



# Flavonoid glucodiversification with engineered sucrose-active enzymes

Yannick Malbert

## ► To cite this version:

Yannick Malbert. Flavonoid glucodiversification with engineered sucrose-active enzymes. Biotechnology. INSA de Toulouse, 2014. English. NNT : 2014ISAT0038 . tel-01219406

**HAL Id: tel-01219406**

**<https://theses.hal.science/tel-01219406>**

Submitted on 22 Oct 2015

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



# 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 :**

**Yannick MALBERT**

le jeudi 10 juillet 2014

**Titre :**

Flavonoid glucodiversification with engineered sucrose-active enzymes

---

**École doctorale et discipline ou spécialité :**

ED SEVAB : Ingénieries microbienne et enzymatique

**Unité de recherche :**

Unité de recherche : LISBP (UMR CNRS 5504, UMR INRA 792), INSA de Toulouse

**Directeur/trice(s) de Thèse :**

Pr. Magali Remaud-Siméon

**Jury :**

Pr. Stéphane Quideau (Université Bordeaux 1), rapporteur

Pr. Eva Nordberg-Karlsson (Lund University, Sweden), rapporteur

Dr. François Lefoulon (Les Laboratoires SERVIER, Orléans), examinateur

Dr. Sandrine Morel (Maître de Conférences, LISBP, INSA de Toulouse)

Pr. Magali Remaud-Siméon (LISBP, INSA de Toulouse)

Pr. Charles Tellier (Université de Nantes), président de jury



**Last name:** MALBERT

**First name:** Yannick

**Title: Flavonoid glucodiversification with engineered sucrose-active enzymes**

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

**Year:** 2014

**Number of pages:** 257

---

Flavonoid glycosides are natural plant secondary metabolites exhibiting many physicochemical and biological properties. Glycosylation usually improves flavonoid solubility but access to flavonoid glycosides is limited by their low production levels in plants. In this thesis work, the focus was placed on the development of new glucodiversification routes of natural flavonoids by taking advantage of protein engineering. Two biochemically and structurally characterized recombinant transglucosylases, the amylosucrase from *Neisseria polysaccharea* and the  $\alpha$ -(1 $\rightarrow$ 2) branching sucrose, a truncated form of the dextransucrase from *L. Mesenteroides* NRRL B-1299, were selected to attempt glucosylation of different flavonoids, synthesize new  $\alpha$ -glucoside derivatives with original patterns of glucosylation and hopefully improved their water-solubility. First, a small-size library of amylosucrase variants showing mutations in their acceptor binding site was screened in the presence of sucrose (glucosyl donor) and luteolin acceptor. A screening procedure was developed. It allowed isolating several mutants improved for luteolin glucosylation and synthesizing of novel luteolin glucosides, which exhibited up to a 17,000-fold increase of solubility in water. To attempt glucosylation of other types of flavonoids, the  $\alpha$ -(1 $\rightarrow$ 2) branching sucrose, naturally designed for acceptor reaction, was preferred. Experimental design and Response Surface Methodology were first used to optimize the production of soluble enzyme and avoid inclusion body formation. Five parameters were included in the design: culture duration, temperature and concentrations of glycerol, lactose inducer and glucose repressor. Using the predicted optimal conditions, 5740 U. L<sup>-1</sup><sub>of culture</sub> of soluble enzyme were obtained in microtiter plates, while no activity was obtained in the soluble fraction when using the previously reported method of production. A structurally-guided approach, based on flavonoids monoglucosides docking in the enzyme active site, was then applied to identify mutagenesis targets and generate libraries of several thousand variants. They were screened using a rapid pH-based screening assay, implemented for this purpose. This allowed sorting out mutants still active on sucrose that were subsequently assayed for both quercetin and diosmetin glucosylation. A small set of 23 variants, constituting a platform of enzymes improved for the glucosylation of these two flavonoids was retained and evaluated for the glucosylation of a six distinct flavonoids. Remarkably, the promiscuity generated in this platform allowed isolating several variants much more efficient than the wild-type enzyme. They produced different glucosylation patterns, and provided valuable information to further design and improve flavonoid glucosylation enzymatic tools.

---

**Keywords:** Flavonoid glycosylation, glucansucrase, transglucosylase, Protein engineering, Antioxidant, Glucoconjugates diversification, amylosucrase,  $\alpha$ -(1 $\rightarrow$ 2) branching sucrose, screening, response surface methodology, enzyme expression optimization.

---

**Doctoral school:** SEVAB (Ecological, Veterinary, Agronomic Sciences and Bioengineering)

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



**NOM:** MALBERT

**Prénom:** Yannick

**Titre: Glucodiversification des flavonoïdes par ingénierie d'enzymes actives sur saccharose**

*Spécialité: Sciences Ecologiques, Vétérinaires, Agronomiques et Bioingénieries, Filière : Ingénierie microbienne et enzymatique.*

**Année:** 2014

**Nombre de pages:** 257

Les flavonoïdes glycosylés sont des métabolites secondaires d'origine végétale, qui présentent de nombreuses propriétés physico-chimiques et biologiques intéressantes pour des applications industrielles. La glycosylation accroît généralement la solubilité de ces flavonoïdes mais leurs faibles niveaux de production dans les plantes limitent leur disponibilité. Ces travaux de thèse portent donc sur le développement de nouvelles voies de gluco-diversification des flavonoïdes naturels, en mettant à profit l'ingénierie des protéines. Deux transglucosylases recombinantes, structurellement et biochimiquement caractérisées, l'amylosaccharase de *Neisseria polysaccharea* et la glucane-saccharase de branchement  $\alpha$ -(1→2), forme tronquée de la dextran-saccharase de *L. Mesenteroides* NRRL B-1299, ont été sélectionnées pour la biosynthèse de nouveaux flavonoïdes, possédant des motifs originaux d' $\alpha$ -glycosylation, et potentiellement une solubilité accrue dans l'eau. Dans un premier temps, une librairie de petite taille de mutants de l'amylosaccharase, ciblée sur le site de liaison à l'accepteur, a été criblée en présence de saccharose (donneur d'unité glycosyl) et de lutéoline comme accepteur. Une méthode de screening a donc été développée, et a permis d'isoler des mutants améliorés pour la synthèse de nouveaux glucosides de lutéoline, jusqu'à 17000 fois plus soluble dans l'eau que la lutéoline aglycon. Afin de glycosyler d'autres flavonoïdes, la glucane-saccharase de branchement  $\alpha$ -(1→2), a été préférentiellement sélectionnée. Des plans expérimentaux alliés à une méthodologie en surface de réponse ont été réalisés pour optimiser la production de l'enzyme sous forme soluble et éviter la formation de corps d'inclusion. Cinq paramètres ont été ainsi analysés : le temps de culture, la température, et les concentrations en glycérol, lactose (inducteur) et glucose (répresseur). En appliquant les conditions optimales prédites, 5740 U.L<sup>-1</sup> de culture d'enzyme soluble ont été produites en microplaques, alors qu'aucune activité n'était retrouvée dans la fraction soluble, lors de l'utilisation de la méthode de production précédemment utilisée. Finalement, Une approche de modélisation moléculaire, structurellement guidée par l'arrimage de flavonoïdes monoglucosylés dans le site actif de l'enzyme, a permis d'identifier des cibles de mutagenèse et de générer des bibliothèques de quelques milliers de variants. Une méthode rapide de criblage sur milieu solide, basée sur la visualisation colorimétrique d'un changement de pH, a été mise au point. Les mutants encore actifs sur saccharose ont été sélectionnés puis analysés sur leur capacités à glycosyler la quercétine et la diosmétine. Une petite série de 23 mutants a ainsi été retenue comme plate-forme d'enzymes améliorées dédiées à la glycosylation de flavonoïdes et a été évaluée pour la glycosylation de six flavonoïdes distincts. La promiscuité, remarquablement générée dans cette plateforme, a permis d'isoler quelques mutants beaucoup plus efficaces que l'enzyme sauvage, produisant des motifs de glycosylation différents et fournissant des informations intéressantes pour le design et l'amélioration des outils enzymatiques de glycosylation des flavonoïdes.

**Mots Clés:** Glycosylation, flavonoïdes, glucan-saccharase, transglucosylase, ingénierie des protéines, anti-oxydant, diversification de glucoconjugués, amylosaccharase, glucane-saccharase de branchement  $\alpha$ -(1→2), criblage, méthodologie en réponse de surface, optimisation de production d'enzyme.

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

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

## PUBLICATIONS

---

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

**Optimizing the production of an  $\alpha$ -(1 $\rightarrow$ 2) branching sucrase in *Escherichia coli* using statistical design.** Malbert Y., Vuillemin M., Laguerre S., Remaud-Siméon M., and Moulis C., 2014, *Appl. Microbiol. Biotechnol., methods and protocol*, 1–12.

**Extending the structural diversity of alpha-flavonoid glycosides with engineered glucansucrases.** Malbert Y., Pizzut-Serin S., Massou S., Cambon M., Laguerre S., Monsan P., Lefoulon F., Morel S., André I., and Remaud-Simeon M., 2014, *ChemCatChem*, Accepted (DOI: 10.1002/cctc.201402144R1)

**Interactions between sugars and glucan binding domains: structural evidences from the truncated GH70 glucansucrase  $\Delta$ N<sub>123</sub>-GBD-CD2.** Brison Y., Malbert Y., Mourey L., Remaud-Siméon M. and Tranier S. *In preparation*.

## ORAL COMMUNICATIONS

---

**Biocatalytic production of flavonic gluco-conjugates.** Malbert Y., Pizzut-Serin S., Morel S., Laguerre S., Cambon E., André I., Lefoulon F., Monsan P. and Remaud-Simeon M., GDR CNRS 2825 "Nouvelles approches en évolution dirigée des enzymes", Ax-Les-Thermes (France), 11th-12th Octobre, 2011.

