glc(O4) 2.7 Å	D231(OD1)•••Glc(O3) 3.1 Å	D231(OD1)•••Glc(O3) 2.9 Å
H187(NE2)•••Glc(O6) 3.9 Å	H187(NE2)•••Glc(O6) 3.0 Å	D231(OD2)•••Glc(O2) 2.7 Å	D231(OD2)•••Glc(O2) 2.7 Å
R284(NH2)•••Glc(O2) 2.8 Å	R284(NH2)•••Glc(O2) 2.8 Å	D231(OD2)•••Glc(O3) 3.4 Å	
D286(OD2)•••Glc(O6) 3.1 Å	D286(OD2)•••Glc(O6) 3.0 Å	Q437(N)•••Glc(O2) 2.8 Å	Q437(N)•••Glc(O2) 3.0 Å
Q328(NE2)•••Glc(O1) 3.2 Å	Q328(NE2)•••Glc(O1) 3.4 Å	Y438(N)•••Glc(O3) 2.8 Å	Y438(N)•••Glc(O3) 2.8 Å
	Q328(NE2)•••Glc(O2) 3.5 Å	S508(N)•••Fru(O1') 2.8 Å	S508(N)•••Fru(O1') 3.1 Å
H392(NE2)•••Glc(O2) 2.9 Å	H392(NE2)•••Glc(O2) 2.9 Å		S508(OG)•••Fru(O1') 3.4 Å
H392(NE2)•••Glc(O3) 2.9 Å	H392(NE2)•••Glc(O3) 3.2 Å		
D393(OD1)•••Glc(O3) 2.9 Å	D393(OD1)•••Glc(O3) 2.8 Å		
D393(OD2)•••Glc(O2) 2.5 Å	D393(OD2)•••Glc(O2) 2.6 Å		
R509(NH1)•••Glc(O3) 3.4 Å	R509(NH1)•••Glc(O3) 3.5 Å		
R509(NH1)•••Glc(O4) 3.0 Å	R509(NH1)•••Glc(O4) 3.2 Å		
Q328(OE1)•••Fru(O1') 2.6 Å			
Q328(NE2)•••Fru(O1') 2.8 Å	Q328(NE2)•••Fru(O1') 3.0 Å		
Q328(NE2)•••Fru(O3') 3.4 Å	Q328(NE2)•••Fru(O3') 3.4 Å		
D393(OD2)•••Fru(O3') 2.8 Å	D393(OD2)•••Fru(O3') 2.9 Å		
D394(OD1)•••Fru(O6') 2.8 Å	D394(OD1)•••Fru(O6') 2.7 Å		
D394(OD2)•••Fru(O6') 3.4 Å	D394(OD2)•••Fru(O6') 3.4 Å		
R446(NH2)•••Fru(O6') 2.8 Å	R446(NH2)•••Fru(O6') 2.7 Å		
Water molecules			
Wat1110•••Glc(O3) 2.6 Å	Wat1110•••Glc(O3) 2.7 Å	Wat1264•••Glc(O3) 2.7 Å	
	Wat1110•••Glc(O4) 3.5 Å	Wat1267•••Glc(O4) 2.7 Å	
Wat1102•••Fru(O3') 3.5 Å		Wat1351•••Fru(O4') 3.3 Å	
Wat1111•••Fru(O3') 3.0 Å	Wat1111•••Fru(O3') 3.1 Å		
Wat1134•••Fru(O4') 2.8 Å	Wat1134•••Fru(O4') 2.8 Å		
Wat1273•••Fru(O1') 2.7 Å	Cl702•••Fru(O1') 3.1 Å		
Wat1379•••Fru(O4') 2.7 Å			

Supplementary Table S3: Degenerate primers used for the construction of libraries 2 to 3 (A289-F290 and I228-F229, respectively)

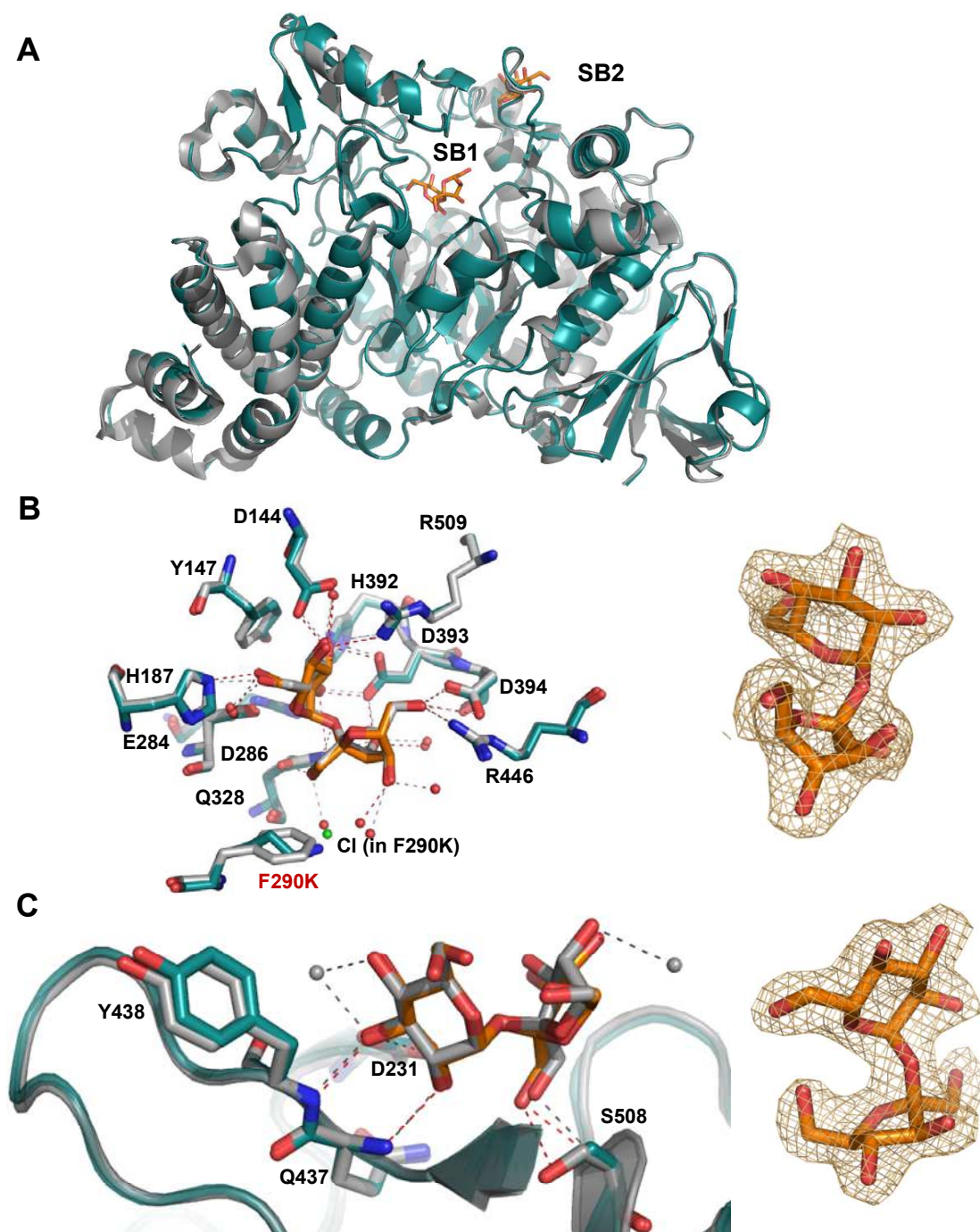
Primer Name	Nucleotide Sequence
228-229 for	5'-CTG CGC GAA NNS NNS CCC GAC CAG CAC CCG GGC G-3'
228-229 rev	5'-CTG GTC GGG WNN WNN TTC GCG CAG GGT GCG GTC G-3'
289-290 for	5'-GCG GTT NNS NNW ATT TGG AAA CAA ATG GGG ACA AGC TGC G-3'
289-290 rev	5'-CCA AAT SNN WNN AAC CGC ATC CAT ACG CAG GAT GTC AAC GCC-3'



Supplementary Figure S6: Graphical representation of T_m values of amylosucrases determined by differential scanning fluorimetry

Panel A. T_m values for active amylosucrases.

Panel B. T_m values for inactive amylosucrases (containing the E328Q mutation) with (100, 300 or 600 mM) or without sucrose (0 mM)



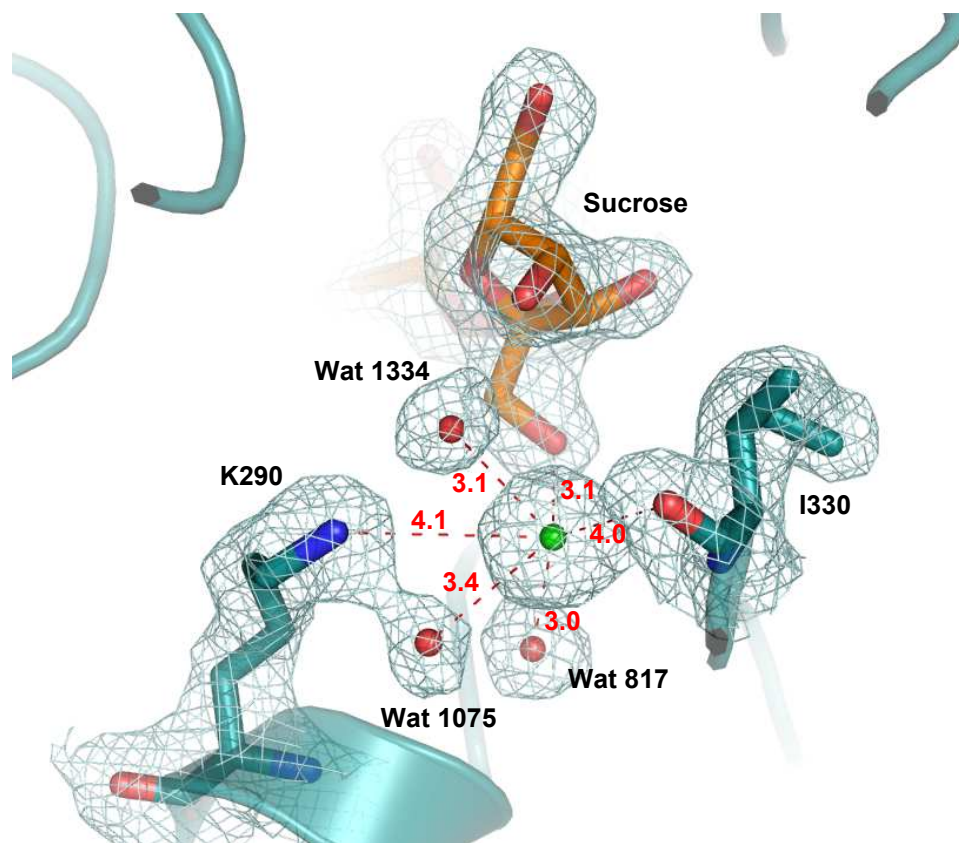
Supplementary Figure S7: Comparison of the X-ray structures of NpAS F290K-E328Q and NpAS E328Q in complex with sucrose

Panel A. Superimposition of NpAS F290K-E328Q (blue) with NpAS E328Q (PDB code: 1JGI, grey), sucrose ligands are represented in orange stick. Two sucrose binding sites are observed: SB1 located at the catalytic site and SB2 on the surface of the protein.

Panel B. View of catalytic Sucrose Binding site (SB1).

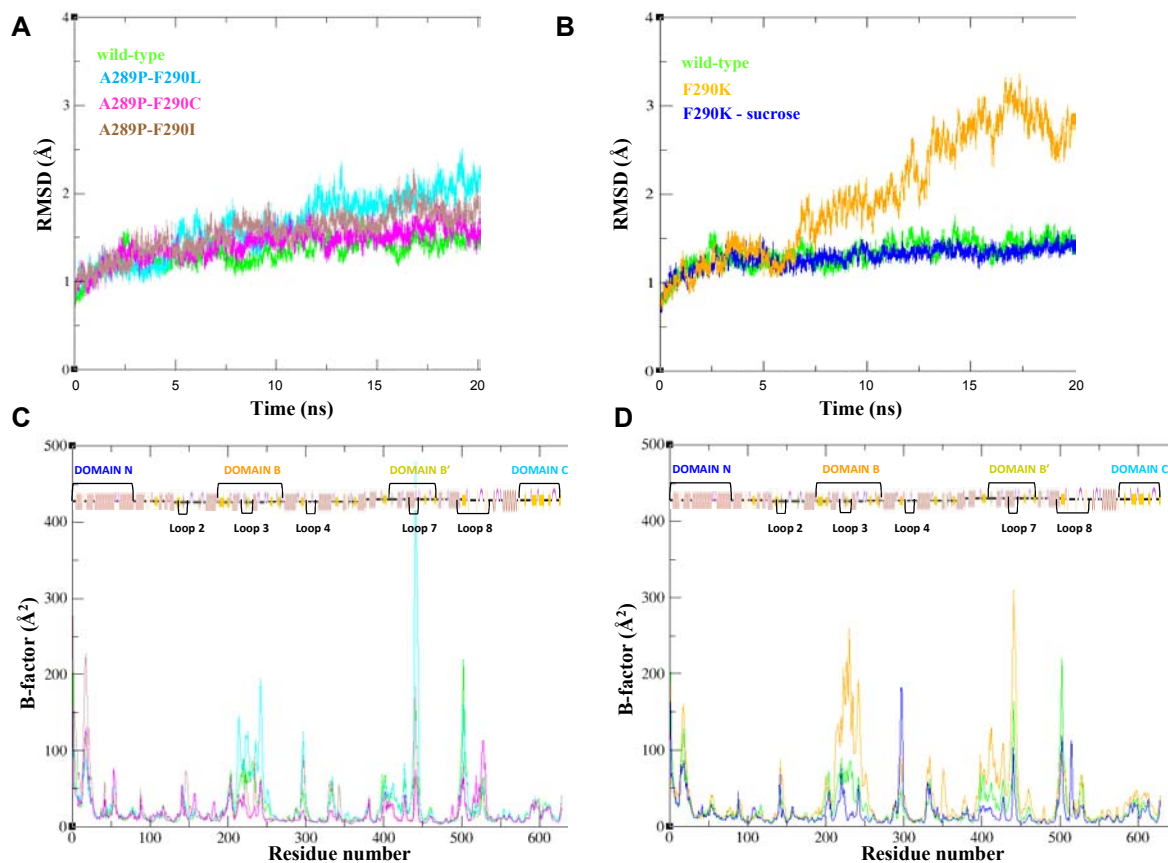
Panel C. View of SB2 located at the surface of the domain B'.

The SigmaA weighted $2F_o - F_c$ electron density maps contoured at 1.0σ of sucrose molecule are provided at the right of the figure for sucrose in each site.