**Développement d'une méthodologie en surface de réponse pour la définition de conditions optimales de production d'une enzyme de branchement en  $\alpha$ -(1 $\rightarrow$ 2) récalcitrante à son expression chez *Escherichia coli*.** Malbert Y., Vuillemin M., Laguerre S., De la Bourdonnaye G., Moulis C. and Remaud-Siméon M., 4ème Congrès Pluridisciplinaire sur les Glycosciences (Glucidoc), Landéda (France), 8th-11th April, 2013.

**Enzymatic engineering through both random and semi-rational directed evolution to improve an  $\alpha$ -(1 $\rightarrow$ 2) branching glucansucrase.** Malbert Y., Pizzut-Serin S., Rivière V., Vuillemin M., Morel S., André I., Moulis C., and Remaud-Siméon M., Congrès de la Société Française de Biochimie et Biologie Moléculaire (SFBBM), Paris (France), 4th-6th Septembre, 2013.

**$\alpha$ -Glucosylation de flavonoïdes par des glucane-saccharases modifiées par ingénierie.** Morel S., Malbert Y., Brison Y., Remaud-Simeon M, 25e Colloque du Club Biocatalyse et Synthèse Organique (CBSO), Rouet (France), 3th-6th June, 2014.

## POSTER COMMUNICATIONS

---

**Screening of an engineered Amylosucrase library to improve polyphenol transglucosylation.** Malbert Y., Cambon E., Pizzut S., Serin, Morel S., Laguerre S., André I., Lefoulon F., Monsan P., and Remaud-Simeon M., Congrès de la Société Française de Biochimie et Biologie Moléculaire (SFBBM), Ax-Les-Thermes (France), 12th-14th Octobre, 2011.

**Enhanced soluble expression of an  $\alpha$ -(1 $\rightarrow$ 2) branching sucrase in *Escherichia coli* using a response surface methodology.** Malbert Y.\*, Vuillemin M.\*, Laguerre S., De la Bourdonnaye G., Moulis C. and Remaud-Siméon M., 10th Carbohydrate Bioengineering Meeting, Prague (Czech Republic), 21st – 25th April, 2013 (\*Both authors contributed equally).

**Enhanced soluble expression of an  $\alpha$ -(1 $\rightarrow$ 2) branching sucrase in *Escherichia coli* using a response surface methodology.** Malbert Y.\*, Vuillemin M.\*, Laguerre S., De la Bourdonnaye G., Moulis C. and Remaud-Siméon M., Congrès COSM'innov, Orléans (France), 8th-9th October, 2013.

**Screening of an engineered Amylosucrase library to improve polyphenol transglucosylation.** Malbert Y., Cambon E., Pizzut S., Serin, Morel S., Laguerre S., André I., Lefoulon F., Monsan P., and Remaud-Simeon M., 25e Colloque du Club Biocatalyse et Synthèse Organique (CBSO), Rouet (France), 3th-6th June, 2014.

## PATENTS

---

**Projet GLYCOFLAV - Modification à façon de flavonoïdes pour créer des composés à forte valeur ajoutée.** Morel S., André I., Brison Y., Cambon E., Malbert Y., Pompon D., Remaud-Siméon M., and Urban P., *Toulouse White Biotechnology, INRA UMS 1337*. Patent FR1452461, 2014.



# REMERCIEMENTS

*Une thèse, dans une vie, c'est un peu comme s'envoler pour un voyage initiatique de quelques années au cœur de la galaxie des sciences. Oui, un voyage dans une autre galaxie, car il arrive, parfois, voir souvent, à un thésard, de se sentir vêtu d'une peau d'extra-terrestre dans le regard de ses proches. En bien. Mais pas toujours... Mais peu importe, car ce voyage est tout sauf ennuyeux !*

*Il regorge de péripéties et de rebondissements aux parfums de sueur, de rire, de larme, de doute, d'espoir, de joie, bref, d'un tumulte de sentiments qui se mélangent dans le tourbillon des neurones en ébullition. Et le plus génial ? C'est qu'il nous entraîne vers des lieux encore inexplorés de l'univers des possibles. Et lorsqu'on en découvre un, même tout petit, l'émerveillement est total ! Et puis, lorsque l'on prend un peu de recul, on se rend compte de l'étendue gigantesque des autres possibles qui l'entourent, cachés sous la brume des inconnues. Alors, d'un coup, on comprend avec joie que ce voyage n'est que le début d'une grande aventure, de notre quête scientifique.*

*Et comme toute aventure digne de ce nom, elle ne se mène pas seule, loin de là. C'est pourquoi, je tiens à remercier de tout mon cœur mes guides et compagnons de routes, sans qui rien de tout cela n'aurait été possible :*

En premier lieu, je voudrais remercier **Magali Remaud-Siméon**, ma directrice de thèse. Magali, je n'aurais jamais suffisamment de mots pour te remercier de la confiance que tu m'as accordée, à un moment crucial de mon parcours et tout au long de cette thèse. De toute ces dimensions de réflexions scientifique, économique et humaine que tu as insufflées à ma vision de la recherche. Et puis, ta patience et ton soutien dans les moments tourmentés, ton écoute et ton ouverture face à mes idées (parfois farfelues), et tellement encore ... Tu m'as offert la chance de pouvoir travailler avec une grande autonomie tout en étant présente pour me protéger du risque de la dispersion. En résumé : Magali, je n'aurais jamais pu rêver d'une meilleure directrice de thèse.

Mais pour m'encadrer durant cette thèse, tu n'étais pas seule, et c'est pourquoi je voudrais également remercier, du fond du cœur, **Sandrine Morel** et **Claire Moulis**. Même si ce n'était pas stipulé sur le papier, vous m'avez encadré, toutes les deux, comme de vraies co-directrices de thèse. Ah, Sandrine, nos longues discussions sur ces flavonoïdes aussi faciles à manipuler qu'un bâton brûlant des deux côtés. Et, Claire, nos échanges et réflexions intenses sur l'ingénierie de cette sucrase de branchement tellement facile à cribler que je m'en suis contracté une kératite. Alors, je pourrais me répéter en milles remerciements, mais à quoi bon, car ceux que j'ai écrit précédemment pour Magali vous sont également destinés.

Un grand merci également à **Isabelle André**. Isabelle, tu m'as fait découvrir la modélisation moléculaire et tout ce qu'elle peut apporter à l'ingénierie enzymatique, que ce soit lorsque l'on cherchait à voir pourquoi la lutéoline était mieux glucosylée ou pour créer de nouveaux mutants spécifiques des flavonoïdes. Et ce n'était pas toujours à faire, mais grâce à toi c'est devenu possible.

Merci également à **Pierre Monsan**. Pierre, je te remercie de m'avoir ouvert les portes du laboratoire et pour ces partages scientifiques (et autres) que nous avons eus. Ton habilité à mettre en avant les points cruciaux est tout simplement incroyable, et reste pour moi une grande source d'inspiration.

Je voudrais aussi remercier l'ensemble des membres du jury pour avoir accepté de se pencher sur mon travail de thèse et d'en juger le contenu et la qualité. Un très grand merci,

**Cherles Tellier**, d'avoir accepté de présider ce jury. Merci, **Stéphane Quideau** et Eva **Nordberg-Karlsson**, pour avoir accepté d'évaluer ce manuscrit, pour l'intérêt que vous y avez porté ainsi que pour vos conseils avisés lors de la soutenance. Et pour cela, ainsi que nos échanges au cours de ce projet de thèse, je tiens également à remercier **François Lefoulon**.

Je tiens à remercier les Laboratoires Servier ainsi que Tate&Lyle pour les supports financiers accordés à cette thèse.

Je voudrais également remercier tout spécialement les personnes qui se sont impliquées à mes côtés dans ces travaux : **Sandra Pizzut-Serin** et **Marlène Vuillemin**. Sandra, je me rappelle, comme si c'était hier, de mon arrivée au labo, de ces premiers temps où tu m'as formé au fonctionnement de l'équipe, au DNS, etc... Jamais je n'oublierais cet esprit d'équipe qui en a découlé et nous a soudé pour faire suer haut et fort les robots (bon, moins que nous, je l'avoue) et obtenir de superbes flavonoïdes glucosylés. Marlène, je suis super fier de ce projet qu'on a réussi à faire naître, à cheval entre nos deux thèses, et qu'on a porté de nos petites mains jusqu'à sa publication. C'était complètement dingue, ces plans d'expériences nous faisant revenir à n'importe quelle heure du jour et de la nuit, en parallèle de nos autres travaux. C'était dur et épuisant, mais voilà, travailler avec quelqu'un d'ultra compétent et de fun, comme toi, c'était tout simplement génial. Au final, on a réussi à la produire au top, cette enzyme récalcitrante !

Merci également à **Stéphane Massou**, **Sandrine Laguerre**, **Yoann Brison**, **Nelly Monties**, **Anne-Claire Ornaghi** et **Virginie Rivière** pour leur contribution à ces travaux.

*Aïe, je suis bavard et la place me manque...*

C'est pourquoi je vais remercier plus brièvement dans les mots mais pas en pensées **l'ensemble des membres passés et présent de l'EAD1**, stagiaires, doctorants, postdoctorants et permanents avec qui j'ai partagé de nombreux moments. Merci à **Manue**, **Greg**, **le duo des Pablo(s)** et **le trio des Sophie(s)**, **Bastien**, **Yoann**, **Julien D.**, **Marc**, **Fred**, **David D.**, **Lisa**, **Seydou**, **Florent**, **Alizée**, **Marion C.**, **Florence**, **Juliette**, **Martial**, **Marion S.**, **Etienne**, **Letian**, **Beau**, **Tuck**, **Faten**, **Sabrina**, **Chris**, **Nathalie C**, **Mélusine**, **Pauline**, **Betty**, **Elisabeth**, **Hélène**, **Géraldine**, **Marie-Laure**, **Claire D**, **David G.**, **Régis**, **Peco**, **Cédric**...ainsi que **tous ceux que je n'ai pas cité**, à cause de ma mémoire parfois défaillante en cette fin de thèse.

Un grand merci général à **tout le LISBP**.

Merci à tous les membres du collectif des précaires de la recherche pour m'avoir intégré à leurs réflexions, et particulièrement à **Mathieu**, **Hélène** et **Marie**.

Un merci spécial à **Antoine** pour avoir sauvé mes données lorsque mon PC m'a brusquement lâché en fin de rédaction.

Je remercie **ma famille**, et plus particulièrement mes parents et ma sœur, ainsi que tous **mes amis** pour leur soutien et tous ces bons moments partagés.

Et bien entendu, avec mon amour infini et tout autant de remerciement pour m'avoir soutenu et supporté sans failles pendant cette thèse : ma femme et ma fille. **Morgane et Arwen**, je vous aime !!!

In memory of Pr. Daniel Thomas...

Research is to see what everybody else has seen,  
and to think what nobody else has thought.

**Albert Szent-Gyorgi**

The most exciting phrase to hear in science, the one that heralds new discoveries, is not  
'Eureka!' but 'That's funny...'

**Isaac Asimov**

Je suis responsable de chaque mot dans mes ouvrages.  
Je peux me tromper, mais je ne triche pas.  
J'explique ce que je crois avoir compris, raconte ce que je crois savoir. »

**Yasmina Khadra**

# **Table of content**





<b>General .....</b>	<b>17</b>
<b>Introduction.....</b>	<b>17</b>
<b>Chapter I. Literature Review .....</b>	<b>23</b>
<b>I. Generalities on flavonoids.....</b>	<b>25</b>
<b>I.1. Flavonoid structure and biosynthesis .....</b>	<b>26</b>
I.1.1. Biosynthesis of flavonoids .....	26
I.1.2. Flavonoid structures and classification.....	28
<b>I.2. Physicochemical and biological properties .....</b>	<b>30</b>
I.2.1. Antioxidant properties .....	30
I.2.1.a. Direct scavenging of Reactive Oxygen Species (ROS).....	31
I.2.1.b. Metal chelation .....	32
I.2.1.c. Activation of antioxidant enzyme gene expression .....	33
I.2.1.d. Inhibition of oxidases.....	33
I.2.2. UV radiations absorption and photo-protection activity.....	34
I.2.3. Inhibition of enzymes non-directly involved in oxidation .....	35
<b>I.3. Applications fields of flavonoids .....</b>	<b>37</b>
<b>I.4. Flavonoid solubility .....</b>	<b>39</b>
<b>I.5. Bioavailability of flavonoids .....</b>	<b>40</b>
<b>I.6. Modifications of flavonoids to improve their properties .....</b>	<b>43</b>
<b>II. Glycosylation of flavonoids .....</b>	<b>45</b>
<b>II.1. Natural glycosylation of flavonoids in plants.....</b>	<b>45</b>
II.1.1. General Considerations .....	45
II.1.2. Glycosyltransferases (GTs) overview.....	46
II.1.3. Glycosylation of flavonoids in plants .....	52
<b>II.2. Production of flavonoid glycosides using glycosyltransferases (GTs) .....</b>	<b>55</b>
II.2.1. Recombinant GTs for in vitro production of flavonoid glycosides .....	55
II.2.2. Metabolic engineering of whole-cell production systems .....	57
II.2.3. Enzymatic engineering of GTs .....	62
II.2.3.a. Site-directed mutagenesis .....	62
II.2.3.b. Domain swapping .....	64
II.2.3.c. Random directed evolution .....	65
<b>II.3. Glycoside-Hydrolases for flavonoid glycosylation.....</b>	<b>67</b>
II.3.1. Glycoside-Hydrolases (GH) short overview .....	67
II.3.2. Flavonoid Glycosylation.....	70
<b>III. Glucansucrases (GS) .....</b>	<b>76</b>
<b>III.1. Few generalities.....</b>	<b>76</b>
III.1.1. Origins and Classification .....	76
III.1.1.a. GH family 13.....	77
III.1.1.b. GH family 70 .....	78
III.1.2. Main reactions and product applications .....	78

III.1.3. <i>Np</i> As and GBD-CD2, the two GS templates of this study.....	80
III.1.3.a. Amylosucrase from <i>N. polysaccharea</i> ( <i>Np</i> AS).....	80
III.1.3.b. The $\alpha$ -(1→2) branching sucrose (GBD-CD2) .....	81
<b>III.2. Structure/function relationships studies .....</b>	<b>83</b>
III.2.1. Catalytic mechanism .....	83
III.2.2. Polymerization process .....	83
III.2.3. Sequences and 3D-structures .....	87
III.2.3.a. General structural organization.....	88
III.2.3.b. Structure-based analysis of glucansucrase active sites .....	92
<b>III.3. Glucansucrases: Attractive tools for glucodiversification .....</b>	<b>97</b>
III.3.1. <i>Np</i> As and $\alpha$ -(1→2) branching sucrose kinetic parameters .....	97
III.3.2. Engineering of glucansucrases .....	99
III.3.3. Glucosylation of flavonoids with glucansucrases .....	102
<b>IV. Thesis objectives .....</b>	<b>107</b>
<b>V. References.....</b>	<b>109</b>
 <b>Chapter II. Extending the structural diversity of <math>\square</math>-flavonoid glycosides with engineered glucansucrases .....</b>	 <b>131</b>
<b>I. Introduction .....</b>	<b>134</b>
<b>II. Results .....</b>	<b>136</b>
II.1. Library and design of the screening assay .....	136
II.2. Production and structural characterization of luteolin glucosides L1, L2 and L3 .....	139
II.3. Physico-chemical characterization of the glucosides.....	141
<b>III. Discussion .....</b>	<b>142</b>
<b>IV. Conclusion .....</b>	<b>146</b>
<b>V. Experimental Section.....</b>	<b>146</b>
V.1. Bacterial Strains, Plasmids, and Chemicals .....	146
V.2. Semi-automated procedure for <i>Np</i> AS single mutant library screening .....	147
V.3. Analytical methods .....	147
V.4. Production of recombinant <i>Np</i> AS and enzyme assay .....	148
V.5. Design of Response Surface Methodology (RSM) Experiment.....	149
V.6. Production of luteolin glucosides (L1, L2 and L3) .....	149
V.7. Nuclear Magnetic Resonance (NMR) analysis .....	150
V.8. Water Solubility, UV spectra and antioxidant properties of luteoline glucosides .....	150
V.9. Docking studies .....	151
V.10. Characterization of luteolin glucosides L2 and L3 .....	151
<b>VI. Acknowledgements .....</b>	<b>151</b>
<b>VII. References.....</b>	<b>152</b>
<b>VIII. Supplementary information.....</b>	<b>154</b>

## **Chapter III. Optimizing the production of an $\alpha$ -(1→2) branching sucrase in *Escherichia coli* using statistical design..... 161**

<b>I. Introduction .....</b>	<b>164</b>
<b>II. Results .....</b>	<b>165</b>
<b>II.1. Pre-selection of the bacterial host and expression vector .....</b>	<b>165</b>
<b>II.2. Optimizing the production of soluble <math>\Delta N_{123}</math>-GBD-CD2.....</b>	<b>166</b>
II.2.1. Model fitting and statistical analysis .....	166
II.2.2. 3D response surface plots and experimental validation.....	168
<b>II.3. Large-scale productions.....</b>	<b>170</b>
<b>II.4. Protein expression pattern .....</b>	<b>171</b>
<b>II.5. Affinity purification of recombinant protein .....</b>	<b>171</b>
<b>III. Discussion .....</b>	<b>173</b>
<b>IV. Materials and methods .....</b>	<b>175</b>
<b>IV.1. Bacterial strains and expression vectors.....</b>	<b>175</b>
<b>IV.2. Molecular cloning / Construction of expression clones .....</b>	<b>176</b>
<b>IV.3. Growth media .....</b>	<b>176</b>
<b>IV.4. Response surface design and statistical analyses .....</b>	<b>176</b>
<b>IV.5. Productions of recombinant <math>\Delta N_{123}</math>-GBD-CD2 .....</b>	<b>177</b>
IV.5.1. Starter cultures .....	177
IV.5.2. Culture in microplates .....	178
IV.5.3. Cultures in 5L-Erlenmeyer flask.....	178
<b>IV.6. Enzymatic assays .....</b>	<b>178</b>
<b>IV.7. Protein expression analysis .....</b>	<b>179</b>
IV.7.1. Electrophoresis gel.....	179
IV.7.2. Western blots .....	179
<b>IV.8. Affinity purification of recombinant protein.....</b>	<b>180</b>
IV.8.1. Affinity 6xHis tag chromatography. ....	180
IV.8.2. Affinity Strep tag II chromatography.....	180
<b>V. Acknowledgments .....</b>	<b>180</b>
<b>VI. References.....</b>	<b>181</b>
<b>VII. Supplementary material.....</b>	<b>183</b>

## **Chapter IV. Rational design of an $\alpha$ -(1→2) branching sucrase for flavonoid glucoside diversification..... 191**