Supplementary Figure S8: View of bound sucrose in the F290K-E328Q mutant active site

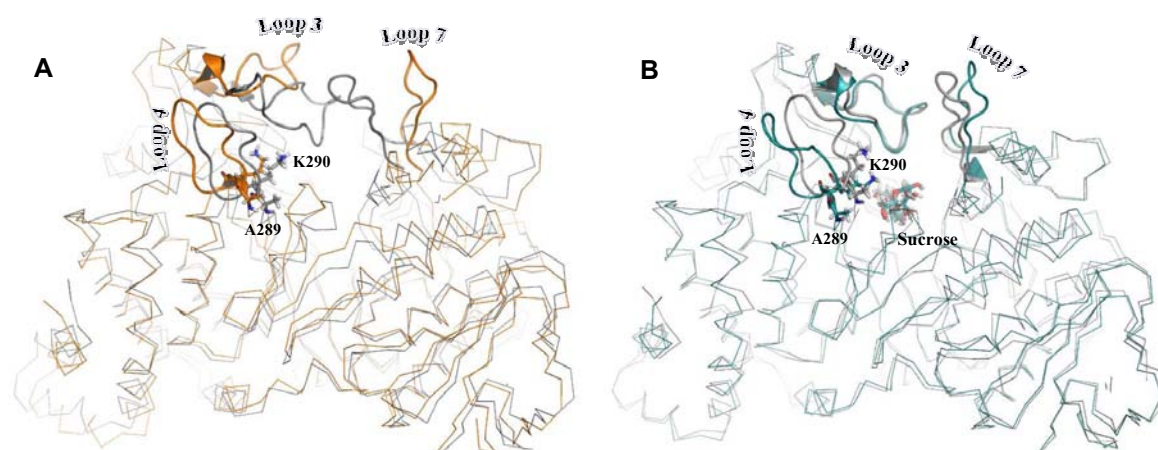
The Chloride ion (green) is found hexacoordinated with Ile 330, the O1' from the sucrose fructosyl moiety and 3 water molecules (red), bonds and length bonds are showed in red. The sigmaA weighted $2F_o-F_c$ electron density map is contoured at 1.0σ around the atoms.



Supplementary Figure S9: Analysis of 20 ns MD simulations performed on wild-type NpAS and mutants in their apo or sucrose-bound forms in explicit water, starting from their respective X-ray structures

Panel A and B. Root mean squared deviation of overall AS backbone atoms as a function of time.

Panel C and D. Calculated B-factors of AS residues from MD simulations.



Supplementary Figure S10: Conformation rearrangements occurring during the course of 20 ns molecular dynamics simulations carried out in explicit water

Panel A. Superimposition of snapshots from MD simulation of F290K-E328Q mutant taken at the beginning (grey) and the extreme conformation identified along 20 ns MD simulation (orange).

Panel B. Superimposition of snapshots from MD simulation of F290K-E328Q mutant in complex with sucrose taken at the beginning (grey) and the extreme conformation identified along 20 ns MD simulation (cyan).

References

1. Niyogi, S. K. (2005) Shigellosis *J Microbiol* **43**, 133-143
2. Foster, R. A., Carlin, N. I., Majcher, M., Tabor, H., Ng, L. K., and Widmalm, G. (2011) Structural elucidation of the O-antigen of the *Shigella flexneri* provisional serotype 88-893: structural and serological similarities with *S. flexneri* provisional serotype Y394 (1c) *Carbohydr Res* **346**, 872-876
3. Passwell, J. H., Harlev, E., Ashkenazi, S., Chu, C., Miron, D., Ramon, R., Farzan, N., Shiloach, J., Bryla, D. A., Majadly, F., Roberson, R., Robbins, J. B., and Schneerson, R. (2001) Safety and immunogenicity of improved *Shigella* O-specific polysaccharide-protein conjugate vaccines in adults in Israel *Infect Immun* **69**, 1351-1357
4. Passwell, J. H., Ashkenzi, S., Banet-Levi, Y., Ramon-Saraf, R., Farzam, N., Lerner-Geva, L., Even-Nir, H., Yerushalmi, B., Chu, C., Shiloach, J., Robbins, J. B., and Schneerson, R. (2010) Age-related efficacy of *Shigella* O-specific polysaccharide conjugates in 1-4-year-old Israeli children *Vaccine* **28**, 2231-2235
5. Belot, F., Guerreiro, C., Baleux, F., and Mulard, L. A. (2005) Synthesis of two linear PADRE conjugates bearing a deca- or pentadecasaccharide B epitope as potential synthetic vaccines against *Shigella flexneri* serotype 2a infection *Chemistry* **11**, 1625-1635
6. Phalipon, A., Tanguy, M., Grandjean, C., Guerreiro, C., Belot, F., Cohen, D., Sansonetti, P. J., and Mulard, L. A. (2009) A synthetic carbohydrate-protein conjugate vaccine candidate against *Shigella flexneri* 2a infection *J Immunol* **182**, 2241-2247
7. Said Hassane, F., Phalipon, A., Tanguy, M., Guerreiro, C., Belot, F., Frisch, B., Mulard, L. A., and Schuber, F. (2009) Rational design and immunogenicity of liposome-based diepitope constructs: application to synthetic oligosaccharides mimicking the *Shigella flexneri* 2a O-antigen *Vaccine* **27**, 5419-5426
8. Lindberg, A. A., Karnell, A., and Weintraub, A. (1991) The lipopolysaccharide of *Shigella* bacteria as a virulence factor *Rev. Infect. Dis.* **13 Suppl 4**, S279-284
9. Champion, E., André, I., Mulard, L. A., Monsan, P., Remaud-Siméon, M., and Morel, S. (2008) Synthesis of L-rhamnose and N-acetyl-D-glucosamine derivatives with glucansucrases *J. Carbohydr. Chem.* **28**, 142-160
10. Champion, E., André, I., Moulis, C., Boutet, J., Descroix, K., Morel, S., Monsan, P., Mulard, L. A., and Remaud-Siméon, M. (2009) Design of alpha-transglucosidases of controlled specificity for programmed chemoenzymatic synthesis of antigenic oligosaccharides *J Am Chem Soc* **131**, 7379-7389
11. André, I., Potocki-Véronèse, G., More, S., Monsan, P., and Remaud-Siméon, M. (2010) Sucrose-Utilizing Transglucosidases for Biocatalysis *Top. Curr. Chem.* **294**, 25-48
12. Neylon, C. (2004) Chemical and biochemical strategies for the randomization of protein encoding DNA sequences: library construction methods for directed evolution *Nucleic Acids Res* **32**, 1448-1459
13. Yuan, L., Kurek, I., English, J., and Keenan, R. (2005) Laboratory-directed protein evolution *Microbiol Mol Biol Rev* **69**, 373-392
14. Monchois, V., Willemot, R. M., and Monsan, P. (1999) Glucansucrases: mechanism of action and structure-function relationships *FEMS Microbiol Rev* **23**, 131-151
15. Henrissat, B., and Davies, G. (1997) Structural and sequence-based classification of glycoside hydrolases *Curr Opin Struct Biol* **7**, 637-644
16. Coutinho, P. M., and Henrissat, B. (1999) Carbohydrate-active enzymes: an integrated database approach. In: Gilbert, H. J., Davies, G., Henrissat, B., and Svensson, B. (eds).

- Recent Advances in Carbohydrate Bioengineering*, The Royal Society of Chemistry, Cambridge
17. Reetz, M. T., Wang, L. W., and Bocla, M. (2006) Directed evolution of enantioselective enzymes: iterative cycles of CASTing for probing protein-sequence space *Angew. Chem. Int. Ed. Engl.* **45**, 1236-1241
 18. Reetz, M. T., and Carballeira, J. D. (2007) Iterative saturation mutagenesis (ISM) for rapid directed evolution of functional enzymes *Nat. Protoc.* **2**, 891-903
 19. De Montalk, G. P., Remaud-Siméon, M., Willemot, R. M., Planchot, V., and Monsan, P. (1999) Sequence analysis of the gene encoding amylosucrase from *Neisseria polysaccharea* and characterization of the recombinant enzyme *J. Bacteriol.* **181**, 375-381
 20. Ghosh, M., Dulina, R. G., Kakarla, R., and Sofia, M. J. (2000) Efficient synthesis of a stereochemically defined carbohydrate scaffold: carboxymethyl 2-acetamido-6-azido-4-O-benzyl-2-deoxy-alpha-D-glucopyranoside *J Org Chem* **65**, 8387-8390
 21. Champion, E., Moulis, C., Morel, S., Mulard, L. A., Monsan, P., Remaud-Siméon, M., and André, I. (2010) A pH-Based High-Throughput Screening of Sucrose-Utilizing Transglucosidases for the Development of Enzymatic Glucosylation Tools *ChemCatChem* **2**, 969-975
 22. Sumner, J., and Howell, S. A. (1935) A method for the determination of invertase activity. *J. Biol. Chem.* **108**, 51-54
 23. Skov, L. K., Mirza, O., Henriksen, A., Potocki de Montalk, G., Remaud-Siméon, M., Sarçabal, P., Willemot, R. M., Monsan, P., and Gajhede, M. (2000) Crystallization and preliminary X-ray studies of recombinant amylosucrase from *Neisseria polysaccharea* *Acta Crystallogr D* **56**, 203-205
 24. Skov, L. K., Mirza, O., Henriksen, A., De Montalk, G. P., Remaud-Siméon, M., Sarçabal, P., Willemot, R. M., Monsan, P., and Gajhede, M. (2001) Amylosucrase, a glucan-synthesizing enzyme from the alpha-amylase family *J Biol Chem* **276**, 25273-25278
 25. Battye, T. G., Kontogiannis, L., Johnson, O., Powell, H. R., and Leslie, A. G. (2011) iMOSFLM: a new graphical interface for diffraction-image processing with MOSFLM *Acta Crystallogr D* **67**, 271-281
 26. Evans, P. (2006) Scaling and assessment of data quality *Acta Crystallogr D* **62**, 72-82
 27. Collaborative Computational Project, N. (1994) The CCP4 suite: programs for protein crystallography *Acta Crystallogr D* **50**, 760-763
 28. Potterton, E., Briggs, P., Turkenburg, M., and Dodson, E. (2003) A graphical user interface to the CCP4 program suite *Acta Crystallogr D* **59**, 1131-1137
 29. Murshudov, G. N., Vagin, A. A., and Dodson, E. J. (1997) Refinement of macromolecular structures by the maximum-likelihood method *Acta Crystallogr D* **53**, 240-255
 30. Emsley, P., and Cowtan, K. (2004) Coot: model-building tools for molecular graphics *Acta Crystallogr D* **60**, 2126-2132
 31. Case, D. A., Darden, T. E., Cheatham, I. T. E., Simmerling, C. L., Wang, J., Duke, R. E., Luo, R., Merz, K. M., Pearlman, D. A., Crowley, M., Walker, R. C., Zhang, W., Wang, B., Hayik, S., Roitberg, A., Seabra, G., Wong, K. F., Paesani, F., Wu, X., Brozell, S., Tsui, V., Gohlke, H., Yang, L., Tan, C., Mongan, J., Hornak, V., Cui, G., Beroza, P., Mathews, D. H., Schafmeister, C., Ross, W. S., and Kollman, P. A. (2006) AMBER 9, University of California, San Francisco.
 32. Duan, Y., Wu, C., Chowdhury, S., Lee, M. C., Xiong, G., Zhang, W., Yang, R., Cieplak, P., Luo, R., Lee, T., Caldwell, J., Wang, J., and Kollman, P. (2003) A point-

- charge force field for molecular mechanics simulations of proteins based on condensed-phase quantum mechanical calculations *J Comput Chem* **24**, 1999-2012
33. Lee, M. C., and Duan, Y. (2004) Distinguish protein decoys by using a scoring function based on a new AMBER force field, short molecular dynamics simulations, and the generalized born solvent model *Proteins* **55**, 620-634
 34. Kirschner, K. N., Yongye, A. B., Tschampel, S. M., Gonzalez-Outeirino, J., Daniels, C. R., Foley, B. L., and Woods, R. J. (2008) GLYCAM06: a generalizable biomolecular force field. *Carbohydrates J Comput Chem* **29**, 622-655
 35. Pastor, R. W., Brooks, B. R., and Szabo, A. (1988) An Analysis of the Accuracy of Langevin and Molecular-Dynamics Algorithms *Mol Phys* **65**, 1409-1419
 36. Berendsen, H. J., Postma, J. P., Van Gunsteren, W. F., Di Nola, A., and Haak, J. R. (1984) Molecular dynamics with coupling to an external bath *J Chem Phys* **81**, 3684
 37. Essmann, U., Perera, L., Berkowitz, M. L., Darden, T., Lee, H., and Pedersen, L. G. (1995) A Smooth Particle Mesh Ewald Method *J Chem Phys* **103**, 8577-8593
 38. Ryckaert, J.-P., Ciccotti, G., and Berendsen, H. J. C. (1977) Numerical integration of the cartesian equations of motion of a system with constraints: molecular dynamics of n-alkanes *J Comput Phys* **23**, 327-341
 39. Potocki de Montalk, G., Remaud-Siméon, M., Willemot, R. M., Sarcabal, P., Planchot, V., and Monsan, P. (2000) Amylosucrase from *Neisseria polysaccharea*: novel catalytic properties *FEBS Lett.* **471**, 219-223
 40. Niesen, F. H., Berglund, H., and Vedadi, M. (2007) The use of differential scanning fluorimetry to detect ligand interactions that promote protein stability *Nat Protoc* **2**, 2212-2221
 41. Skov, L. K., Mirza, O., Sprogø, D., Dar, I., Remaud-Siméon, M., Albenne, C., Monsan, P., and Gajhede, M. (2002) Oligosaccharide and sucrose complexes of amylosucrase. Structural implications for the polymerase activity *J. Biol. Chem.* **277**, 47741-47747
 42. Mirza, O., Skov, L. K., Remaud-Siméon, M., Potocki de Montalk, G., Albenne, C., Monsan, P., and Gajhede, M. (2001) Crystal structures of amylosucrase from *Neisseria polysaccharea* in complex with D-glucose and the active site mutant Glu328Gln in complex with the natural substrate sucrose *Biochemistry* **40**, 9032-9039
 43. Aghajari, N., Feller, G., Gerday, C., and Haser, R. (2002) Structural basis of alpha-amylase activation by chloride *Protein Sci* **11**, 1435-1441
 44. 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
 45. Cortés, J., Siméon, T., Remaud-Siméon, M., and Tran, V. (2004) Geometric algorithms for the conformational analysis of long protein loops *J Comput Chem* **25**, 956-967
 46. Albenne, C., Skov, L., Tran, V., Gajhede, M., Monsan, P., Remaud-Simeon, M., and André-Leroux, G. (2007) Towards the molecular understanding of glycogen elongation by amylosucrase *Proteins* **66**, 118-126

Conclusions Générales et Perspectives

Au cours de mon travail de thèse, je me suis attaché à apporter des éléments de compréhension sur les relations unissant l'activité, la structure et la dynamique des amylosaccharases.

Nos travaux avaient initialement comme principal objectif l'identification des facteurs moléculaires responsables des propriétés physicochimiques et biochimiques des amylosaccharases depuis la compréhension de leur stabilité à la température jusqu'à l'identification de paramètres gouvernant leur spécificité de substrats et de produits. Pour progresser sur ces différents aspects, il fallait tout d'abord procéder à une analyse comparative des structures de différentes amylosaccharases judicieusement sélectionnées sur la base de leurs propriétés distinctes. Au début de nos travaux, seule la structure de l'amylosaccharase de *N. polysaccharea* (ASNp) sous forme *apo*, en complexe avec le saccharose (substrat de la réaction), en complexe avec le maltoheptaose (produit naturel de la réaction) ainsi qu'un intermédiaire covalent étaient disponibles.

Pour notre étude, nous avons sélectionné plusieurs amylosaccharases : l'amylosaccharase de *Deinococcus geothermalis* qui était connue pour être plus stable que l'ASNp ainsi que des mutants de l'ASNp déployant une spécificité très marquée pour la glucosylation d'un accepteur non naturel, l' α -allyl-*N*-acétyl-glucosamine. Nous avons produit ces protéines sous forme recombinante, les avons purifiées à homogénéité et avons parfois construit des mutants inactifs pour au final pouvoir disposer d'un ensemble de structures tri-dimensionnelles listé ci-dessous.

Structures tri-dimensionnelles résolues au cours de ce travail :

- ✓ ASDg *apo* à 1,97 Å de résolution
- ✓ ASDg en complexe avec le turanose à 2,10 Å de résolution
- ✓ ASNp en complexe avec le turanose à 1,85 Å de résolution
- ✓ ASNp A289P-F290L à 2,40 Å de résolution
- ✓ ASNp A289P-F290C à 2,20 Å de résolution
- ✓ ASNp A289P-F290I à 2,50 Å de résolution
- ✓ ASNp F290K-E328Q en complexe avec le saccharose à 2,30 Å de résolution

Les analyses structurales complétées par des caractérisations biochimiques et biophysiques ont ainsi pu être initiées pour répondre à un certain nombre de questions initialement posées et ainsi avancer dans nos objectifs. Ces interrogations étaient les suivantes :

1. Quels sont les déterminants moléculaires ou dynamiques impliqués dans la thermostabilité accrue de l'ASDg par rapport à l'ASNp ?
2. Comment la spécificité de produits de l'ASDg et de l'ASNp est-elle reliée à la structure ?
3. Comment les mutations introduites dans l'ASNp permettent-elles le gain d'efficacité catalytique observé vis-à-vis de l' α -allyl-*N*-acétyl-glucosamine ?

Nous dresserons ici le bilan des avancées permettant d'apporter des réponses à ces questions et d'avancer des perspectives.

1. Quels sont les déterminants moléculaires ou dynamiques impliqués dans la thermostabilité accrue de l'ASDg par rapport à l'ASNp ?

La différence majeure entre les structures de l'ASNp et de l'ASDg, mise à jour par la résolution de la structure tri-dimensionnelle de l'ASDg ainsi que par les données biochimiques et biophysiques obtenues en solution, provient de la structure quaternaire homodimérique de l'ASDg. En revanche, nous avons confirmé que l'ASNp est présente sous forme monomérique en solution. De plus, la comparaison des protomères a révélé que la première hélice α du domaine N de l'ASDg est bien plus longue que son équivalente dans l'ASNp et qu'elle présente peu d'interactions avec son entourage alors que les deux premières hélices α du domaine N de l'ASNp s'organisent en un domaine très compact. Ces différences impactent, comme nous l'avons montré, le comportement dynamique du domaine N de l'ASNp qui est très stable car maintenu par son réseau d'interactions, alors que le domaine N de l'ASDg se réarrange rapidement au niveau de la première hélice α .

Nous proposons que l'organisation en dimère de l'ASDg, stabilisant le domaine N, très flexible, soit impliquée dans la thermostabilité de l'enzyme. Cependant, il est probable que d'autres facteurs interviennent. En effet, en comparaison de l'ASNp, l'ASDg possède plus de résidus chargés (28,8 % vs 22,9 % pour l'ASNp), plus de résidus hydrophobes (46,9 % vs 43,7 % pour l'ASNp), plus de proline (36 % vs 30 % pour l'ASNp) et moins de résidus polaires non chargés (24,3 % vs 33,4 % pour l'ASNp). Ces particularités, très souvent relevées

dans les protéines thermostables, pourraient aussi être en partie responsables de la stabilité de l'ASDg.

Pour mieux saisir la part de ces différents phénomènes sur la stabilité, nous pourrions tenter de réaliser des simulations de dynamique moléculaire sur le dimère ou encore introduire des mutations sur la zone d'interaction entre protomères afin de déstabiliser l'homodimère et d'évaluer l'activité et la stabilité thermique de la forme monomérique de l'ASDg par rapport à celle du dimère. D'autre part, l'analyse de l'interface de dimérisation de l'ASDg nous permet de suggérer une séquence signature de la dimérisation qui est retrouvée plus ou moins conservée dans d'autres amylosaccharases. Il serait intéressant d'analyser certains de ces dimères putatifs pour renforcer nos hypothèses. En marge de ce travail, une analyse détaillée de séquences et l'obtention de nouvelles données structurales pourraient permettre d'approfondir l'identification d'une signature moléculaire liée à une organisation quaternaire et de proposer des outils de prédiction de la structure tri-dimensionnelle des amylosaccharases à partir de la simple structure primaire.

2. Comment la spécificité de l'ASDg est assurée par ses particularités structurales ?

Nous avons observé que l'ASDg, synthétise le turanose et le tréhalulose en quantité équimolaire contrairement à l'ASNp, qui catalyse préférentiellement la synthèse de turanose. La structure des complexes formés par ces deux enzymes avec le turanose a permis de mettre en évidence d'importantes variations au niveau du mode de fixation du turanose plus particulièrement, de l'unité fructosyle. En effet, dans le site actif de l'ASNp, le résidu fructosyle se trouve exclusivement sous forme ouverte, alors que dans l'ASDg, cette même unité est présente sous forme furanose en mélange α/β (la carte de densité électronique ne permettant pas d'exclure complètement les autres formes pyranose et ouverte). Le site actif de l'ASDg apparaît donc moins contraint au niveau du sous-site +1. Cela permettrait d'accommoder le fructose accepteur selon différentes conformations possibles, expliquant ainsi les proportions équivalentes de synthèse de turanose et de tréhalulose. A l'inverse, la poche catalytique de l'ASNp semble beaucoup plus dirigiste dans la fixation du fructose, spécialisant en partie l'enzyme pour la synthèse de turanose.

Pour conforter ces interprétations, l'obtention de complexe avec le tréhalulose pourrait être informative. Ces données structurales couplées à des études de dynamique moléculaire pourraient permettre d'appréhender le mode d'entrée de substrat ou de sortie de produit et

donc d'aller plus loin dans la compréhension du mode de reconnaissance des ligands au sein du site actif et ainsi aider à mieux comprendre les différences de profils de produits.

Enfin, pour spécialiser ces enzymes dans la production d'isomères de saccharose, des travaux d'ingénierie visant à réduire la formation de polymère pourraient être entrepris. De plus, les différences observées entre ASNp et ASDg pourraient aussi être mises à profit dans d'autres objectifs. Par exemple, la substitution de l'arginine 226 (ASNp) par la proline 219 dans le sous-site +4 de l'ASDg aboutit à une ouverture du site catalytique de l'ASDg qui pourrait être exploitée pour la glucosylation d'accepteurs plus encombrants que ceux que pourrait accommoder l'ASNp.

3. Comment les mutations introduites dans l'ASNp permettent-elles le gain d'efficacité catalytique observée vis-à-vis de l' α -allyl-*N*-acétyl-glucosamine ?

L'introduction de mutations en positions 289 et 290 du site actif de l'ASNp a permis d'améliorer l'efficacité catalytique de la glucosylation de l' α -allyl-*N*-acétyl-glucosamine, un accepteur d'intérêt. Dès lors, comprendre les déterminants moléculaires affectés par ces mutations et conduisant aux gains catalytiques observés est devenu une question primordiale à adresser afin de pouvoir à terme mieux contrôler la construction de mutants aux propriétés désirées. Pour approfondir cette question, nous avons entrepris l'étude structurale des différents mutants par cristallographie des rayons X. Malheureusement, les données collectées n'indiquent que de très faibles changements structuraux par rapport à l'enzyme sauvage ne permettant pas d'expliquer les gains catalytiques observés. Néanmoins, un ensemble de techniques biophysiques a été mis en œuvre afin d'apporter des éléments d'informations complémentaires. Pour cela, l'inactivation de chaque mutant par l'introduction de la mutation E328Q a été nécessaire pour mener les différentes études en présence de substrats ou produits de la réaction et éviter une hydrolyse des ligands. Grâce à des mesures de T_m par DSF, nous avons constaté que l'introduction des mutations influence significativement la stabilité thermique de l'ASNp avec une diminution du T_m pouvant aller jusqu'à 6 °C. A l'inverse, l'ajout de ligand, en particulier de saccharose, est suivi d'un gain significatif de stabilité des enzymes native ou mutantes. La déstabilisation liée à l'introduction des mutations est néanmoins accompagnée d'une remarquable augmentation de l'efficacité catalytique des mutants de l'ASNp. En parallèle, des simulations de dynamique moléculaire menées sur les mutants suggèrent une flexibilité accrue de certains éléments structuraux des mutants par comparaison avec l'ASNp sauvage. Il est à noter que la présence de saccharose lié au site actif

conduit à des réarrangements conformationnels beaucoup moins importants de l'ASNp et de ses mutants.

Pour répondre brièvement à cette question, nous pouvons donc conclure que :

- ✓ les mutations introduites dans l'ASNp ne modulent que très peu la structure *apo* des mutants par rapport à l'enzyme sauvage, cependant, un effet est clairement visible sur leur stabilité et leur dynamique, ces éléments pourraient être la clef expliquant les gains spectaculaires d'efficacité catalytique.
- ✓ les simulations de dynamique moléculaire menées sur des formes *apo* et en complexe avec du saccharose suggèrent une flexibilité importante des boucles 3, 4 et 7 du tonneau catalytique qui définissent la topologie du site actif. Ainsi, les mutations introduites sur ces boucles pourraient jouer un rôle sur leur flexibilité, et par conséquent sur la reconnaissance des ligands au niveau du site actif et l'activité catalytique ainsi que la spécificité des amylosaccharases.

Plusieurs perspectives reliées à ces travaux sont envisageables.

Il est essentiel de poursuivre les efforts sur l'étude en solution des mutants en présence de ligand (substrat, produit ou accepteur non naturel). L'influence des ligands sur la stabilité de l'enzyme combinée à la détermination des constantes de dissociation pourrait permettre de mettre en lumière des différences importantes entre les différents mutants qui fourniraient des éléments de réponse supplémentaires sur l'activité et l'efficacité catalytique. Sur le plan structural, les mutants sous forme *apo* ne montrant que des différences mineures, l'enjeu est maintenant de résoudre la structure :

- ✓ des mutants de l'ASNp en complexe avec le produit de la réaction
et
- ✓ des intermédiaires covalents β -D-glucosyl-enzyme en complexe avec un accepteur cible.

L'analyse de ces structures pourrait peut-être aider à l'identification des réarrangements structuraux pouvant intervenir lors de la reconnaissance du ligand et au piégeage éventuel de structures adoptant d'autres conformations.

D'une manière générale, l'analyse de nouveaux complexes couplée à l'étude de la dynamique moléculaire et de son impact sur la catalyse doit être poursuivie et intensifiée. L'entrée et la sortie des substrats et/ou produits, phénomènes primordiaux pour la catalyse et qui constituent

bien souvent les étapes limitantes de la réaction semblent associées au réarrangement des domaines B et B' et, plus particulièrement, à la flexibilité des boucles 3 et 7 de ces deux domaines. Ces réarrangements devraient être étudiés plus en détail par le biais de simulation ou d'analyses de changements de conformations par des techniques de FRET ou autres. L'accès à ces données est d'une importance capitale pour la compréhension des relations entre la structure, la dynamique et l'activité des amylosaccharases et sa prise en compte dans les stratégies rationnelles d'évolution de ces protéines visant à élargir le panel d'applications de ces biocatalyseurs.

Documentations Annexes I
Résultats Complémentaires

1 Effet du tampon sur la stabilité de l'ASNp

Afin de mener les études sur l'ASNp en solution, nous avons déterminé le conditionnement optimal de stabilité pour l'ASNp tout en restant dans un conditionnement relativement proche des conditions de tampon utilisées lors de la purification. L'optimisation a été effectuée sur l'enzyme sauvage par mesure du T_m à l'aide de la technique de DSF. Le criblage a été réalisé en faisant varier la concentration de l'ASNp sauvage allant de 2 μM (0,14 g.L^{-1}) à 10 μM (0,72 g.L^{-1}), la concentration en NaCl de 0 à 1,5 M et avec différents tampons (HEPES, Tris, phosphate et MES).

Ce criblage a montré que la concentration en protéine a peu d'influence sur la stabilité thermique dans la gamme étudiée. Une concentration de 2 μM a été identifiée comme idéale pour une étude par DSF, aussi bien pour sa stabilité que pour des raisons d'économie d'échantillon.

Les tampons Tris, HEPES ou MES 50 mM pH 7,3 confèrent la meilleure stabilité à l'ASNp par rapport au tampon phosphate (un exemple de résultat en tampon Tris est fourni en figure 38). La concentration du tampon HEPES n'a pas montré d'influence, ou très peu, sur la stabilité de l'enzyme. Ce tampon étant utilisé pour la cristallisation de l'ASNp, nous l'avons donc choisi et avons fixé sa concentration à 150 mM.

D'autre part, l'utilisation du tampon Tris est à proscrire car cette molécule est fréquemment retrouvée dans le site actif des structures cristallographiques de l'ASNp. La présence de cette molécule de Tris pourrait alors induire un effet sur la stabilité de l'enzyme et/ou entrer en compétition avec un ligand d'intérêt, ce qui pourrait fausser les analyses d'interactions entre protéine et ligand effectuées en solution.

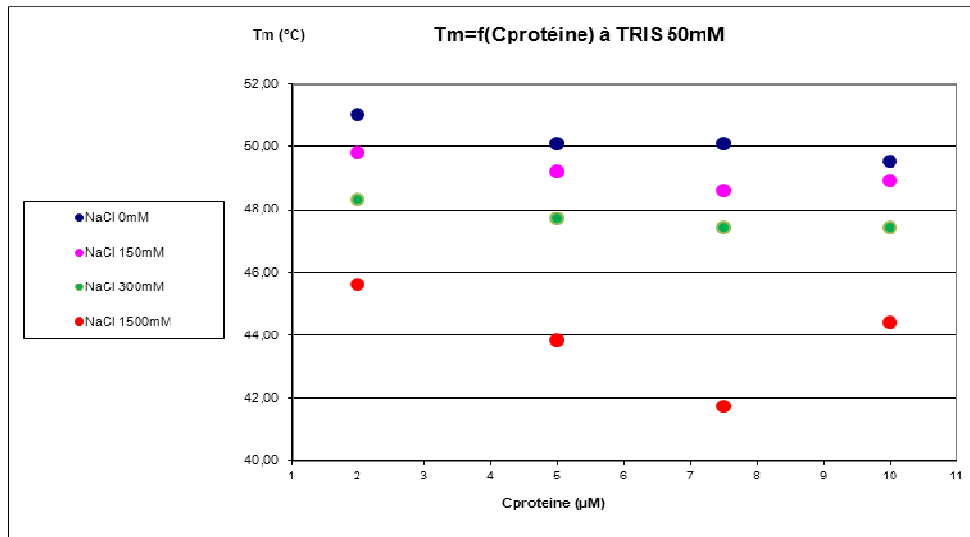


Figure 38 : Exemple de résultats obtenus par DSF sur du tampon Tris 50 mM en fonction de la concentration en NaCl et en fonction de la concentration en enzyme

En augmentant la force ionique, le T_m de l'ASNp est drastiquement diminué dans les tampons Tris, HEPES et MES à pH 7,3. L'augmentation de la force ionique induit donc une diminution du T_m de la protéine. Par exemple, le T_m de l'ASNp diminue progressivement pour passer de 51 °C à 45 °C en passant de 0 M à 1,5 M de NaCl dans 150 mM de tampon HEPES à pH 7,3. Ainsi, une concentration de 150 mM en NaCl permet de rester dans une condition proche de celle utilisée pour la purification de l'enzyme, mais provoque une baisse de T_m de 1,5 °C.