<b>I. Introduction .....</b>	<b>194</b>
<b>II. Results and discussion.....</b>	<b>195</b>
<b>II.1. Semi-rational design and <math>\Delta N_{123}</math>-GBD-CD2 mutant libraries creation .....</b>	<b>195</b>
<b>II.2. Setting up the primary solid-state colorimetric screening assay .....</b>	<b>197</b>
II.2.1. General principle of the screening method.....	197
II.2.2. Development and validation of the screening method .....	198
II.2.3. Screening of the libraries .....	200

<b>II.3. Secondary screening: diosmetin and quercetin glucosylation.....</b>	<b>201</b>
II.3.1. Analysis of the conversion rates.....	202
II.3.2. Analysis of the product profile and sub-selection of the mutants.....	202
<b>II.4. <math>\Delta N_{123}</math>-GBD-CD2 best mutants as a flavonoid general glucosylation platform .....</b>	<b>207</b>
II.4.1. Conversion rate analyses .....	208
II.4.1.a. Reproducibility of the screening assays .....	208
II.4.1.b. Comparison of flavonoid conversion.....	208
II.4.2. Sucrose consumption analysis .....	210
II.4.3. Glucosylation profile analysis .....	211
II.4.3.a. Quercetin glucosylation profile (Figure 65).....	212
II.4.3.b. Luteolin glucosylation profile (Figure 66).....	213
II.4.3.c. Morin glucosylation profile (Figure 67) .....	214
II.4.3.d. Naringenin glucosylation profile (Figure 68) .....	215
<b>III. Conclusion.....</b>	<b>215</b>
<b>IV. Material and methods .....</b>	<b>218</b>
IV.1. Bacterial Strains, Plasmids, and Chemicals .....	218
IV.2. Production of recombinant $\alpha$ -(1 $\rightarrow$ 2) branching sucrase .....	218
IV.3. Sucrose enzymatic activity assay of the $\alpha$ -(1 $\rightarrow$ 2) branching sucrase .....	218
IV.4. Buffer and pH analyses .....	219
IV.5. Molecular modeling and docking studies .....	219
IV.6. Molecular biology for $\alpha$ -(1 $\rightarrow$ 2) branching sucrase libraries creation.....	220
IV.7. $\alpha$ -(1 $\rightarrow$ 2) branching sucrase mutant libraries primary screening .....	221
IV.8. Semi-automated procedure for $\alpha$ -(1 $\rightarrow$ 2) branching sucrase mutant libraries secondary screening.....	222
IV.9. Analytical methods .....	222
<b>V. References.....</b>	<b>222</b>
<b>Conclusion &amp; Prospects .....</b>	<b>225</b>
<b>Artworks &amp; Table .....</b>	<b>237</b>
<b>Contents .....</b>	<b>237</b>
I. Artwork content.....	239
II. Table content .....	241
<b>Résumé en Français.....</b>	<b>243</b>
<b>Abbreviations.....</b>	<b>255</b>

# **General Introduction**



Nowadays, sustainable development is a major concern of our society and industries. In this socio-economic context, the demand for new biosourced molecular structures produced via eco-aware industrial processes is incredibly increasing. These trends stimulate academic research efforts for the development of alternative biotransformation pathways of extracts and active ingredients from plant origin, such as flavonoids or sugars.

**Flavonoids are natural pigments widespread throughout the plant kingdom** where they act, among others, as UV filters and protective agents against oxidation and pathogens. They have been widely studied for their physicochemical and biological properties that make them very promising candidates to **develop of new active compounds of interest** for the pharmaceutical, cosmetic and food industry fields.

However, their **drastic lacks of solubility** in either aqueous or hydrophobic environments highly complicate their formulation, may limit their bioavailability and, consequently, the biologic effects expected. **In plants, glycosylation** is an efficient way to generate flavonoid derivatives, which usually present a **higher water-solubility** than their referent aglycons. Flavonoid glycosides are widespread in nature. However, they are usually produced in low amounts and their extraction is quite tedious. Chemical glycosylation of flavonoids could be a way to generate such structures but the complexity of sugar chemistry limits such a development. In the last decades, **enzymatic glycosylation of flavonoids** has attracted considerable interests. This alternative takes advantage of the **stereo and regioselectivity** of enzymes and enables the synthesis of novel compounds under environmentally friendly and mild conditions.

Among the glyco-enzymes that can be envisaged for flavonoid glycoside synthesis, **glucansucrases** are very attractive biocatalysts. Indeed, they are **promiscuous transglucosylases**, which use the cheap and renewable **sucrose as a donor of glucosyl units**. In addition, they can naturally transfer this glucosyl unit onto a broad variety of hydroxylated acceptors, including flavonoids. Nevertheless, their catalytic efficiency for the glucosylation of many flavonoids remains strongly limited by their specificity. To overcome such a limitation, **glucansucrases engineering** can provide effective solutions by reshaping their active sites to promote flavonoid accommodation in a productive mode of glucosylation.

In this thesis work, two biochemically and structurally characterized recombinant transglucosylases were selected to attempt **glucosylation of different flavonoids**, synthesize **new  $\alpha$ -glucoside derivatives with original patterns of glucosylation** and hopefully **improved their water-solubility**.



These transglucosylases are:

- **The amylosucrase from *Neisseria polysaccharea* (NpAS)**, belonging to glycoside-hydrolase family 13. This enzyme was the first structurally characterized glucansucrase and has already been engineered for several applications.
- **The  $\alpha$ -(1→2) branching sucrase**, which is a truncated form of the dextransucrase DSR-E from *Leuconostoc citreum* B1299, belonging to GH family 70 and very recently characterized.

In the **first chapter** of this thesis, we provide a **literature review** to set the context of our work. First, a general **overview of the flavonoids** is proposed to give insights in their biosynthesis, physicochemical properties, and potential uses as active compounds. In a second part, the focus is placed on their glycosylation as, naturally occurring in plants, or performed *in vitro* using enzyme-based process involving **glycosyltransferases of Leloir-type or glycoside hydrolases**. Finally, the transglucosylases and particularly the **NpAS and the  $\alpha$ -(1→2) branching sucrase**, are presented in details, including enzyme engineering studies and flavonoids glucosylation attempts.

The next chapters are dedicated to the research work conducted in this thesis, presented in the form of research articles.

**Chapter II** shows how it was possible to take advantage of one **NpAS library of 171 variants** that were mutated in their acceptor binding site, at subsite +1 and +2 (Cambon et al., 2014; Champion et al., 2009) to **glucosylate luteolin** and generate original derivatives with **improved water-solubility**. This work was supported by “Les Laboratoires Servier”.

Some NpAS mutants were shown to be very efficient to synthesize novel and remarkably water-soluble luteolin glucosides, but appeared far less effective for the glucosylation of other flavonoids. To overcome these limitations, another transglucosylase, naturally designed for acceptor reaction, the  $\alpha$ -(1→2) branching sucrase, was selected and evolved to generate catalysts able to efficiently glucosylate various flavonoid structures with original glucosylation patterns. One of the major limitations for the usage of this enzyme was due to the very low level of production in recombinant *E. coli*. The enzyme was mainly produced in inclusion bodies. That was not compatible with the development of miniaturized screening procedures.

**Chapter III** deals with the **optimization of the recombinant  $\alpha$ -(1→2) branching sucrase production** by *E. coli*. An **experimental design** approach using **Response Surface Methodology** was conducted and allowed producing very high level of soluble enzyme, enabling the implementation of screening assays.

Finally, **Chapter IV**, is focused on the  **$\alpha$ -(1→2) branching sucrase engineering**. It describes the **semi-rational strategy** which started with molecular modeling and docking experiments to proposed mutation targets. Libraries of 2 to 3,000 mutants were generated through saturation mutagenesis. **Specific screening assays** were implemented to isolate a small platform of improved mutants for the glucosylation of quercetin. Then, the **promiscuity of this platform was explored** by testing the variants for the glucosylation of **five other flavonoids**.

---

## References

- Cambon, E., Barbe, S., Pizzut-Serin, S., Remaud-Simeon, M., and André, I. (2014). Essential role of amino acid position 226 in oligosaccharide elongation by amylosucrase from *Neisseria polysaccharea*: Amylosucrase Engineering for High Amylose Production. *Biotechnol. Bioeng.* n/a–n/a.
- 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.



# **Chapter I**

## **Literature Review**



This first chapter is dedicated to the study of literature knowledge available at the beginning of this study and its recent updates. It allows exposing the context and highlighting the goal of our **study about the enzymatic glycosylation of flavonoids**. In a **first part, general considerations on flavonoids** are presented, such as their origin, structure, properties and applications. The **second part** of this chapter focuses on their ***in-vivo*, and *in-vitro* chemical and enzymatic glycosylation**. In the **third part, glucosylation of flavonoids, specifically catalyzed by glucansucrases** is developed. In this last part, an overview of glucansucrase enzymes is first presented. Then, flavonoid glucosylation attempts performed with glucansucrases are summarized. Finally, the potential of enzyme engineering for improving enzyme properties and, in particular, specificity toward unnatural substrates is exposed.

## I. Generalities on flavonoids

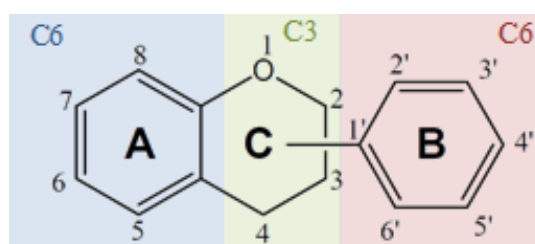
The flavonoids (the term is derived from the Latin word “flavus”, meaning yellow) were **first isolated in the 1930’s by Albert Szent-Györgyi** who was nominated Nobel laureate in Physiology and Medicine, in 1937, for “his discoveries in connection with the biological combustion process with special reference to vitamin C and the catalysis of fumaric acid” (Szent-Györgyi, 1963). At first, he suggested to name flavonoids “vitamin P” (Rusznayák and Szent-Györgyi, 1936) but this term was rapidly replaced by “bioflavonoids” as the isolated compound was revealed to be a mixture of various molecules. During the past 30 years, huge advances in the research field of bioflavonoids resulted in the identification and characterization of many individual polyphenolic compounds found in bioflavonoid mixtures which were finally classified under the general term “flavonoids”.