Nous avons donc sélectionné le milieu suivant : tampon HEPES 150 mM pH 7,3 ; NaCl 150 mM pour étudier l'ASNp en solution en absence ou en présence de différents ligands par DSF ou ITC. La concentration optimale d'enzymes pour lancer des analyses par DSF est de 2 μM.

2 Structure de l'ASNp cristallisée en tampon HEPES

La structure de l'ASNp sauvage (code PDB : 1G5A, classiquement cristallisée en tampon HEPES 0,1 mM pH 7, PEG 6 000 20 %, protéine conditionnée dans du Tris 50 mM pH 7, NaCl 150 mM, EDTA 1 mM, DTT 1 mM, 5 g.L⁻¹) présente au sein du site actif une molécule de Tris provenant du tampon de conditionnement de l'enzyme. Or, les simulations de dynamique moléculaire menées en absence de ligand suggèrent une ouverture du site actif de l'ASNp au niveau des domaines B et B', avec un écartement des boucles 3 et 7 pouvant aller

jusqu'à 25 Å (cf chapitre 3). Cette molécule de Tris pourrait stabiliser la structure, notamment les boucles 3 et 7, contraignant ainsi le site actif de l'enzyme à adopter une forme "fermée". Afin d'éviter tout artefact structural lié à la présence de la molécule de Tris, nous avons entrepris de produire et cristalliser l'ASNp en remplaçant le tampon Tris par du tampon HEPES. Habituellement, l'ASNp cristallise spontanément en trois jours à 12 °C. Lorsque l'enzyme est conditionnée directement en tampon HEPES, elle ne cristallise qu'après un micro-ensemencement par des micro-cristaux d'ASNp obtenus en condition classique.

Les résultats obtenus sont surprenants car, une fois résolue, la structure a montré une molécule de Tris au sein du site actif. Cette présence de Tris peut être expliquée par l'utilisation de la PreScission protéase pour cliver l'étiquette GST co-produite avec la protéine d'intérêt. En effet, la PreScission® protéase est conditionnée, entre autre, dans un tampon contenant 50 mM de Tris. La concentration résiduelle de Tris (~1 mM) pendant la phase de coupure pourrait être suffisante pour permettre une croissance cristalline, mais pas la nucléation, expliquant ainsi la nécessité de réaliser un micro-ensemencement.

L'occupation du site actif par le Tris ou un autre ligand tel que le saccharose, pourrait être une condition nécessaire à la nucléation et à la croissance des cristaux d'ASNp.

3 Structure basse résolution de l'ASDg

Les premiers essais de cristallisation ont permis d'obtenir des cristaux d'ASDg en présence de 25 % de PEG 400, 0,1 M de MgCl₂, 0,1 M de MES à pH 6. Ces cristaux appartiennent au groupe d'espace P4₁2₁2 et diffractent à 3,1 Å de résolution. Les paramètres de maille sont a=b=151 Å, c=187 Å, $\alpha=\beta=\gamma=90^\circ$. Cet empilement cristallin présente la particularité d'avoir deux molécules d'ASDg dans l'unité asymétrique, soit la structure du dimère. Malheureusement, la faible résolution a rendu impossible une analyse détaillée des interactions. Cependant, la structure quaternaire est identique à celle identifiée pour la structure publiée sous le code 3UCQ.

4 Détermination d'un nouveau conditionnement pour la cristallisation du mutant I228Y de l'ASNp par DSF

Les premiers essais de cristallisation réalisés aussi bien dans les conditions classiques pour l'ASNp sauvage qu'avec des cribles commerciaux en micro-plaques 96 puits (criblage de 960 conditions différentes), n'ont pas permis d'obtenir des cristaux du mono-mutant ASNp I228Y. Nous avons criblé par DSF différents conditionnements afin de déterminer ceux dans lesquels le mutant serait le plus stable. Les premiers criblages par DSF ont été effectués à l'aide des kits commerciaux pHClear et AmSO₄ (Qiagen). Nous avons observé que la valeur de T_m du mutant ASNp I228Y augmentait de 2 °C dans le tampon HEPES pH 7,0 en présence de sulfate d'ammonium à 1,6 M.

A partir des résultats obtenus sur les premiers criblages, nous avons évalué l'effet de différents additifs (de type sel de sulfate : sulfate de lithium LiSO₄, sulfate de magnésium MgSO₄ et sulfate d'ammonium AmSO₄) couplés au glycérol (connu pour son effet stabilisateur sur les protéines) sur la thermostabilité du mono-mutant. En présence d'AmSO₄ 1,2 M ou LiSO₄ 1,2 M ou d'un mélange de LiSO₄ 1,2 M avec du glycérol 15 % en tampon HEPES pH 7,0, des T_m plus élevés ont été observés pour les mono-mutants (Tableau 7). L'AmSO₄ étant volatile, nous avons décidé de rajouter, à la place de l'AmSO₄, du LiSO₄ 1,2 M en tampon HEPES pH 7,0 seul ou en mélange avec du glycérol 15 % pour le conditionnement des mono-mutants afin de lancer de nouveaux essais de cristallisation en micro-plaques 96 puits.

Tableau 7 : Température de dénaturation de l'ASNp I228Y en fonction de différents additifs

Additif	T_m ASNp I228Y (°C)
Sans additif	48,8
Glycérol 15 %	52,0
AmSO₄ 1,6 M glycérol 15 %	55,6
MgSO₄ 1,2 M glycérol 15 %	50,6
LiSO₄ 1,2 M glycérol 15 %	54,8

Les essais de cristallisation ont été lancés, toutefois, à ce jour aucun cristal n'a été observé dans ces conditions.

5 Analyse des interactions protéine-ligand par ITC

Une étude par ITC a été entreprise afin de déterminer la constante de dissociation K_d de l'ASNp, ainsi que des mutants améliorés vis-à-vis du saccharose et des accepteurs cibles (ou produits glucosylés). Par manque de temps, il n'a malheureusement pas été possible dans le cadre de cette thèse de mener à son terme l'étude envisagée. Néanmoins, les travaux préliminaires réalisés sur le saccharose et l'ASNp nous ont permis de vérifier la faisabilité des expériences par ITC et de mettre au point une ébauche des conditions de l'étude. Afin d'empêcher l'hydrolyse du saccharose au cours de la réaction catalytique, les enzymes utilisées sont préalablement inactivées par l'introduction de la mutation E328Q.

Les courbes de titrage obtenues pour le mutant E328Q de l'ASNp sur saccharose montrent une bonne corrélation avec un modèle à 3 sites de fixation du saccharose (figure 39).

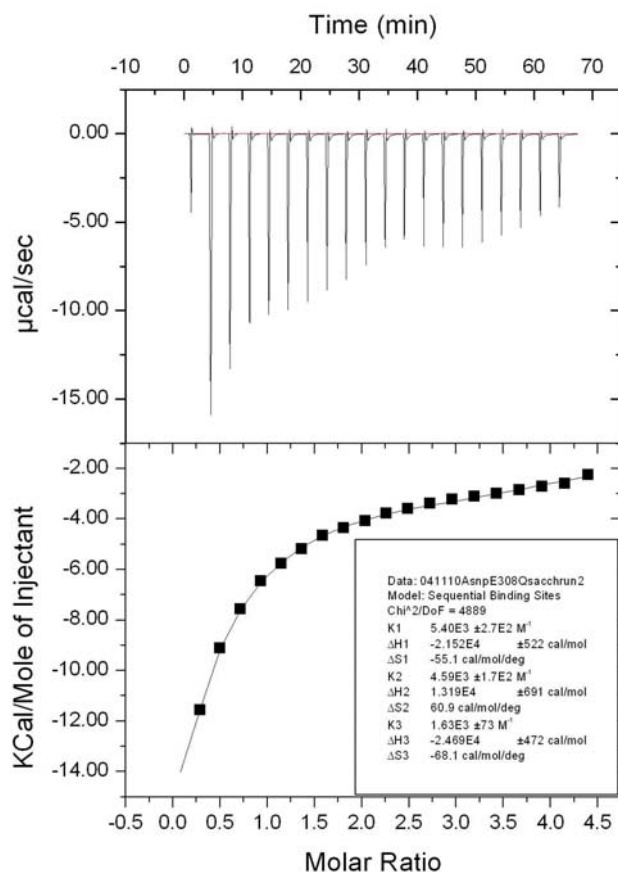


Figure 39 : Exemple de thermogramme obtenu (essai 2) avec une concentration en protéine de 420µM et un rapport saccharose/E328Q-ASNp : 20

Les résultats des différents essais réalisés à diverses concentrations en protéine sont regroupés dans le tableau 8

Tableau 8 : Résultats préliminaires des expériences d'ITC réalisées à partir de saccharose et du mutant E328Q de l'ASNp

		Essai 1	Essai 2	Essai 3
Site 1	Ka (M ⁻¹)	1,25.10 ⁴	5,4.10 ³	5,0.10 ³
	Kd (mM)	0,08	0,18	0,20
	ΔH (cal.mol ⁻¹)	-1,7.10 ⁴	-2,2.10 ⁴	-1,9.10 ⁴
	ΔS (cal.mol ⁻¹ .deg ⁻¹)	-37,5	-55,1	-46,7
Site 2	Ka (M ⁻¹)	3,7.10 ⁴	4,6.10 ³	7,9.10 ³
	Kd (mM)	0,03	0,22	0,012
	ΔH (cal.mol ⁻¹)	1400	1,3.10 ⁴	1,8.10 ⁴
	ΔS (cal.mol ⁻¹ .deg ⁻¹)	25,7	60,9	77,3
Site 3	Ka (M ⁻¹)	272	1630	1090
	Kd (mM)	3,6	0,61	0,92
	ΔH (cal.mol ⁻¹)	-6,53.10 ⁴	-2,5.10 ⁴	-2,66.10 ⁴
	ΔS (cal.mol ⁻¹ .deg ⁻¹)	-208	-68,1	-75,3

Essai 1 :

$$C_{\text{protéine}} = 573 \mu\text{M}$$

$$C_{\text{saccharose}} = 8,82 \text{ mM}$$

$$\text{Rapport saccharose/protéine} = 15$$

Essai 2 :

$$C_{\text{protéine}} = 420 \mu\text{M}$$

$$C_{\text{saccharose}} = 8,82 \text{ mM}$$

$$\text{Rapport saccharose/protéine} = 21$$

Essai 3 :

$$C_{\text{protéine}} = 280 \mu\text{M}$$

$$C_{\text{saccharose}} = 8,82 \text{ mM}$$

$$\text{Rapport saccharose/protéine} = 31$$

Les études cristallographiques réalisées sur l'ASNp ont mis en évidence la présence de 3 sites de fixation du saccharose et de malto-oligosaccharides, un premier au sein du site actif (SB1/OB1) et deux autres en surface de la protéine (un au niveau du domaine B' (SB2/OB2) et un au niveau du domaine C (OB3)). Les données structurales seraient donc en accord avec le modèle à 3 sites de fixation du saccharose proposé par les expériences d'ITC.

D'après ces résultats préliminaires, nous pouvons conclure que deux sites sont d'affinité comparable (ordre de la centaine de micromolaire), dont une conduite par l'enthalpie (site 1) et la seconde conduite par l'entropie (site 2). Le troisième site est d'affinité moindre (millimolaire). Ces manipulations devraient être répétées afin de vérifier la reproductibilité de la technique avec l'ASNp.

Ces travaux seront poursuivis très prochainement au sein du laboratoire pour l'étude de l'affinité des mutants améliorés de l'ASNp pour le saccharose ainsi que pour les produits de la réaction de glycosylation des accepteurs cibles.

Documentations Annexes II

Les travaux menés au cours de cette thèse ont permis de résoudre la structure cristallographique de l'amylosaccharase de *Deinococcus geothermalis* (ASDg), et d'identifier les facteurs architecturaux impliqués dans sa spécificité de produits. Nous avons pu mettre en évidence que cette enzyme se distingue de l'amylosaccharase de *Neisseria polysaccharea* (ASNp) par une structure quaternaire homodimérique, alors que l'ASNp est monomérique (cf chapitre 2). A la suite de cette étude, nous avons poursuivi la caractérisation structurale de ces deux enzymes, cette fois en solution et au cours de la catalyse, par diffusion des rayons X aux petits angles (SAXS) sur la ligne SWING du synchrotron Soleil.

Les résultats montrent que la structure des deux amylosaccharases en solution est identique à leur structure cristallographique. De plus, les données SAXS acquises en absence et en présence de saccharose sont elles aussi identiques, ce qui permet d'affirmer que ces deux enzymes n'adoptent pas de changements conformationnels majeurs durant la réaction, que ce soit à 30 °C pour l'ASNp ou à 50 °C pour l'ASDg.

De plus, ces travaux ont permis d'analyser les premiers stades d'enchevêtrement du polymère de type amylose produit par l'ASDg et l'ASNp, et ce, simultanément avec la caractérisation de la conformation des protéines en solution. En effet, après soustraction des signaux SAXS de l'enzyme, du saccharose et des oligosaccharides produits, il est possible d'extraire les courbes SAXS du polymère à différents temps de synthèse. Pour le polymère produit à 30 °C par l'ASNp, la structure la plus probable correspond à un empilement de 4 double hélices. Le modèle structural permettant le meilleur ajustement de la courbe de diffusion implique un décalage d'un demi-pas entre 2 doubles hélices voisines, comme observé dans les structures cristallines d'amidon. Dans le cas de l'ASDg, pour une concentration finale d'amylose identique à celle produite par l'ASNp, la taille moyenne des objets est plus petite et l'amylose ne s'auto-associerait pas sous la forme de structures spécifiques. En effet, dans ces conditions, le polymère produit correspondrait à un mélange de pelotes statistiques, de cylindres mal structurés et de particules de type "worm like". Ces résultats sont probablement liés à la plus haute température de synthèse utilisée, à laquelle l'amylose est beaucoup plus soluble qu'à 30 °C.

A notre connaissance, cette analyse par SAXS de la conformation d'enzymes au cours de la catalyse est unique. En effet, la conformation en solution d'enzymes est généralement étudiée à partir de mutants inactifs, ou en présence d'inhibiteurs analogues de substrats pour lesquelles. Par ailleurs, l'analyse simultanée de la conformation d'enzymes et des polymères

en cours de synthèse ouvre de nouvelles perspectives pour la compréhension des mécanismes de biosynthèse des polymères.

SAXS conformational tracking of amylose synthesis by amylosucrases

P. Roblin, G. Veronese, D. Guieysse, F. Guérin, M. Axelos, J. Perez and A. Buleon

Introduction

Amylose is one of the two polysaccharides that constitute starch, the major energy reserve for higher plants such as cereals, legumes, and tubers. Amylose is a linear polymer, consisting of α -(1,4)-linked glucosyl residues [1]. Self-association of amylose in aqueous solutions has been widely studied to understand both the crystallization of starch during biosynthesis and the structural changes occurring during starch processing [2]. Nonetheless, accessing amylose conformation during its polymerisation is still challenging. Indeed, granule-bound starch synthases [3] and starch phosphorylase [4], the plant enzymes that are responsible for amylose biosynthesis, require α -glucan primers and they are difficult to use *in vitro*, as most of them need to be bound to starch granule to maintain activity.

Bacterial amylosucrases represent a good alternative to mimic *in vitro* amylose polymerisation, as they are the only known enzymes that catalyze the synthesis of amylose from sucrose, a cheap agroressource, without any primer [5]. Amylosucrases belong to family 13 of glycoside-hydrolases (GH), as classified in the CAZy database [6]. The recombinant amylosucrase from *Neisseria polysaccharea* (NpAS) has been extensively studied for *in vitro* synthesis and modification of amylopolysaccharides [7]. The reaction starts by sucrose hydrolysis. The released glucose is then used as the first glucosyl acceptor to produce, after successive transfers via a multichain elongation process, maltooligosaccharides and longer α -glucan chains which precipitate when they reach critical size and concentration. Besides, trehalulose and turanose are produced as side reaction products, synthesized by glucosylation of the fructose that is released in the reaction medium. Amylose synthesis yield, morphology and crystallinity only depend on the initial sucrose concentration [8,9]. The structure and physico-chemical properties of the α -glucan polymer synthesized by amylosucrases were thought to be related to chains self-association phenomena [9], but the first stages of entanglement during synthesis before precipitation and crystallization of the polymer have never been investigated.

Another amylosucrase, amylosucrase from *Deinococcus geothermalis* (DgAS), is considered as very promising for industrial applications. Indeed, it catalyses the same reactions as NpAS,

but it is the most thermostable amylosucrase known to date, with an optimal temperature of 50 °C, while it is 37 °C for NpAS [10]. The size distribution of amylose synthesized by DgAS is slightly different than that produced by NpAS, with a lower amount of α -glucan chains of polymerisation degree (DP) higher than 30.

The only crystallographic structures of amylosucrases available to date are those of NpAS [11–16] and DgAS [17]. Their three-dimensional structures are organized in five domains, including a all helical N-terminal domain, a $(\beta/\alpha)_8$ -barrel folded catalytic domain A, domains B and B' (between β -strand 3 and α -helix 3 and between β -strand 7 and α -helix 7 of the catalytic barrel, respectively), and a β -stranded C-terminal domain adopting a greek key motif. Domains B' and N are specific to amylosucrases whereas the three other ones are conserved amongst GH13 enzymes. However, recent studies revealed one major structural difference between NpAS and DgAS crystallographic structures [17]: NpAS is monomeric, while DgAS presents an unusual homodimeric quaternary organization, that probably accounts for the enhanced thermostability of this protein, in particular by reducing the mobility of its N-terminal domain. However, amylosucrase conformation in solution during amylose synthesis reaction has never been studied, neither for NpAS nor for DgAS.

Small Angle X-ray Scattering (SAXS) has been widely used to determine the conformation in solution of proteins (see for review [18]). In particular, enzymes in complex with their substrates have been investigated, either inactive mutants with natural substrates or wild-type enzymes with non-reactive substrate analogs. However, there is no example, to our knowledge, of SAXS analysis of the enzyme conformation during catalysis. Besides, SAXS has been extensively used for investigating conformation in solution, entanglement or gelation of synthetic polymers, and in a lesser extent gelling or texturizing properties of polysaccharides, as carrageenan [19], gellan [20], xanthan [21], scleroglucan [22] or arabic gum [23]. Few papers describe the conformation of amylose in solution from SAXS measurements [24], computer simulation of SAXS curves [25,26] or gelation of amylose [27] or maltodextrins [28] but at a rather high concentration rapidly leading to recrystallization. In the present paper we describe the first SAXS analysis of the conformation in solution of glucan-synthesizing enzymes (*ie* NpAS and DgAS) during catalysis. Both the enzyme conformation in the reaction medium and the first stages of amylose entanglement during its *in vitro* synthesis were analyzed. Comprehension of amylose enzymatic synthesis is of utmost interest to design starchy products with optimal morphology and physico-chemical properties.

Experimental procedures

Production and purification of NpAS and DgAS. *Escherichia coli* JM109 cells containing pGST-NpAS or pGST-DgAS, respectively, were used to produce the fusion proteins glutathione-S-transferase/amylosucrase (GST-NpAS and GST-DgAS), as previously described [10,29]. Fusion protein purification by affinity chromatography using Glutathion-Sepharose-4B support (Amersham Biosciences), and GST tag removal were performed as previously described [29]. Protein content was measured with a Nano-Drop Spectrophotometer (NanoDrop Technologies, Wilmington, DE, USA).

Amylosucrase assay. One unit of amylosucrase activity corresponds to the amount of enzyme that catalyzes the release of 1 μmol of reducing sugars per minute from 50 g.L^{-1} sucrose and 0.1 g.L^{-1} glycogen in 50 mM Tris-HCl buffer, pH 7.0. The concentration of reducing sugars was determined using the dinitrosalicylic method [30] and fructose as standard.

Amylose synthesis reactions. Reactions were performed in 50 mM Tris-HCl buffer, pH 7.0, containing 111 and 128 mM sucrose as substrate, for NpAS and DgAS, respectively. Enzyme concentration was 0.18 g.L^{-1} for NpAS and 0.23 g.L^{-1} for DgAS, corresponding to 1,500 U.L^{-1} and 2,500 U.L^{-1} , respectively. Reaction temperature was 30 °C for NpAS and 50 °C for DgAS. During the reactions, samples were collected at regular time intervals for SAXS measurements and carbohydrate quantification.

Carbohydrate quantification. Samples collected during amylose synthesis were heated 5 min at 95 °C to stop reaction, before carbohydrate analysis, by high-performance anion-exchange chromatography with pulsed amperometric detection (HPAEC-PAD). Reaction yields were determined by quantification of monosaccharides (glucose and fructose), DP 2 oligosaccharides (sucrose, maltose, turanose and trehalulose), and DP 3 oligosaccharide (maltotriose) contained in the soluble fraction of the reaction media, that was diluted in water to 17 mg.kg^{-1} total sugars. Separation was performed on a 4x250 mm Dionex Carbo-pack PA-100. The flow rate of 150 mM NaOH was 1.0 mL.min^{-1} and an acetate gradient was applied as follows: 0-20 min, 6-196 mM; 20-70 min, 196-303.5 mM; 70-134 min, 303.5-450 mM; 135-140 min, 6 mM. Detection was performed using a Dionex ED40 module with a gold

working electrode and a Ag/AgCl pH reference. HPAEC-PAD does not enable one to directly quantify the amount of oligosaccharides for which standards are not available (maltooligosaccharides of DP higher than 7). Consequently, polymerization yields (corresponding to the production of DP>3 chains contained in the total reaction media, which comprise soluble and insoluble fractions) were determined as the difference between the amount of sucrose consumed and the amount of monosaccharides, maltose, turanose, trehalulose and maltoriose.

SAXS measurements. Small Angle X-ray Scattering (SAXS) experiments were carried out on the beamline SWING at Synchrotron SOLEIL (Gif-sur-Yvette, France). The beam wavelength was set to $\lambda = 1.033 \text{ \AA}$. The $17 \times 17 \text{ cm}^2$ low-noise Avix CCD detector was positioned at a distance of 1,800 mm from the sample, with the direct beam off-centered. The resulting exploitable Q-range was $0.008\text{-}0.57 \text{ \AA}^{-1}$, where $Q = 4\pi \sin \theta / \lambda$, and 2θ is the scattering angle. The solutions were circulated in a thermostated Quartz capillary with a diameter of 1.5 mm and $10 \text{ }\mu\text{m}$ wall thickness, positioned within a vacuum chamber. The transmitted intensity was permanently measured using a diode embedded in the beam-stop, with an accuracy of 0.1 %.

The experiences were started with small volumes of each sample ($20 \text{ }\mu\text{l}$), circulated into the flow-through capillary cell at $40 \text{ }\mu\text{l}\cdot\text{min}^{-1}$ and separated from the pushing liquid (water) by two air volumes of $6 \text{ }\mu\text{l}$ each, as described previously [31]. 15 frames of 2 s each were collected, normalized to transmitted intensity and subsequently averaged using “Foxtrot”, a dedicated home-made application. The same protocol was applied to buffer scattering.

Enzyme conformation determination. To obtain an unquestionable form factor of the two amylosucrases, a special protocol was applied. A small volume of a concentrated solution ($15 \text{ }\mu\text{l}$ at $15 \text{ mg}\cdot\text{mL}^{-1}$) was injected into a Size Exclusion Column (SHODEX KW402.5) using an Agilent © High Performance Liquid Chromatography (HPLC) system and eluted directly into the SAXS flow-through capillary cell at a flow rate of $150 \text{ }\mu\text{l}\cdot\text{min}^{-1}$. SAXS data were collected online throughout the whole elution time, with a frame duration of 2 s and a dead time between frames of 0.5 s. For each sample, the frames corresponding to the main elution peak were inspected for stability of their corresponding radius of gyration and a selection of them were averaged according to the procedure described by David & Perez [31].

SAXS analysis of amylose synthesis reaction media. Samples collected at different time intervals were analysed at the temperature used for the reaction (ie 30 and 50 °C for NpAS and DgAS respectively). After each sample analysis, 3 blanks were analysed: a protein control without any added sucrose, a buffer blank and a buffer containing sucrose.

SAXS data processing. For the NpAS kinetics, scattering curves were measured at different times after mixing with sucrose: 4, 180, 280, 480, 840 and 1,320 minutes, without any stop of the reaction. For the DgAS kinetics, three scattering curves were collected at 375,660 and 1,080 minutes. The first step in the SAXS data processing after the usual buffer subtraction consisted in subtracting the contribution of the enzyme, thus obtaining the signal of the mixtures of the carbohydrates present in the reaction media (α -glucan chains of DP>3, sucrose, glucose, fructose, maltose, maltotriose, turanose and trehalulose). Then, in order to obtain the signal of the polymer alone, the contribution of the other products was removed from the scattering curve. For each sample, the mass concentration of monosaccharide (glucose, fructose), sucrose and the other disaccharides produced by the enzyme (turanose, trehalulose and maltose) was separately determined by HPLC. With the range of Q values used for the experience, the scattering curves of glucose and fructose are identical, named $I(Q)_{\text{monosaccharide}}$, and the four types of disaccharide present in the reaction have the same SAXS signature, named $I(Q)_{\text{disaccharide}}$. Hence, to determine the SAXS signal of α -glucan chains of DP>3, we applied the following relationship:

$$I(Q)_{\text{polymer}} = I(Q)_{\text{reaction}} - I(Q)_{\text{protein buffer}} - X\% \text{ of mass} * I(Q)_{\text{disaccharide}} - Y\% \text{ of mass} * I(Q)_{\text{monosaccharide}}$$

Where $I(Q)_{\text{disaccharide}}$ and $I(Q)_{\text{monosaccharide}}$ were measured experimentally using pure solutions of glucose and sucrose and normalized to $c = 1 \text{ g/l}$.

SAXS data processing for the modeling of the polymer conformation modeling. The radius of gyration R_g from the scattering curve of the polymer was subsequently evaluated using the Guinier approximation [32] assuming that at very small angles ($Q \cdot R_g < 0.8$), intensity may be represented as:

$$I(Q) = I(0) \exp(-(Q \cdot R_g)^2/3) \quad (1)$$

At intermediate angles ($0,1 < Q < 0,15 \text{ \AA}^{-1}$), the section radius of gyration R_c of a rod shape particle was evaluated with the relationship [33]:

$$I(Q) * Q = I(0) \exp(-(Q \cdot R_c)^2/42) \quad (2)$$

When considering a cylinder model for the polymer in solution, the height H of the cylinder can be deduced from the values of R_g and R_c with the following relationship [33]:

$$H^2/12 = R_g^2 - R_c^2 \quad (3)$$

SAXS data processing for protein structure. The experimental intensity $I(Q)$ was fitted with a calculated curve derived from the crystallographic model of NpAS or DgAS using CRY SOL [34]. In this process, the excluded volume, the average atomic radius and the contrast of the hydration layer surrounding the particle in solution are optimized to minimize the following factor:

$$\chi^2 = \frac{1}{N-1} \sum_i \left[\frac{I_{\text{exp}}(Q_i) - c \cdot I_{\text{calc}}(Q_i)}{\sigma(Q_i)} \right]^2 \quad (4)$$

Where N is the number of experimental points, $I_{\text{exp}}(Q_i)$, $I_{\text{calc}}(Q_i)$ and $\sigma(Q_i)$ are the experimental and calculated intensity and the experimental error at the momentum transfer Q_i , respectively, and c is a scaling factor. The experimental error $\sigma(Q_i)$ was evaluated for each frame from the scattered number of photons $N(Q_i)$ as $\sigma(Q_i) = \sqrt{N(Q_i)}$ and adequately propagated.

Low-resolution shapes of NpAS and DgAS were generated ab initio, using the dummy atom model method, as implemented in the program DAMMIF [35], which uses simulated annealing to search a compact model that fits the Porod-modified experimental data by minimizing the factor χ^2 (4).

Results and discussion

Conformation of amylosucrases in solution

First, pure preparations of NpAS and DgAS were analysed to gain insights into enzyme conformation in solution. This also allowed to access the contribution of proteins to scattering, to further analyse the conformation of the products they synthesize. SAXS data from pure NpAS and DgAS preparations were collected directly at the exit of a SEC-HPLC system. For each of the two enzymes, the curve produced by averaging the frames of equal R_g within the main peak of elution is shown in Figure 1. The corresponding R_g values drawn

from the Guinier domain (equation 1) are close to 25.2 Å and 35.5 Å for NpAS and DgAS respectively, while these two proteins have similar molecular weights, as calculated from their sequences (71,491 and 73,166 Da, respectively). The autocorrelation functions $P(R)$, determined using the total set of scattering points, are presented as an inset in Figure 1. The maximal intramolecular distances are equal to 80 Å and 110 Å for NpAS and DgAS respectively. The $P(R)$ curves are nicely symmetric showing that both enzymes have a globular and compact shape. However, the differences of D_{\max} , R_g and SEC-HPLC elution volume (data not shown) between these two enzymes indicate that DgAS has a dimeric structure in solution, by contrast to NpAS which is monomeric.