Flavonoids are natural products that belong to the **general class of polyphenol**. They are **low molecular weight secondary metabolites produced by plants** and generally described as non-essential for plant survival, unlike primary metabolites (Quideau et al., 2011). They occur in a large variety of fruits, vegetables, grains, nuts, stems, leaves, flowers and roots and play a critical role in plant growth and development (Harborne, 1973). With **more than 8,000 known structures**, flavonoids represent the most abundant secondary metabolites in plants (Corcoran et al., 2012; Ferrer et al., 2008; Quideau et al., 2011; Williams and Grayer, 2004). Recently, several databases have been developed to organize the knowledge acquired about flavonoids (Arita and Suwa, 2008; Bhagwat et al., 2013; Kinoshita et al., 2006; Moco et al., 2006). The most important one is proposed by the USDA (United States Department of Agriculture) and compiles the flavonoid content of many different foods according to the data available from literature (USDA Database for the

Flavonoid Content of Selected Foods, Release 3.1 (December 2013), <http://www.ars.usda.gov> (Bhagwat et al., 2013).

## I.1. Flavonoid structure and biosynthesis

Due to their common biosynthetic origin, most of the flavonoids share the same structural basic motif, which is constituted by 15 carbon atoms, organized in a **C6-C3-C6 skeleton** (Figure 1) forming a **phenylbenzopyrane structure** (Harborne and Williams, 2000; Marais et al., 2006).

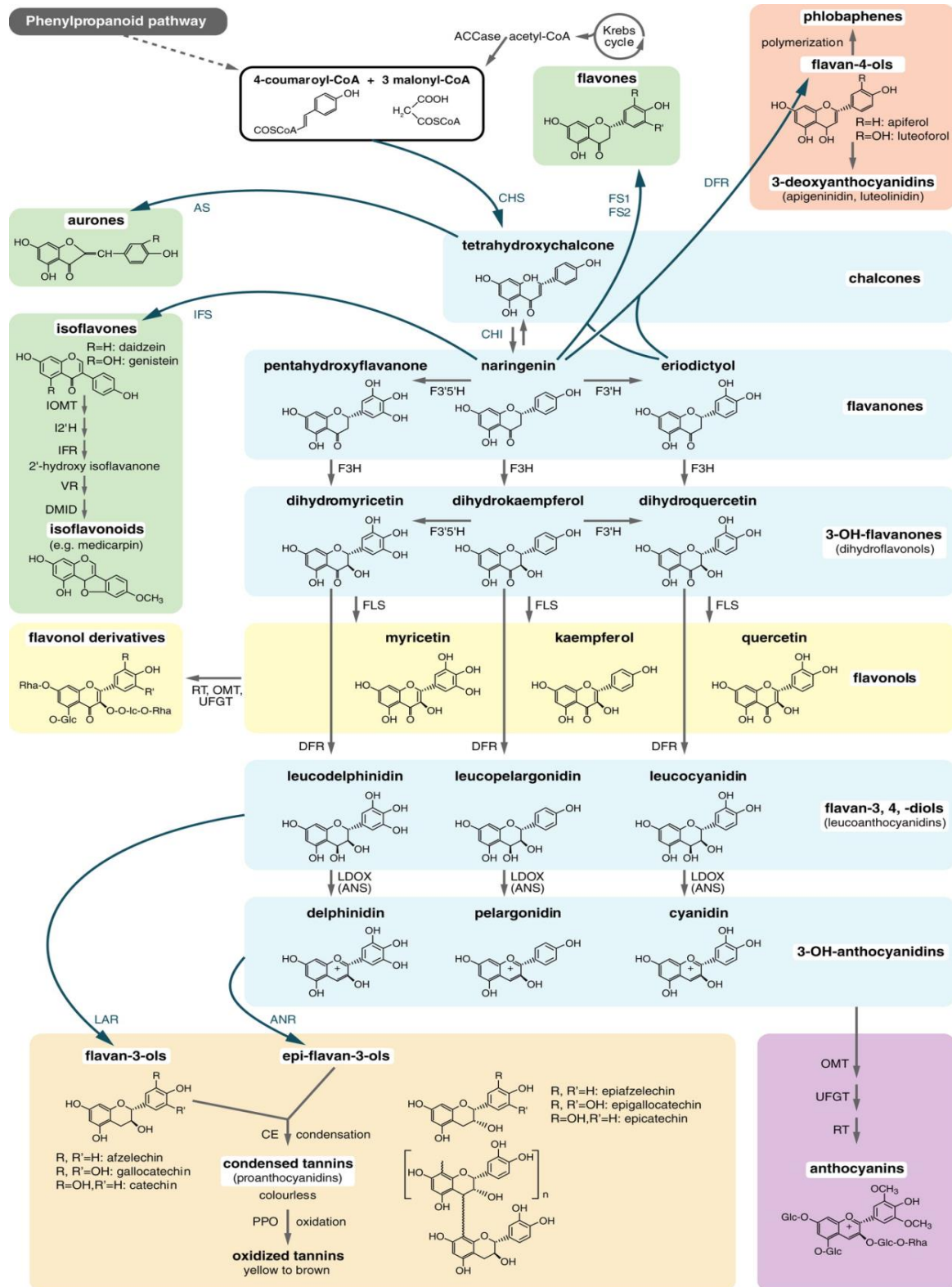


**Figure 1. Basic C6-C3-C6 skeleton of flavonoids.** The individual carbon atoms are based on a numbering system which uses classic atoms numbering with ordinary numerals for the A and C cycles and the “primed” numerals for B-cycle (Cook and Samman, 1996).

### I.1.1. Biosynthesis of flavonoids

Nowadays, the flavonoid biosynthetic pathway (Figure 2) is well elucidated (Saito et al., 2013; Winkel-Shirley, 2001, 2002), compared to other secondary biosynthetic pathways. **All flavonoids derived from the shikimate pathway** (Tripoli et al., 2007), the common route leading to the production of aromatic amino acids, such as phenylalanine. The latter is subsequently converted to 4-hydroxycinnamate-coenzyme A (4-coumaroyl-CoA) through the **phenylpropanoid pathway**. Then, chalcone synthase catalyzes the condensation of three molecules of malonyl-CoA (A ring) together with one 4-coumaroyl-CoA molecule (B and C rings) forming a 2', 4, 4', 6'-tétrahydroxychalcone molecule, which represents the common intermediate of all the flavonoids. From this stereospecific chalcone, the chalcone isomerase catalyzes its cyclisation to yield a (S)-4', 5, 7-trihydroxyflavanone molecule, leading to the common flavonoid C6-C3-C6 skeleton.

From this intermediate, many enzymes (synthase, reductase, hydroxylase) contribute to the biosynthesis of the different flavonoids classes. **In each class, molecules are further diversified with hydroxylation** catalyzed by flavonoid 3'-hydroxylases or flavonoid 3',5'-hydroxylases, methylation with O-methyltransferases, glycosylation with rhamnosyl transferases or other types of flavonoid glycosyl transferases, acylation with acyl-CoA transferase or polymerization with laccases.



**Figure 2. Flavonoid biosynthetic pathway from (Lepiniec et al., 2006).** ACCase, acetyl CoA carboxylase; ANS, anthocyanidin synthase; AS, aureusidin synthase; DFR, dihydroflavonol 4-reductase; DMID, 7,2'-dihydroxy, 4'-methoxyisoflavanol dehydratase; F3H, flavanone 3-hydroxylase; F3'H, flavonoid 3'-hydroxylase; F3'5'H, flavonoid 3'5' hydroxylase; FLS, flavonol synthase; FSI/FS2, flavone synthase; I2'H, isoflavone 2'-hydroxylase; IFR, isoflavone reductase; IFS, isoflavone synthase; IOMT, isoflavone O-methyltransferase; LAR, leucoanthocyanidin reductase; LDOX, leucoanthocyanidin dioxygenase; OMT, O-methyltransferase; RT, rhamnosyl transferase; UFGT, UDP flavonoid



The intra-cellular localization and organization of **enzymes involved in flavonoid biosynthesis** initially studied by Helen Stafford more than 35 years ago is still an ongoing question. In 1974, Helen Stafford proposed that they **operate as multi-enzymatic complexes** (Stafford, 1974). Many reports agreed with this first suggestion and pointed out that flavonoid biosynthesis pathway was likely to be an example of a **dynamic**, rather than a stable, enzyme complex (Burbulis and Winkel-Shirley, 1999; Dixon and Steele, 1999; Hrazdina, 1992; Hrazdina and Wagner, 1985). More recently, it has been established that, in higher plants, flavonoid synthesis starts when the enzymes involved in the synthesis **assemble at the cytosolic side of the endoplasmic reticulum** (Jørgensen et al., 2005). Flavonoids have also been found in the nucleus, the vacuole, the cell membranes and the cytoplasm, providing evidences of **intra-cellular flavonoid movements and transports** (Erlejman et al., 2004; Hutzler et al., 1998; Marinova et al., 2007; Naoumkina and Dixon; Saslowsky et al., 2005; Weston and Mathesius, 2013; Zhao and Dixon, 2010). Moreover, the increase in cell wall flavonoids contents, together with a decrease in soluble flavonoids, suggests a vacuolar efflux of these metabolites toward the cell wall, and underline the **extra-cellular transport** of flavonoids in plants (Agati et al., 2012).