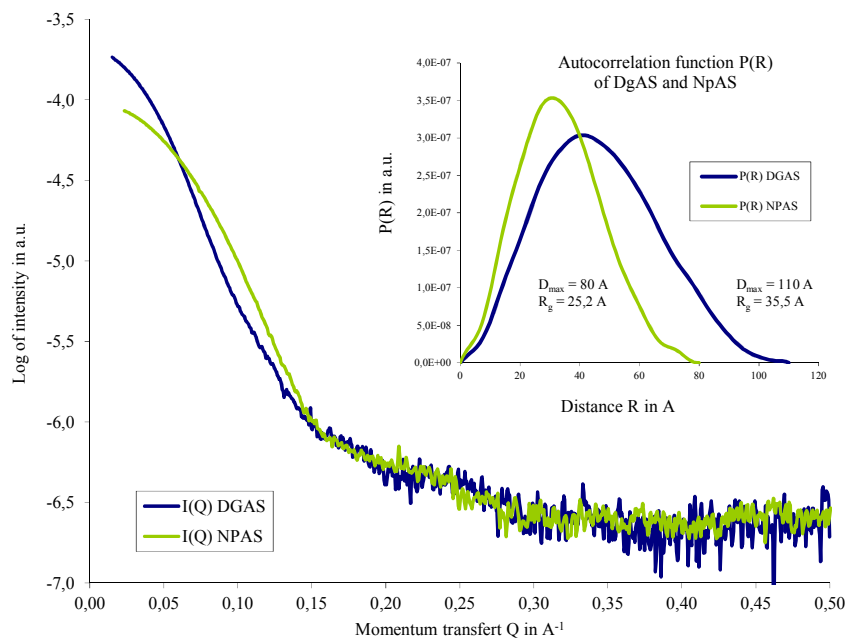


Figure 1: Superimposition of the SAXS curves of DgAS and NpAS

From the $P(R)$ results, low resolution envelopes of NpAS and DgAS were determined at 25 Å resolution ($Q = 0.25 \text{ \AA}^{-1}$) and were compared by visual superimposition to the high resolution 3D structures of NpAS (code PDB: 1G5A) and DgAS (code PDB: 3UCQ) determined by crystallography. The agreement in dimensions and shape was very good. The scattering curve calculated from the crystallographic structure with CRY SOL fits well with the experimental SAXS curve for both NpAS (chi value = 1.82) and DgAS (chi value = 2.16) (Figure 2). It confirms that the dimeric structure determined by SAXS for DgAS in solution is similar to that determined by crystallography, as for the monomeric structure of NpAS. Only small differences were observed at the N-terminal domain of DgAS that might be related to

the larger flexibility of the N-extremity observed for DgAS compared to NpAS [17]. Indeed, large scale molecular dynamics simulations carried out in explicit water suggested some important conformational rearrangements of the first helix (residues 1-24) of DgAS that were not observed for NpAS whose N-ter domain was found strongly stabilised by a network of ionic interactions.

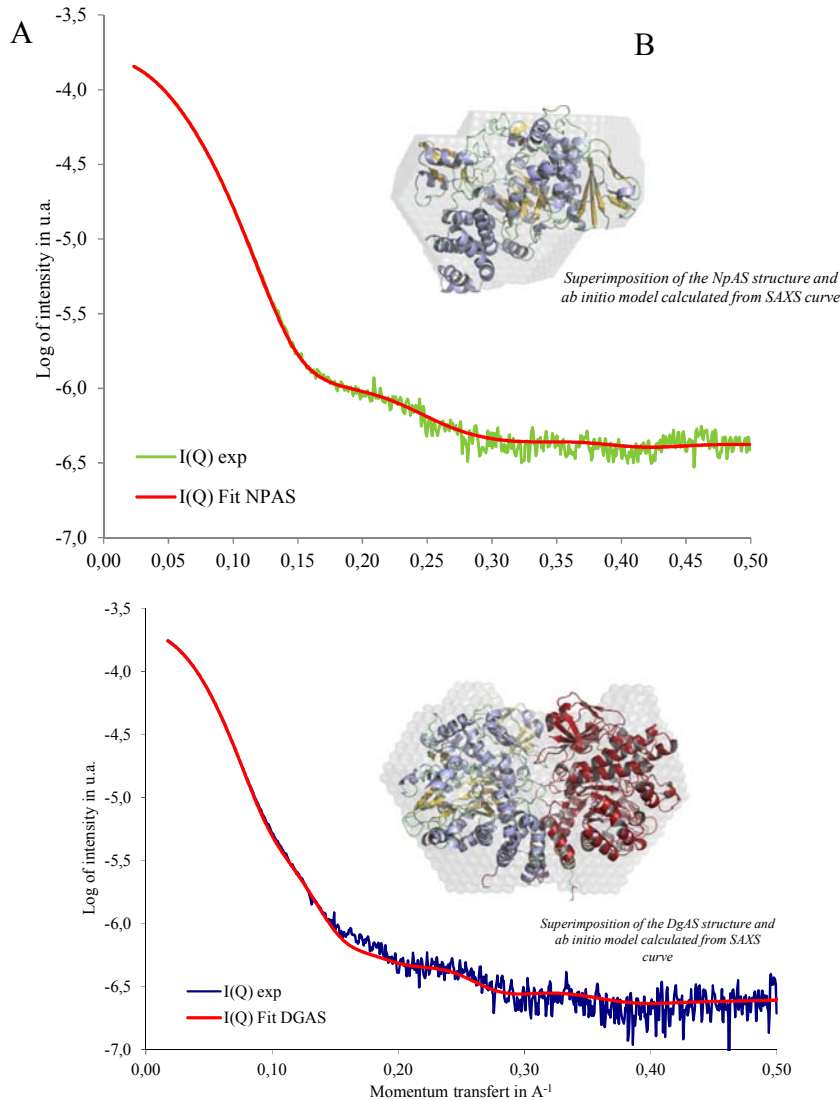


Figure 2: Comparison between experimental SAXS data and calculated data from crystallographic structures

Panel A. Comparison between experimental SAXS curve of NpAS and scattering profile calculated from crystallographic structure of NpAS.

Panel B. Comparison between experimental SAXS curve of DgAS and scattering profile calculated from crystallographic structure of DgAS

SAXS measurements were then performed in the presence of sucrose, to analyse the enzyme conformation in the first stage of amylose synthesis. The first scattering curve was recorded after 4 min of reaction. In these conditions, only the enzyme and the sucrose contribute to scattering, as still no polymer was synthesized (tables 1 and 2).

Table 1: Yields of sucrose consumption, of DP 1 to 3 oligosaccharides and α -glucan (DP > 3) production by NpAS from sucrose

Time (h)	Consumed Sucrose (%)	DP 1 to 3 (mg.L ⁻¹)	DP 1 to 3 (% total mass)	Polymer DP>3 (mg.L ⁻¹)	Polymer DP>3 (% total mass)
0,0	0	37982	100.0	0	0.0
2.7	40.4	37054	97.6	928	2.4
4.7	65.5	33178	87.3	4804	12.7
8.0	93.6	28418	74.8	9565	25.2
14.0	99.5	27653	72.8	10330	27.2

Table 2: Yields of sucrose consumption, of DP 1 to 3 oligosaccharides and α -glucan (DP > 3) production by DgAS from sucrose

Time (h)	Consumed Sucrose (%)	DP 1 to 3 (mg.L ⁻¹)	DP 1 to 3 (% total mass)	Polymer DP>3 (mg.L ⁻¹)	Polymer DP>3 (% total mass)
0.1	0	43808	100.0	0	0.0
6.3	24.4	43430	99.1	378	0.9
11.0	50.2	40640	92.8	3168	7.2
18.0	71.4	32137	73.7	11671	26.3

By subtracting the sucrose signal, the SAXS curves of NpAS and DgAS in the conditions of reaction were then obtained. As shown in Figure 3, these curves are superimposable, especially at low Q, with the corresponding curves obtained after SEC-HPLC for pure NpAS and DgAS in reaction buffer (50 mM Tris-HCl buffer, pH 7.0) without any substrate. Same R_g and form factors were thus obtained in the presence or absence of sucrose. This result suggests that NpAS and DgAS structures do not undergo any major conformational change, detectable at this resolution, when they are in conditions of amylose synthesis and saturated in substrate ([protein] = 2.6 μ M and [sucrose] = 0.11 M).

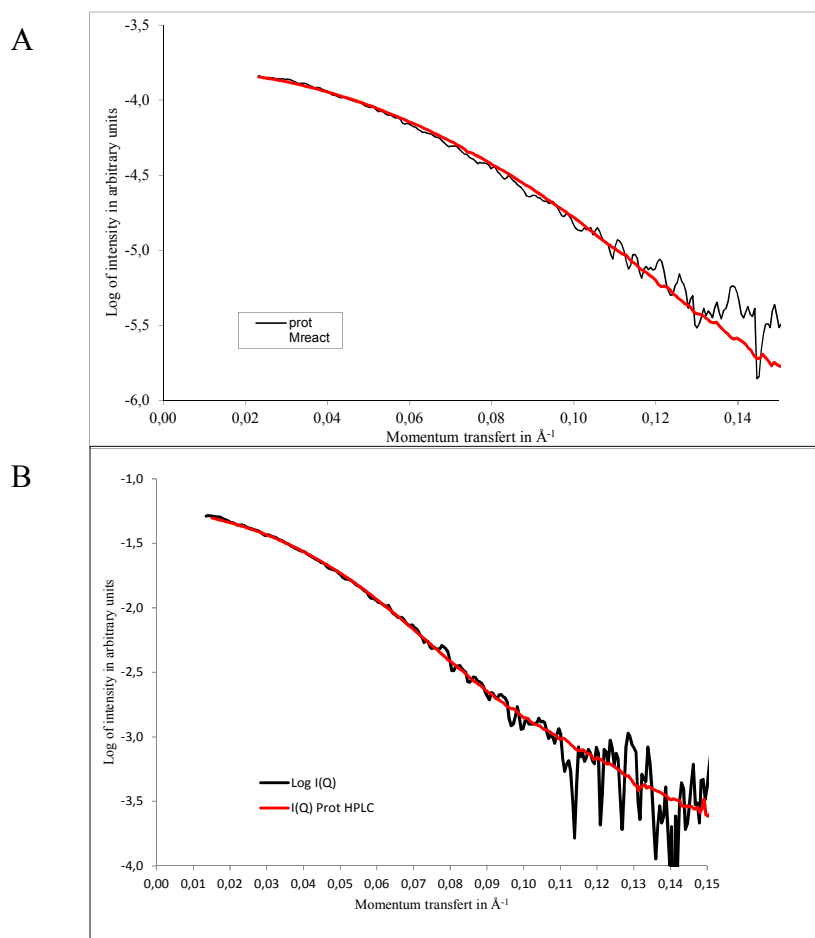


Figure 3: Superimposition of the SAXS signal of the reaction medium and the signal obtained with HPLC

Panel A. Signals of NpAS.

Panel B. Signals of DgAS.

SAXS modeling of amylose conformation during synthesis

SAXS analysis of the amylose synthesis reaction media was thus performed to model the conformation of the α -glucan growing chains. The signal corresponding to NpAS (0.18 g.L^{-1}) and DgAS (0.25 g.L^{-1}) at the different reaction times was first subtracted to SAXS curves, to access the contribution of carbohydrates that are present in the reaction media. In figure 4, the resulting "carbohydrate SAXS curves" are represented as $\log I(Q)$ versus $\log(Q)$ in order to highlight different regimes. Besides, carbohydrate amounts at the sampling times used for SAXS measurements were also quantified by HPAEC-PAD (Table 1).

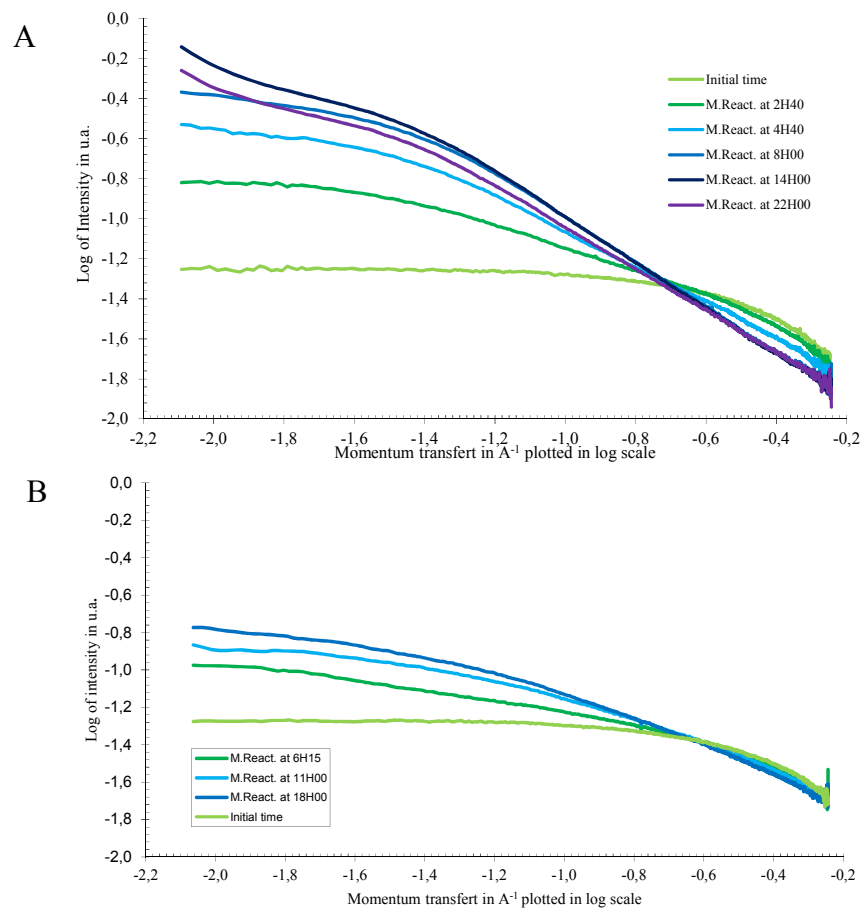


Figure 4: Kinetic of amylose polymerization followed by SAXS

Panel A. Kinetic of amylose polymerisation catalyzed by NpAS followed by SAXS, complete media (polymer, mono and disaccharide without protein contribution).

Panel B. Kinetic of amylose polymerisation catalyzed by DgAS followed by SAXS, complete media (polymer, mono and disaccharide without protein contribution).

For both enzymes, the carbohydrate SAXS curves at the initial synthesis time, result only from sucrose contribution to scattering, as no other oligosaccharide nor polymer was present in the reaction media. Then, two different behaviors were observed for NpAS and DgAS.

NpAS was able to consume 93.6 % and 99.5 % of sucrose after 8 h and 14 h, respectively. At these reaction times, α -glucan fraction reached 9,565 and 10,330 mg.L^{-1} (25.2 and 27.2 % of total mass), which is consistent with the usual NpAS behavior [9].

As a consequence, the carbohydrate SAXS signal intensity increases regularly up to 8 h. In that range, the curves are parallel at very low angles ($\text{Log}Q < -1.8$) indicating that the R_g remains constant. The intensity increase is thus mainly due to the increase of the polymer

particle concentration, which apparently reached a size limit after 8 h reaction. The ratio of $I(0)$ determined from the SAXS curve at 8 h to that from the curve at 2.7 h is equal to about 10^3 , which is in agreement with the ratio of α -glucan determined at these 2 synthesis times (table 1).