### I.1.2. Flavonoid structures and classification

The various types of flavonoids differ in their level of oxidation and pattern of substitution of the C ring leading to 9 classes of flavonoids (Di Carlo et al., 1999; Harborne and Williams, 2000) (Figure 3):

- **Flavones** are characterized by the presence of a keto group at C4 and a C2,3-double bond. The A-ring of the majority of flavones is derived from phloroglucinol and the B-ring may be substituted at positions C3', C4' or C5'.
- **Flavonols** are flavones which possess a hydroxyl group at C3.
- **Flavanones** show a keto group at C4 without the C2,3- double bond, found in flavones. They possess a chiral center at C2.
- **Dihydroflavonols** are flavanones substituted at C3 with a hydroxyl group. They exhibit two asymmetric carbons, C2 and C3, thus giving two couples of enantiomers.
- **Isoflavones** present a B-ring linked to the C-ring at C3 instead of the classical C2 position for other flavonoid subclasses.

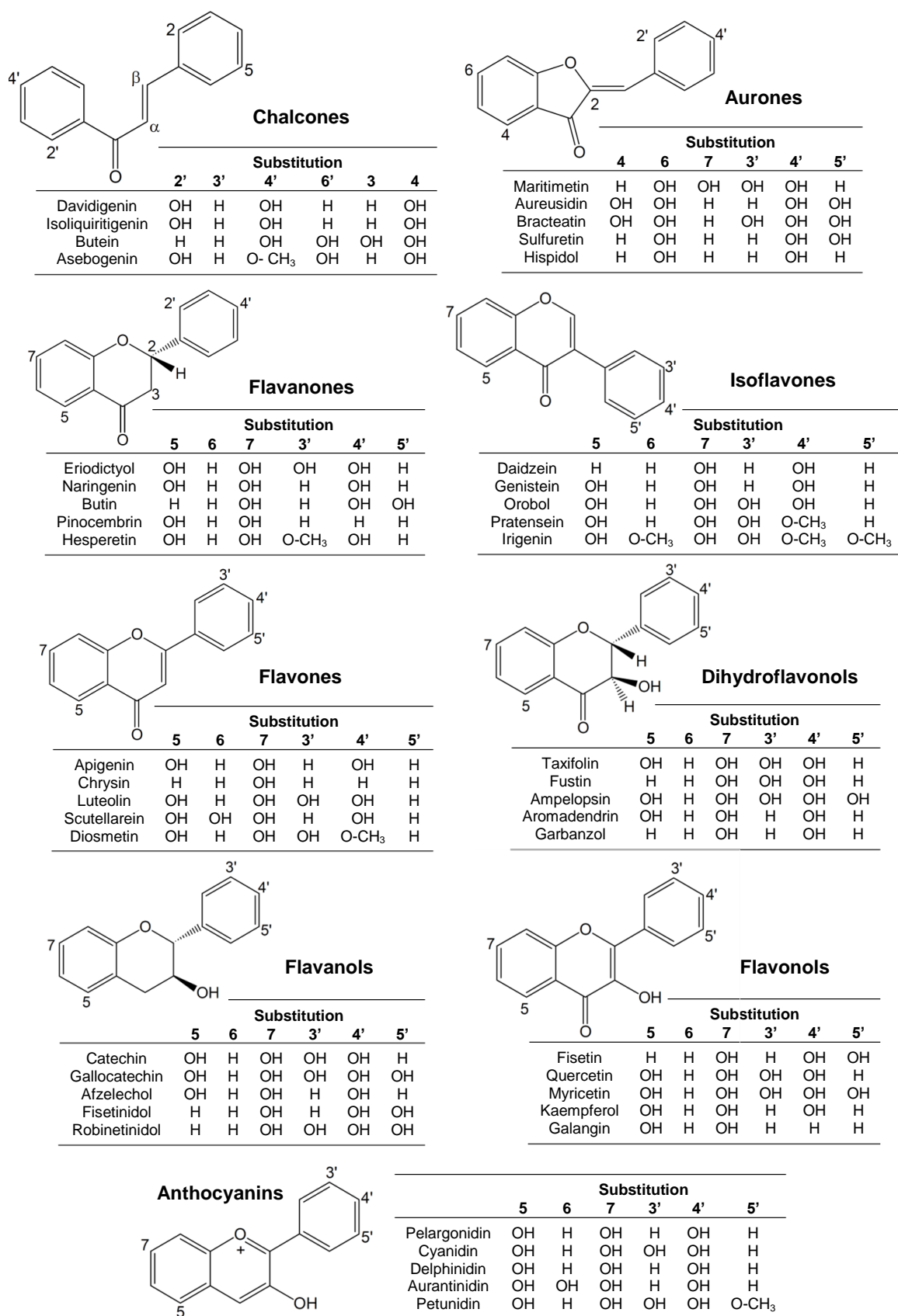


Figure 3. Classes of flavonoids: chemical structures and examples of compounds for each group.

- **Flavanols** are characterized by the absence of the 4-keto function and a C2,3-double bond. Two subclasses exist, depending on C4 hydroxylation or not, and named flavan-3,4-diol and flavan-3-ol, respectively.
- **Anthocyanidins** possess a skeleton based on pyrylium cation. They are characterized by the presence of a cationic charge and are typically not found as free aglycons.
- **Chalcones** and **aurones** are considered to be minor flavonoids. Chalcones are the open-chain flavonoids in which the two aromatic rings are linked by a three carbon  $\alpha,\beta$ -unsaturated carbonyl system. Numbering in chalcones is reversed as compared to the other flavonoids *i.e.* the A-ring is numbered 1'-6' and the B-ring 1-6. In aurone, considered as a chalcone-like group, a 5-atoms ring is present instead of the 6-atoms ring more typical of the flavonoid central C ring.

Individual substances within a subclass differ in the pattern of substitution of the A and B rings onto which mainly hydroxylated positions appeared to be C5, C7, C3', C4' and C5' (Figure 3).

## I.2. Physicochemical and biological properties

A large number of epidemiologic studies showed that flavonoid-rich diets are associated with a low incidence of chronic diseases, such as cardiovascular diseases (CVDs), type II diabetes, neurodegenerative diseases, and possibly cancers (Del Rio et al., 2013). Thus, **interest in understanding the relation between flavonoid structure, reactivity and health benefits has rapidly increased over the past decades** (Cheynier et al., 2012). The antioxidant property of flavonoids is one of the most specific features attributed to these molecules that play a critical role in the protection against chronic diseases. Besides their antioxidant activity, flavonoids also possess various physicochemical properties related to biological activities (Lu et al., 2013) that will be further described in the following section.

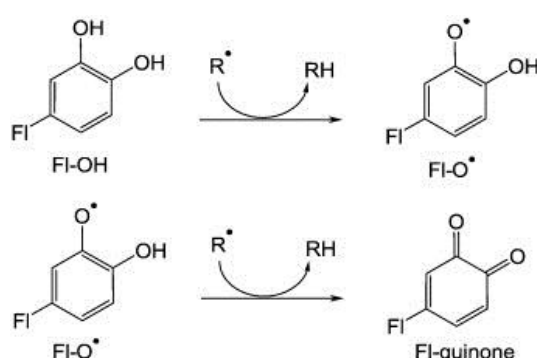
### I.2.1. Antioxidant properties

The term antioxidant is commonly used in scientific literature but can be defined in many different ways according to the methods used to evaluate the antioxidant activity. Halliwell and Gutteridge (1985) proposed a definition of an antioxidant as “**any substance that delays, prevents or removes oxidative damage to a target molecule**”. The oxidant protective effects of flavonoids in biological systems may be ascribed to

**different processes either direct** via the transfer of electrons from free radical species to flavonoids **or indirect** via the chelation of metal catalysts, the activation of antioxidant enzymes, the reduction of alpha-tocopherol radicals, the inhibition of oxidases, the mitigation of oxidative stress caused by nitric oxide or the increase of uric acid level (Heim et al., 2002; Procházková et al., 2011).

#### *1.2.1.a. Direct scavenging of Reactive Oxygen Species (ROS)*

Flavonoids can scavenge free radicals through a hydrogen atom donation and electron exchange. Radicals are made inactive according to the following global reaction (Figure 4), where  $R^\bullet$  is a free radical and  $FI-O^\bullet$  is a flavonoid phenoxyl radical (Pietta, 2000).



**Figure 4. Scavenging of reactive oxygen species ( $R^\bullet$ ) by flavonoid.** The free radical  $FI-O^\bullet$  may react with a second radical to yield a stable quinone structure (Pietta, 2000).

***In vitro*, flavonoid antioxidant activity depends on the arrangement of the functional groups on its core structure.** In particular, the presence of a catechol moiety (C3', C4' hydroxylated B ring) is the most significant determinant of ROS scavenging (Heim et al., 2002), whereas substitutions of the A and C rings with hydroxyl groups have little impact (Amic et al., 2007).

The main structural contributors to efficient radical scavenging are summarized below:

- An ortho-dihydroxy (catechol) structure in the B ring meaning conjugation of C3' and C4' hydroxyl groups for electron delocalization (Figure 5, blue box).
- C2,3-double bond in conjugation with a C4-oxo function in the C ring providing electron delocalization from the B ring (Figure 5, green box).
- Hydroxyl groups at positions C3 (heterocycle C) and C5 (ring A) providing hydrogen bonding to the oxo group (Figure 5, red box).