Then, the contribution of α -glucan to the carbohydrate SAXS signal was extracted by subtraction of monosaccharide and oligosaccharide ($DP \leq 3$) signals. The resulting α -glucan curves presented in Figure 5 can be divided into two parts. The first one, $-2.2 < \text{Log } Q < -1.5$ ($6.3 \times 10^{-3} < Q < 0.032 \text{ \AA}^{-1}$), corresponds to the Guinier region in which the radius of gyration, R_g , can be estimated. R_g is equal to 30 \AA at 2.40 h and increases to 35 \AA at 8 h and 14 h. The second part ranges from $\text{Log } Q = -1.2$ ($Q = 0.063 \text{ \AA}^{-1}$) to $\text{Log } Q = -0.6$ ($Q = 0.25 \text{ \AA}^{-1}$), and varies as $I(Q) = Q^{-d}$. For the first curve obtained after 2.7 h, the slope d is equal to 1.2, that is characteristic of rod-like objects, *i.e.* with a high height to width ratio. After 4.7 h, the d value increases up to 1.7 and remains stable until the end of reaction. This value suggests that the initial rod-like particles are growing in section, not in length.

To obtain a more quantitative view, the curves were analyzed using a rod-like model. After 2.7 h reaction, the section radius of gyration R_c (equation 2) is close to 4.4 \AA . With the corresponding R_g value, we found a height equal to 108 \AA using equation 3. After 4.7 h reaction, R_c reaches 8,5 \AA , then remains stable giving a constant value of the height H at 110 \AA in average. Thus, the rod-like particles grow in section but not in height, suggesting lateral self-association of the individual polymeric molecules such as measured at 2.7 h. For a quasi-one-dimensional object, R_g variation is more sensitive to the variation of height H than to that of section. Thus, while R_c increases significantly (multiplied by 2), R_g is proportionally less affected (30 \AA to 35 \AA).

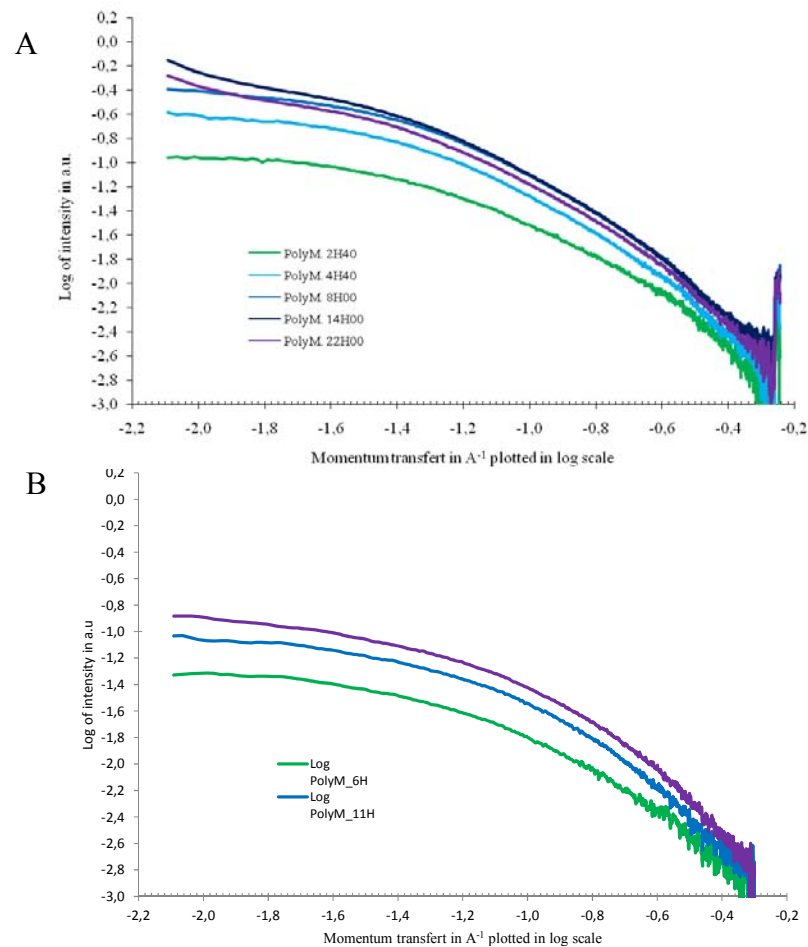


Figure 5: *Kinetic of amylose polymerization followed by SAXS, polymer alone contribution*

Panel A. *Kinetic of amylose polymerisation catalysed by NpAS followed by SAXS, polymer alone contribution.*

Panel B. *Kinetic of amylose polymerisation catalysed by DgAS followed by SAXS, polymer alone contribution.*

Finally, α -glucan conformation was modeled based on 1 to 4 amylose-like double helices from which theoretical scattering curves were calculated and compared to experimental scattering data. Indeed, amylose produced from 100 mM sucrose by NpAS was shown to crystallize as B-type networks [9]. B-type structure, usually encountered during amylose crystallization in water, consists of hexagonal packing of double helices [36]. Double helices are packed as duplex with a half pitch shift along the helix axis allowing two neighbouring helices to be more tightly packed. This duplex is also found in A-type structure, the other double helix based structure obtained from amylose crystallization [2,37]. In the present models, three turns of double helices were used, which corresponds in B-type structure to 10.5 nm (Figure 6), the value determined for the height of the cylinder determined from scattering curves. Neighbouring double helices were shifted as in the duplex structure determined for A and B-type amylose. When considering both R_c which increases from 4.2 Å

after 2.7 h reaction to 8.5 Å after 4.7 h reaction, the most probable structure would consist of an average of 4 associated double helices before further aggregation and precipitation which yield to too big particles to be studied by SAXS in the Q range used.

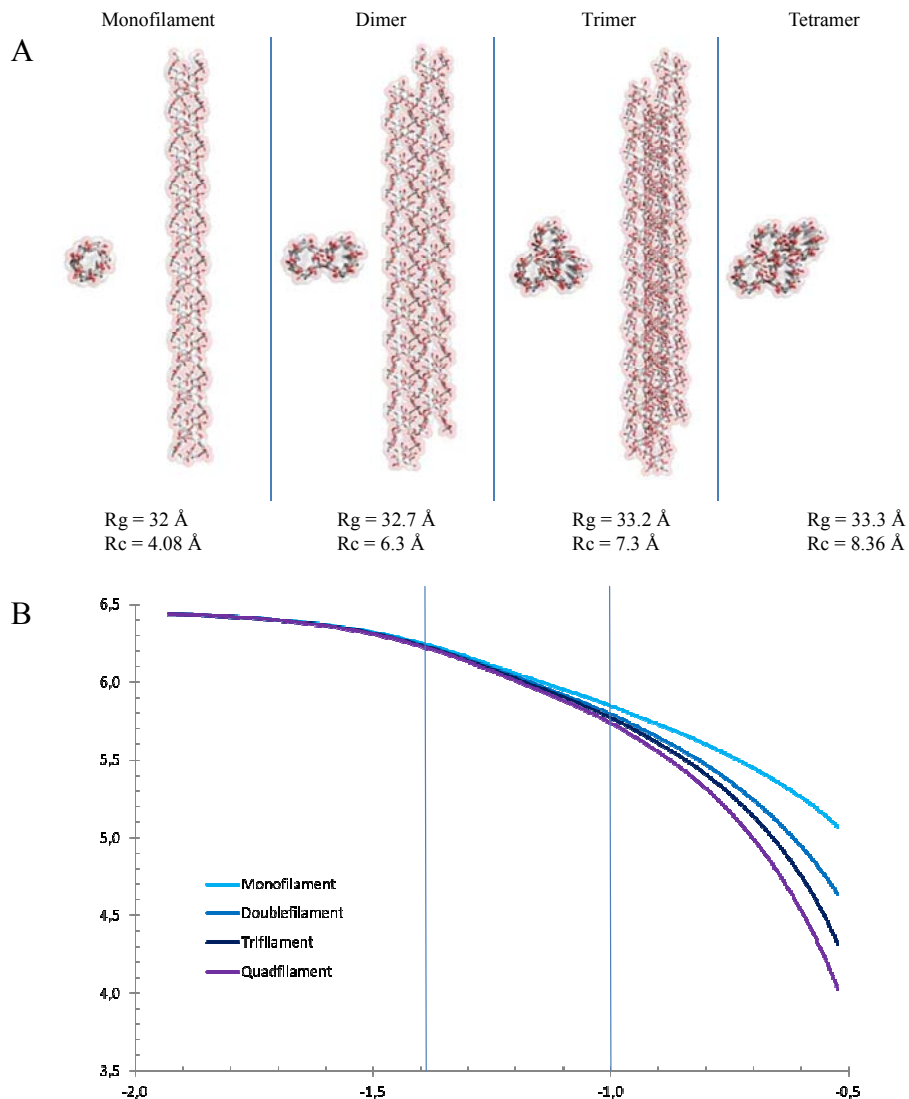


Figure 6: Models of amylose

Panel A. Structure views of amylose models

Panel B. SAXS curve of amylose models

The case of the reaction performed with DgAS is different. Indeed, sucrose consumption was slower compared to the NpAS catalysed reaction, and total substrate consumption was still not reached after 18 h reaction, when SAXS measurements were stopped. Such a low substrate consumption was already observed for DgAS [10], even if the initial activity introduced in the reaction medium was higher than that used for NpAS. We exclude that it

could be due to enzyme aggregation and inactivation during reaction. Indeed, we checked by SAXS measurements of control samples containing DgAS alone at each sampling times that no aggregation occurred. We thus rather speculate that sucrose could be re-synthesized during reaction by glucosyl transfer onto the fructose that is released in the medium. This correlates with the differences observed between DgAS and NpAS regarding fructosyl moiety accommodation in their +1 subsite [17]. However, despite this low sucrose conversion rate, the DgAS polymerization yield reached after 18 h reaction similar values than for NpAS after 14 h reaction, as 11,671 mg.L⁻¹ of polymer were produced.

The SAXS responses obtained for entire reaction media (Figure 4B) and for polymers (5B) during DgAS synthesis are different from those recorded for NpAS reactions. Indeed, for DgAS the three SAXS curves recorded after 6, 11 and 18 h of reaction have similar shapes with $I(Q)$ increasing with synthesis time, due to increasing polymer concentration (Figure 5B). Until 18 h of DgAS reaction, $I(0)$ value at low Q increased to reach the $I(0)$ value obtained for NpAS after only 2.7 h of reaction. R_g and R_c remain constant at 30 Å and 4.5 Å respectively throughout the whole reaction with values similar to those obtained with NpAS after 2.7 h reaction. However, α -glucan concentration obtained after 18 h of DgAS reaction was 12.6 times higher that obtained after 2.7 h with NpAS. As the value of $I(0)$ is directly correlated to the size and concentration of the objects in solution, this indicates that the mean size of the α -glucan particles produced by DgAS is smaller than that of particles produced by NpAS. On the other hand, in the region of intermediate Q , the curves do not follow a power law. This suggests that amylose synthesized by DgAS does not self associate under the form of specific structure such as the rod-like particles observed with NpAS.

These phenomena are probably due to the higher temperature used for polymerisation by DgAS synthesis (50 °C) than by NpAS (30 °C), favoring higher amylose solubility and preventing from important entanglement or self-association. This is also in agreement with the size distribution of the α -glucan chains produced by NpAS and DgAS from 100 mM sucrose, which is known to differ slightly. Indeed, chains of DP>30 are less concentrated in the reaction medium of DgAS in comparison to that of NpAS [10]. Finally, these data correlate with the fact that it is not possible to recover any water insoluble amylose fraction by centrifugation of the final DgAS reaction medium, when amylose synthesis is performed from 100 mM sucrose. For comparison, the insoluble amylose fraction obtained with NpAS in the same conditions (except the synthesis temperature used) reaches 2.9 g.L⁻¹ [9]. Finally, due to higher synthesis temperature limiting amylose self-association and aggregation, and to the

usual polydispersity of polymers synthesized by amylosucrases [9,10], a mixture of particles having worm like, cylindrical and statistical coils shape could be present in the DgAS reaction medium, which prevents from determining specific parameters as persistence or contour lengths.

Conclusion

This paper presents a SAXS study of amylose synthesis by two amylosucrases, NpAS and DgAS. For the first time, both conformations of enzyme during catalysis and of reaction products were characterized simultaneously.

We first demonstrated that the conformation in solution of the two amylosucrases NpAS and DgAS corresponds to their respective monomeric and dimeric crystallographic structures, respectively. SAXS analysis allowed to show that these overall conformations are conserved when the enzymes are complexed with substrates. Such experiment is unique, as SAXS analysis of protein conformation is being usually performed using substrate analogous on which enzymes are not active, or with inactive enzyme mutants.

In addition, for the first time, it was possible to access the first stages of polysaccharide chain entanglement and to deduce its conformation by subtracting the contribution of other compounds that are present in the synthesis medium to the SAXS scattering. The amylose chains produced by NpAS were shown to first associate into isolated single helices with R_c close to 4 Å, and then into assemblies containing 4 double helices ($R_c = 8$ Å) before aggregation and crystallization. This was not the case for the polymer produced by DgAS, which never entangles into such assemblies during the synthesis, probably because a higher synthesis temperature was used for this enzyme, favoring higher amylose solubility and limiting chain entanglement and intermolecular association.

Such results are promising for in depth study of semi-crystalline polymer enzymatic synthesis. The use of a wider range of Q for SAXS experiments would allow investigating the mechanisms involved in polymer crystallization and larger particle formation that occur after first stages of entanglement.

References

- 1 Buléon A, Colonna P, Planchot V & Ball S (1998) Starch granules: structure and biosynthesis. *Int. J. Biol. Macromol.* **23**, 85-112.
- 2 Buléon A, Véronèse G & Putaux J (2007) Self-Association and Crystallization of Amylose. *Aust. J. Chem.* **60**, 706-718.
- 3 Wattebled F, Buléon A, Bouchet B, Ral J-P, Liénard L, Delvallé D, Binderup K, Dauvillée D, Ball S & D'Hulst C (2002) Granule-bound starch synthase I. A major enzyme involved in the biogenesis of B-crystallites in starch granules. *Eur. J. Biochem.* **269**, 3810-3820.
- 4 Kumar A & Sanwal GG (1982) Purification and physicochemical properties of starch phosphorylase from young banana leaves. *Biochemistry* **21**, 4152-4159.
- 5 Potocki de Montalk G, Remaud-Siméon M, Willemot RM, Sarçabal P, Planchot V & Monsan P (2000) Amylosucrase from *Neisseria polysaccharea*: novel catalytic properties. *FEBS Lett.* **471**, 219-223.
- 6 Cantarel BL, Coutinho PM, Rancurel C, Bernard T, Lombard V & Henrissat B (2009) The Carbohydrate-Active EnZymes database (CAZy): an expert resource for Glycogenomics. *Nucleic Acids Res.* **37**, D233-238.
- 7 André I, Potocki-Véronèse G, Morel S, Monsan P & Remaud-Siméon M (2010) Sucrose-utilizing transglucosidases for biocatalysis. *Top Curr Chem* **294**, 25-48.
- 8 Albenne C, Skov LK, Mirza O, Gajhede M, Feller G, D'Amico S, André G, Potocki-Véronèse G, van der Veen BA, Monsan P & Remaud-Simeon M (2004) Molecular basis of the amylose-like polymer formation catalyzed by *Neisseria polysaccharea* amylosucrase. *J. Biol. Chem.* **279**, 726-734.
- 9 Potocki-Veronese G, Putaux J-L, Dupeyre D, Albenne C, Remaud-Siméon M, Monsan P & Buléon A (2005) Amylose synthesized in vitro by amylosucrase: morphology, structure, and properties. *Biomacromolecules* **6**, 1000-1011.
- 10 Emond S, Mondeil S, Jaziri K, André I, Monsan P, Remaud-Siméon M & Potocki-Véronèse G (2008) Cloning, purification and characterization of a thermostable amylosucrase from *Deinococcus geothermalis*. *FEMS Microbiol. Lett.* **285**, 25-32.
- 11 Skov LK, Mirza O, Henriksen A, Potocki de Montalk G, Remaud-Siméon M, Sarçabal P, Willemot RM, Monsan P & Gajhede M (2000) Crystallization and preliminary X-ray studies of recombinant amylosucrase from *Neisseria polysaccharea*. *Acta Crystallogr. D Biol. Crystallogr.* **56**, 203-205.
- 12 Skov LK, Mirza O, Henriksen A, De Montalk GP, Remaud-Siméon M, Sarçabal P, Willemot RM, Monsan P & Gajhede M (2001) Amylosucrase, a glucan-synthesizing enzyme from the alpha-amylase family. *J. Biol. Chem.* **276**, 25273-25278.
- 13 Skov LK, Mirza O, Sprogøe D, Dar I, Remaud-Siméon M, Albenne C, Monsan P & Gajhede M (2002) Oligosaccharide and sucrose complexes of amylosucrase. Structural implications for the polymerase activity. *J. Biol. Chem.* **277**, 47741-47747.
- 14 Skov LK, Mirza O, Sprogøe D, van der VEEN BA, Remaud-Siméon M, Albenne C, Monsan P & Gajhede M Crystal structure of the Glu328Gln mutant of *Neisseria polysaccharea* amylosucrase in complex with sucrose and maltoheptaose. *Biocatalysis and biotransformation* **24**, 99-105.
- 15 Mirza O, Skov LK, Remaud-Siméon M, Potocki de Montalk G, Albenne C, Monsan P & Gajhede M (2001) Crystal structures of amylosucrase from *Neisseria polysaccharea* in complex with D-glucose and the active site mutant Glu328Gln in complex with the natural substrate sucrose. *Biochemistry* **40**, 9032-9039.

- 16 Jensen MH, Mirza O, Albenne C, Remaud-Siméon M, Monsan P, Gajhede M & Skov LK (2004) Crystal structure of the covalent intermediate of amylosucrase from *Neisseria polysaccharea*. *Biochemistry* **43**, 3104-3110.
- 17 Guérin F, Barbe S, Pizzut-Serin S, Potocki-Véronèse G, Guieysse D, Guillet V, Monsan P, Mourey L, Remaud-Siméon M, André I & Tranier S (2011) Structural investigation of the thermostability and product specificity of amylosucrase from the bacterium *Deinococcus geothermalis*. *J. Biol. Chem.* **287**, 6642-6654
- 18 Koch MH, Vachette P & Svergun DI (2003) Small-angle scattering: a view on the properties, structures and structural changes of biological macromolecules in solution. *Q. Rev. Biophys.* **36**, 147-227.
- 19 Yuguchi Y (2009) Nano-Structure Analysis of Sugar Chains by Small Angle X-ray Scattering. *Trends in Glycoscience and Glycotechnology* **21**, 1-12.
- 20 Dogsa I, Štrancar J, Laggner P & Stopar D (2008) Efficient modeling of polysaccharide conformations based on Small-Angle X-ray Scattering experimental data. *Polymer* **49**, 1398-1406.
- 21 Quinn FX, Hatakeyama T, Takahashi M & Hatakeyama H (1994) The effect of annealing on the conformational properties of xanthan hydrogels. *Polymer* **35**, 1248-1252.
- 22 Maeda H, Rambone G, Coviello T, Yuguchi Y, Urakawa H, Alhaique F & Kajiwara K (2001) Low-degree oxidized scleroglucan and its hydrogel. *Int. J. Biol. Macromol.* **28**, 351-358.
- 23 Weinbreck F, Tromp RH & de Kruif CG (2004) Composition and structure of whey protein/gum arabic coacervates. *Biomacromolecules* **5**, 1437-1445.
- 24 Braga, Ferracini E, Ferrero A, Ripamonti A, Brant DA, Buliga G & Cesàro A (1985) Amylose Conformation in Aqueous Solution: A Small-Angle X-Ray Scattering Study. *International Journal of Biological Macromolecules* **7**, 161-166.
- 25 Müller JJ, Gernat C, Schulz W, Müller E, Vorweg W & Damaschun G (1995) Computer simulations of x-ray scattering curves: Gelation and crystallization process in amylose solutions. *Biopolymers* **35**, 271-288.
- 26 Shimada J, Kaneko H, Takada T, Kitamura S & Kajiwara K (2000) Conformation of Amylose in Aqueous Solution: Small-Angle X-ray Scattering Measurements and Simulations. *Journal of Physical Chemistry B* **104**, 2136-2147.
- 27 I' Anson KJ, Miles MJ, Morris VJ, Ring SG & Nave C (1988) A study of amylose gelation using a synchrotron X-ray source. *Carbohydrate Polymers* **8**, 45-53.
- 28 Gernat C, Reuther F, Damaschun G & Schierbaum F (1987) Real structure of a partially crystalline maltodextrin gel as revealed by wide-angle x-ray scattering. *Acta Polymerica* **38**, 603-607.
- 29 De Montalk GP, Remaud-Simeon M, Willemot RM, Planchot V & Monsan P (1999) Sequence analysis of the gene encoding amylosucrase from *Neisseria polysaccharea* and characterization of the recombinant enzyme. *J. Bacteriol.* **181**, 375-381.
- 30 Sumner JB & Howell SF (1935) A method for determination of saccharase activity. *The Journal of Biological Chemistry*, 51-54.
- 31 David G & 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. *Journal of Applied Crystallography* **42**, 892-900.
- 32 Guinier A (1939) La diffraction des rayons X aux tres petits angles: applications a l'étude de phenomenes ultramicroscopiques. *Ann. Phys.* **12**, 161-237

-
- 33 Rawiso M (1999) De l'intensité à la structure en physico-chimie des polymères. *Le Journal de Physique IV* **09**, 49.
- 34 Svergun D, Barberato C & Koch MHJ (1995) CRY SOL—a Program to Evaluate X-ray Solution Scattering of Biological Macromolecules from Atomic Coordinates. *Journal of Applied Crystallography* **28**, 768-773.
- 35 Franke D & Svergun DI (2009) DAMMIF, a program for rapid ab-initio shape determination in small-angle scattering. *Journal of Applied Crystallography* **42**, 342-346.
- 36 Imberty A & Perez S (1988) A revisit to the three-dimensional structure of B-type starch. *Biopolymers* **27**, 1205-1221.
- 37 Popov D, Buléon A, Burghammer M, Chanzy H, Montesanti N, Putaux J-L, Potocki-Véronèse G & Riekkel C (2009) Crystal Structure of A-amylose: A Revisit from Synchrotron Microdiffraction Analysis of Single Crystals. *Macromolecules* **42**, 1167-1174.

Table des illustrations

Liste des figures

<i>Figure 1 : Représentation schématique des différentes réactions catalysées par les amylosaccharases</i>	32
<i>Figure 2 : Représentation tridimensionnelle des résidus d'acides aminés conservés dans les sous-sites -1 et +1 des enzymes de la famille GH13</i>	35
<i>Figure 3 : Résidus conservés du tonneau catalytique des GH13</i>	36
<i>Figure 4 : Séquences consensus de GH13 aux activités différentes</i>	37
<i>Figure 5 : Mécanisme réactionnel des glycoside-hydrolases et transglycosidases à rétention de configuration</i>	39
<i>Figure 6 : Comparaison des produits formés par l'ASNp et l'ASDg en fonction des conditions réactionnelles</i>	48
<i>Figure 7 : Formules semi-développées des isomères du saccharose produits par les amylosaccharases</i>	49
<i>Figure 8 : Organisation tridimensionnelle de l'ASNp (code PDB : 1G5A)</i>	53
<i>Figure 9 : Sites de fixation du saccharose dans l'ASNp</i>	55
<i>Figure 10 : Site actif de l'ASNp dont le résidu catalytique D286 est lié de façon covalente au résidu β-D-glucosyle (code PDB : 1S46)</i>	57
<i>Figure 11 : Représentation schématique de la poche catalytique</i>	58
<i>Figure 12 : Sites de fixation des oligosaccharides dans l'ASNp</i>	59
<i>Figure 13 : Arrimage moléculaire des produits cibles dans le site actif de l'ASNp</i>	70
<i>Figure 14 : Bilan de la glucosylation des accepteurs cibles par les monomutants</i>	71
<i>Figure 15 : Lumière polarisée</i>	77
<i>Figure 16 : Spectre d'absorption dans l'UV lointain en fonction de différents types de structure secondaire</i>	79
<i>Figure 17 : Spectre d'absorption dans le proche UV</i>	80
<i>Figure 18 : Courbe de dénaturation</i>	81
<i>Figure 19 : Les 14 réseaux de Bravais</i>	86
<i>Figure 20 : Diagramme de phase</i>	87

<i>Figure 21 : Construction de la sphère d'Ewald</i>	90
<i>Figure 22 : Diagramme de Ramachandran</i>	97
<i>Figure 23 : Dynamique moléculaire</i>	100
<i>Figure 24 : Illustration schématique d'un microcolarimètre</i>	101
<i>Figure 25 : Exemple d'un résultat obtenu</i>	102
<i>Figure 26 : Exemple de résultats en fonction de l'affinité</i>	104
<i>Figure 27: Reactions catalyzed by amylosucrases from sole sucrose</i>	128
<i>Figure 28: Comparison of the three dimensional architecture of DgAS and NpAS</i>	136
<i>Figure 29: Comparison of DgAS and NpAS crystallographic structures</i>	138
<i>Figure 30: Quaternary structure organization of DgAS</i>	140
<i>Figure 31: Biochemical analysis of DgAS and NpAS</i>	142
<i>Figure 32: Stereoview of turanose conformation bound to NpAS and DgAS</i>	148
<i>Figure 33: Reactions catalyzed by amylosucrases from sole sucrose</i>	169
<i>Figure 34: Comparison of malto-oligosaccharide synthesis by wild-type amylosucrase (wtAS), F290K, A289P-F290C, A289P-F290I and A289P-F290L mutants from 250 mM sucrose supplemented or not with 250 mM allyl-2-acetamido-2-deoxy-α-D-glucopyranoside (D')</i>	180
<i>Figure 35: Comparison of the three-dimensional architecture of wild-type NpAS and NpAS mutants</i>	184
<i>Figure 36: Molecular docking of ligands (sucrose and ED') in the active site of wild-type amylosucrase in green and one double-mutant chosen for illustration, A289P-F290L</i> 187	
<i>Figure 37: Superimposition of wild-type NpAS and double-mutant structures at the beginning (0 ns, in grey) and at the end (20 ns) of the Molecular Dynamics simulation</i>	189
<i>Figure 38 : Exemple de résultats obtenus par DSF sur du tampon Tris 50 mM en fonction de la concentration en NaCl et en fonction de la concentration en enzyme</i>	212
<i>Figure 39 : Exemple de thermogramme obtenu (essai 2) avec une concentration en protéine de 420μM et un rapport saccharose/E328Q-ASNp : 20</i>	215

Liste des figures supplémentaires

<i>Supplementary Figure S1: Quantitative measurement of DgAS (purple) and NpAS (green) movements during the course of 60 ns MD simulations carried out in explicit water, starting from the X-ray structures (DgAS PDB code: 3UCQ and NpAS PDB code: 1G5A)</i>	151
<i>Supplementary Figure S2: View of the main interactions involved in the dimer interface of DgAS</i>	152
<i>Supplementary Figure S3: Sequence alignment of amylosucrases</i>	153
<i>Supplementary Figure S4: Stereoview of the glucosyl moiety of sucrose or turanose</i>	155
<i>Supplementary Figure S5: Stereoview of TB2 turanose binding site at the surface of the NpAS domain C</i>	156
<i>Supplementary Figure S6: Graphical representation of T_m values of amylosucrases determined by differential scanning fluorimetry</i>	193
<i>Supplementary Figure S7: Comparison of the X-ray structures of NpAS F290K-E328Q and NpAS E328Q in complex with sucrose</i>	194
<i>Supplementary Figure S8: View of bound sucrose in the F290K-E328Q mutant active site</i>	195
<i>Supplementary Figure S9: Analysis of 20 ns MD simulations performed on wild-type NpAS and mutants in their apo or sucrose-bound forms in explicit water, starting from their respective X-ray structures</i>	196
<i>Supplementary Figure S10: Conformation rearrangements occurring during the course of 20 ns molecular dynamics simulations carried out in explicit water</i>	197

Liste des tableaux

<i>Tableau 1 : Les différentes enzymes constituant la famille 13 des Glycoside-Hydrolases ..</i>	34
<i>Tableau 2 : Conservation de séquences entre les différentes amylosaccharases caractérisées à ce jour</i>	45
<i>Tableau 3 : Bilan des amylosaccharases recombinantes caractérisées à ce jour</i>	46

<i>Tableau 4 : Caractéristiques biochimiques des amylosaccharases caractérisées à ce jour</i>	47
<i>Tableau 5 : Tableau récapitulatif des structures d'ASNp déposées dans la PDB (Octobre 2011) et de leurs conditions de cristallisation</i>	52
<i>Tableau 6 : Activité relative et profil des produits obtenus à partir des différents mutants de l'ASNp</i>	60
<i>Tableau 7 : Température de dénaturation de l'ASNp I228Y en fonction de différents additifs</i>	214
<i>Tableau 8 : Résultats préliminaires des expériences d'ITC réalisées à partir de saccharose et du mutant E328Q de l'ASNp</i>	216

Liste des Tables

<i>Table 1: Data collection and refinement statistics</i>	133
<i>Table 2: Comparison of turanose and trehalulose formation by DgAS and NpAS expressed as the percentage of the glucosyl units coming from sucrose transferred onto fructose</i>	144
<i>Table 3: Interactions of fructose unit from sucrose and turanose with NpAS and DgAS</i>	146
<i>Table 4: Data collection and refinement statistics</i>	175
<i>Table 5: Results of the pH screening of amylosucrase sub-libraries for sucrose utilization</i>	178
<i>Table 6: Improved amylosucrase mutants identified for α-D-Glcp-(1\rightarrow4)-D-GlcpNAc (ED) synthesis</i>	179
<i>Table 7: Kinetic parameters determination of wild-type amylosucrase and improved mutants for sucrose donor (error range is within 5 to 10 %)</i>	181
<i>Table 8: Kinetic parameters determination for wild-type amylosucrase and improved mutants for allyl-2-acetamido-2-deoxy-α-D-glucopyranoside acceptor (D') using a fixed 250 mM sucrose concentration (error range is within 5 to 10 %)</i>	182
<i>Table 9: Average T_m of amylosucrases determined by differential scanning fluorimetry</i>	183

Liste des tables supplémentaires

<i>Supplementary Table S1: Interactions of glucosyl moieties from sucrose and turanose in complex with NpAS and DgAS</i>	150
<i>Supplementary Table S2: Binding interactions of sucrose with E328Q and F290K-E328Q mutants of amylosucrase</i>	192
<i>Supplementary Table S3: Degenerate primers used for the construction of libraries 2 to 3 (A289-F290 and I228-F229, respectively)</i>	192