### *1.2.1.c. Activation of antioxidant enzyme gene expression*

Antioxidant action of flavonoids is also related to their capacity to interact with DNA and gene expression systems in both plants and mammals (Nijveldt et al., 2001). Indeed, **flavonoids may exhibit the capacity to induce the production of phase II important detoxifying enzymes** (NAD(P)H-quinone oxidoreductase, glutathione S-transferase, and UDP-glucuronosyl transferase), and regulate the expression of their encoding genes through an electrophile responsive element (EpRE) (Nerland, 2007). The most effective inducers were reported to be flavonoids containing a hydroxyl group at C3 position of the C ring (flavonols). In the absence of this hydroxyl group, some flavones may also act as low inducers, as demonstrated in mouse hepatoma cells (Lee-Hilz et al., 2006). Accordingly, administration of a flavonoid-rich fraction along with a high fat diet to rat caused a significant increase of superoxide dismutase, catalase and cytosolic glutathione peroxidase activities in erythrocytes (Kaviarasan and Pugalendi, 2009). Moreover, naringin administration to hypercholesterolemic humans led to similar results (Jung et al., 2003).

### *1.2.1.d. Inhibition of oxidases*

Antioxidant activity of flavonoids may also be indirectly linked to their **inhibitory effect onto oxidases responsible for superoxide ( $O_2^{\bullet-}$ ) production** such as xanthine oxidase (Chang et al., 1993), protein kinase (Gamet-Payraastre et al., 1999), cyclooxygenase (García-Mediavilla et al., 2007), lipoxygenase (Redrejo-Rodriguez et al., 2004), cytochrome P450 isoforms (Hodek et al., 2002), succinoxidase, NADH (Bohm et al., 1987) and NADPH oxidase (Steffen et al., 2008). QSAR studies indicate that the presence of free hydroxyl groups, C4-oxo group and the C2,3-double bond is required to observe inhibitory effect onto various protein kinases (Ravishankar et al., 2013).

**Flavonoids also prevent the toxic effect of nitric oxide**, whose toxicity is mainly mediated by peroxynitrite ( $ONO_2^{\bullet}$ ) formed in the reaction between free radicals nitric oxide ( $NO^{\bullet}$ ) and superoxide ( $O_2^{\bullet-}$ ) (Rubbo et al., 1994). The mechanism of action of flavonoids to reduce peroxynitrite production is still ambiguous. It may derived from a direct superoxide scavenging or due to a reduced expression of nitric oxide synthases involved in nitric oxide formation, or both (Kim et al., 2005). Several structural features were proposed to be beneficial for the direct peroxynitrite scavenging activity including the presence of C2,3-double bond with C4-oxo group, C3, C5, C7 hydroxyl and vicinal C3', C4' groups of the B-ring (Heijnen et al., 2001; Matsuda et al., 2003). Concerning nitric oxide synthases inhibition, flavonoids were reported to inhibit the induction of their expression and not directly their activity. The pathways of nitric oxide synthases induction converge in the activation of the essential nuclear transcription factor  $\kappa B$  (NF- $\kappa B$ ) which triggers transcription of nitric oxide

synthase genes. Interestingly, some flavonoids were reported to inhibit NF- $\kappa$ B activation through inhibition of I- $\kappa$ B kinases, and thus nitric oxide synthase induction. Flavonoids structural features such as C2,3 double bond, the C7 and C4' hydroxyl groups, and planar ring system in the flavonoid molecule were highlights as favorable structural elements of flavonoids for this induction inhibitory activity (Olszanecki et al., 2002).

Finally, several studies indicate that the **consumption of flavonoid-rich foods may increase plasma uric acid levels**, which is a major contributor to plasma total antioxidant capacity. However, the underlying mechanism still remains unclear (Halliwell, 2007).

### 1.2.2. UV radiations absorption and photo-protection activity

According to the International Commission on Illumination, UV radiation is divided in three categories: short-wave UVC (200–280 nm), mid-wave UVB (280–320 nm) and long-wave UVA (320–400 nm). Early physiological experiments provided evidences that flavonoids were involved in plant UV resistance and act as UV filters, thereby protecting the underlying photosynthetic tissues from damages (Harborne and Williams, 2000).

Most of the **flavonoid UV/Vis spectra exhibits at least two absorption bands**. The first one occur between 240 and 280 nm (Band II) and the second one between 300 to 400 nm (Band I) (Harborne and Williams, 2000). UV-absorption of polyphenols is **generally attributed to electronic transitions between  $\pi$ -type molecular orbitals over the molecular backbone**, depending on the flavonoid subclass (Table 1).

**Table 1. UV-visible absorption of flavonoid subclasses, in methanol** (Harborne, 1989; Markham, 1982).

Flavonoid subclass	UV-visible $\lambda_{\text{max}}$ absorption (nm)	
	Band II	Band I
Chalcones	230-270	340-390
Aurones	230-270	380-430
Flavanones	275-295	300-330
Isoflavones	245-275	310-330
Flavones	250-280	310-350
Dihydroflavonol / Flavonols	250-280	330-385
Flavanols	270-280	-

Anouar et al. (2012) showed that the absence of the C2,3-double bond in dihydroflavonols and flavanones substantially reduced  $\pi$ -conjugation as well as their UV absorption capacity, when compared to flavonols and flavones. It has also been demonstrated that the substituent pattern (number and position of OH groups) influences the absorption wavelengths of flavonoids. This is mainly due to the mesomeric (+M) effect of the **hydroxyl groups which allows an extension  $\pi$ -conjugation** (Harborne et al., 1975).

The maximum UV-visible ( $\lambda_{\text{max}}$ ) absorption of **flavonoids band I and II** correspond to **UV-A** ( $320 < \lambda_{\text{nm}} < 400$  nm) and **UV-C** ( $200 < \lambda_{\text{nm}} < 280$ ) radiations, respectively. Nevertheless, most of the flavonoids also exhibit UV-B absorption ( $280 < \lambda_{\text{nm}} < 320$ ) (Harborne and Williams, 2000). As UV-B stress enhances the production of ROS, UV-B-responsive **flavonoids can reduce the oxidative damage caused by short solar wavelengths**, by reducing the risk of ROS generation (Agati and Tattini, 2010).

### I.2.3. Inhibition of enzymes non-directly involved in oxidation

Flavonoids are known to inhibit a wide range of enzymes. A non-exhaustive list of enzymes non-directly involved in oxidation pathways is presented in Table 2. Some of the structural elements of flavonoids, identified as favorable or disadvantageous for enzyme inhibition, are also indicated.

As seen in Table 2, **flavonoids exhibit inhibitory effects, depending on their structures, on a huge range of enzymes**. Mechanisms of enzymatic inhibition by flavonoids were reported to be **mainly reversible**. **Competitive inhibition of enzymes**, induced by flavonoids, seems to be the most encountered along the literature. For example, Inhibition of fatty acid synthase was showed to involve competition between flavonoids and acetyl CoA at the acyl transferase catalytic domain (Tian, 2006). Partial antagonism between flavonoids and ATP at the nucleotide binding-site of cytosolic NBD2 domain of P-glycoprotein is also consistent with the competitive inhibition produced by quercetin on the ATPase activity of whole P-glycoprotein (Conseil et al., 1998). In addition, inhibition of kinases was reported to be due to the competitive binding of flavonoids with ATP at catalytic sites on the enzymes (Ferriola et al., 1989); The inhibitory mode of flavonol is usually competitive inhibition for the oxidation of L-dopa by tyrosinase (Chang, 2009); and baicalein was reported to be a competitive inhibitor targeting thiol groups of urease (Tan et al., 2013).



**Table 2. Examples of enzymes inhibited by flavonoids.**

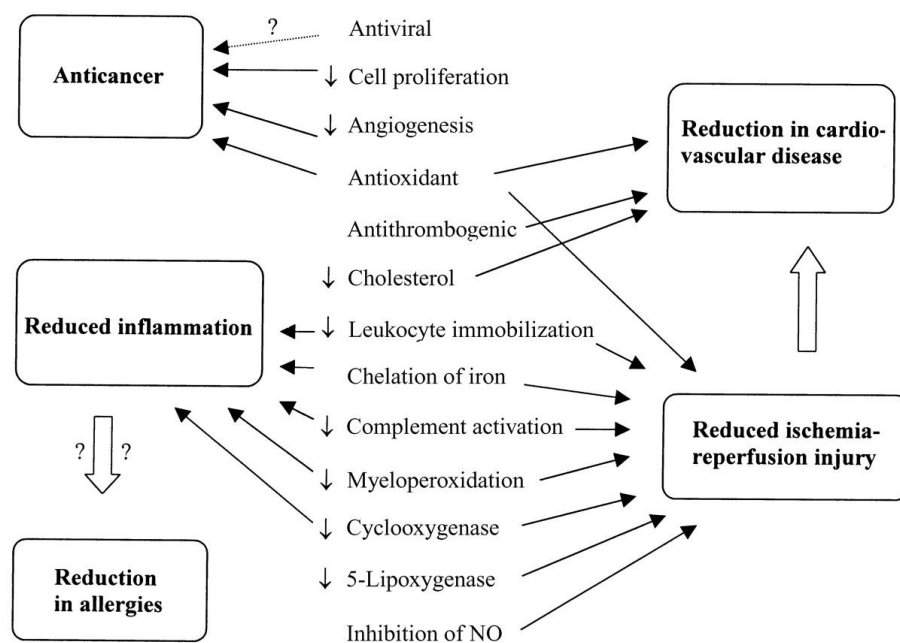
Enzyme	Structure-activity data <sup>a</sup>	Reference
α-glucosidase	(+) C2,3 double bond (+) C4-oxo function (+) A-ring : C5, C6, C7 hydroxylation (+) B-ring : C3', C4', C5' hydroxylation (-) C-ring C3 hydroxylation	(Li et al., 2009; Liu et al., 2013; Yao et al., 2013)
α-amylase	(+) C2,3 double bond (+) C4-oxo function (+) A and C-ring : C3, C6, C7 hydroxylation (+) B-ring : C3', C4' hydroxylation	Lo Piparo et al., 2008; Liu et al., 2013; Yao et al., 2013
Pancreatic lipase	(+) C2,3 double bond (+) C4-oxo function (+) A and C-ring : C3, C5, C6, C7 hydroxylation (+) B-ring : C3', C4' hydroxylation (-) C5' hydroxylation	(Lindahl and Tagesson, 1997; Liu et al., 2013)
DNA Topoisomerases	(+) C2,3 double bond (+) C-ring : C3 hydroxylation (+) B-ring : C3', C4' hydroxylation	(Boege et al., 1996; Snyder and Gillies, 2002; Webb and Ebeler, 2004)
Cyclic AMP Phosphodiesterase	(+) C2,3 double bond (+) C4-oxo function (+) A and C-ring : C3, C7 hydroxylation (+) B-ring : C3', C4' hydroxylation	(Ferrell et al., 1979; Madeswaran et al., 2012)
Aldose reductase	(+) C2,3 double bond (+) A and C-ring : C7 hydroxylation (+) B-ring : C3', C4' hydroxylation (-) C3 hydroxylation	Okuda et al., 1984; Matsuda et al., 2002; Mercader et al., 2008
Angiotensin converting enzyme		Loizzo et al., 2007; Guerrero et al., 2012; Parichatikanond et al., 2012
Reverse transcriptase	(+) C2,3 double bond (+) C4-oxo function (+) A and C-ring : C3, C5, C7 hydroxylation (+) B-ring : C3', C4', C5' hydroxylation	(Ono et al., 1990; Phosrithong et al., 2012)
Fatty acid biosynthetic pathway enzymes	(+) C2,3 double bond (+) C4-oxo function (+) A and C-ring : C3, C5, C7 hydroxylation (+) B-ring : C3', C4', C5' hydroxylation	Li and Tian, 2004; Brusselmans et al., 2005; Tasdemir et al., 2006; Brown et al., 2007
Cytosolic portion of Pglycoprotein	(+) C2,3 double bond (-) A, B and C-ring hydroxylation	Conseil et al., 1998; Kitagawa et al., 2005; Kitagawa, 2006
Proteasome	(+) C2,3 double bond (+) C4-oxo function (-) A, B and C-ring hydroxylation	Maliar et al., 2004; Chen et al., 2005; Jedinák et al., 2006
Cholinesterases	(+) A and C-ring : C3, C5, C7 hydroxylation (-) B-ring hydroxylation	Ji and Zhang, 2006; Plochmann et al., 2007; Katalinić et al., 2010; Uriarte-Pueyo and I. Calvo, 2011
Glucosyltransferases GTF B and C	(+) C2,3 double bond	Koo et al., 2002; Duarte et al., 2006

<sup>a</sup> Flavonoid structural elements necessary for enzyme inhibition: (+), the presence of this structural element promotes cited enzyme inhibition; (-), the presence of this structural element reduces cited enzyme inhibition

Flavonoids were also reported to act as **non-competitive, or mixed non-competitive inhibitors of some enzymes**. For example, flavonoids revealed a competitive–non-competitive mixed type of inhibition toward activity of some cytochromes (Kimura et al., 2010) and a mixed non-competitive type of inhibition of LOX (Ribeiro et al., 2014). Fisetin was reported to be a competitive inhibitor of Protein kinase C with respect to ATP binding but also a non-competitive inhibitor with respect to protein substrate (Ferriola et al., 1989). In many cases, the mode of inhibitory action of flavonoids turned out to be more complex than expected, as they can exhibit a **complicated combination of inhibition modes** (Sadik et al., 2003).

### I.3. Applications fields of flavonoids

Due to their reactivity, flavonoids exhibit a huge potential for human health (Figure 7), cosmetic and food industries. Many scientific publications and patent highlights their properties such as antiviral (Cody et al., 1986), antibacterial (Cushnie and Lamb, 2005), anticarcinogenic (Paliwal et al., 2005), anti-inflammatory (Lee et al., 1993), antiallergic (Tanaka, 2013), antidiabetic (Vessal et al., 2003), neuroprotective (Dajas et al., 2013), cardioprotective (Cook and Samman, 1996), etc.



**Figure 7. Hypothesis of the links between the working mechanisms of flavonoids and their effects on disease.** NO, nitrous oxide. (Nijveldt et al., 2001; Tapas et al., 2008)

Flavonoids naturally occur in food diets but they are also used in **food-processing industry** as dietary supplements, food color additives, and/or antioxidant (Fukumoto et al., 2007; Perlman, 2011). In the **cosmetic field**, flavonoids are mostly used for their antioxidant, UV filter, skin blood vessels protection and/or soothing actions (Arct and Pytkowska, 2008). But most of the data available in the scientific literature focus on flavonoid potential, for **therapeutics, to treat numerous chronic diseases**.

Table 3 summarizes some examples of possible applications in pharmaceutical industry for human or animal health. Relations between flavonoid structural features and biological activities are also indicated when the information were available.

**Table 3. Potential benefic effects of flavonoids in animal and human health**

Applications	Flavonoids	Structure-activity data <sup>a</sup>	References
Anti-ulcer	Catechins, flavanones, flavones, flavonol		(Beil et al., 1995; Raj et al., 2001; Tapas et al., 2008)
Rheumatoid arthritis	Apigenin, rutin		(Guardia et al., 2001; Kang et al., 2009)
Anti-inflammatory	Quercetin, apigenin, catechin, hesperidin, rutin, luteolin, kaempferol, myricetin, fisetin	(+) C3', C4', C5 and C7 hydroxylation (+) C2,3-double bound (+) C4-oxo function (-) C7 glycosylation or methylation (-) C3, C3'and C7 methoxylation	Miller, 1996; Le Marchand, 2002; Kang et al., 2009
Anticancer	Many flavones, flavonols, isoflavones, flavanones and catechins	(+) C3', C4', C5', C6' and C3 hydroxylation (+) C2,3-double bound (-) C4'and C6 methoxylation	Havsteen, 1983, 2002; Lea et al., 1993; Le Marchand, 2002; Ramos, 2007; Rang et al., 2007
Brain protection	Genistein, quercetin, fisetin		Maher et al., 2006; Jung et al., 2010
Antidepressant	Naringenin, 2S-Hesperidin, Linarin, fisetin		Fernández et al., 2006; Yi et al., 2010; Girish et al., 2012; Zhen et al., 2012
Cardiovascular protection	Many flavones, flavonols, isoflavones and flavanones		Knekt et al., 1996; Manach et al., 1996; Matsubara et al., 2004; Tapas et al., 2008; Mulvihill and Huff, 2010
Antidiabetes	Fisetin, quercetin		Russo et al., 2012; Prasath and Subramanian, 2013
Antiallergic	Quercetin,	(+) C3', C4' hydroxylation (+) C2,3-double bound	(Amellal et al., 2007; Samuelsson et al., 2009)
Antihepatotoxic	Quercetin, luteolin, morin, silybin, apigenin, ornitin,		Vijayaraghavan et al., 1991; Oh et al., 2004; Jayaraj et al., 2007; Weng et al., 2011

	gossypin		
Antiviral	Many flavones, flavonols and aurones, daidzein, genistein,		Formica and Regelson, 1995; Li et al., 2000; Mantas et al., 2000; Liu et al., 2008
Antimicrobial infections	Many flavones, flavanones and chalcones	(+) C5, C2' hydroxylation	(Alcaráz et al., 2000; Cushnie and Lamb, 2005)
Analgesic	Hesperidin		(Emim et al., 1994)
Antidiarrheal	Flavones, kaempferol, morin, myrcetin, naringenin, quercetin		(Di Carlo et al., 1993)
Antispasmodic	Catechin, luteolin, apigenin, chrysin, flavone, kaempferol, quercetin		Capasso et al., 1991 a, 1991 b; Ghayur et al., 2007; Engelbertz et al., 2012

<sup>a</sup> Flavonoid structural elements necessary for biological activity, reported from (Ioannou and Ghoul, 2012)): (+), the presence of structural elements promotes the cited activity; (-), the presence of structural elements reduces the cited activity.

## I.4. Flavonoid solubility

The solubility of a substance in a solvent is influenced by environmental parameters such as temperature and pH and also tightly linked to its structural characteristics, which control solute–solvent interaction forces. In this way, flavonoid solubility is strongly affected by the nature of both the solvent and the flavonoid structure according to the following classical interactions:

- **Hydrophobic interactions** between non-polar solvent and flavonoid aromatic rings A and C and carbonated aliphatic substituents, if present.
- **Dipolar interactions** between polar solvents and functional groups of flavonoid (carbonyl, ether, ester, hydroxyl)
- **Hydrogen bonds** between solvent (water, amine, alcohol) and acceptor or donor substituent of the flavonoid, such as sugars.
- **Electrostatic interactions**, between hydroxylic and carboxylic groups, that are strongly influenced by pH.

Numerous studies have established that **flavonoids present a global low aqua- and lipo-solubility** (Rice-Evans, 2001). For example, it has been reported that solubility of naringenin and quercetin in water, at 20°C, were very low, with only 500 and 10 mg.L<sup>-1</sup> respectively (Calias et al., 1996; Pulley, 1936). According to Saidman et al. (2002), the major factor influencing flavonoid solubility in water is their capacity to form hydrogen bond with solvent

molecules. Most of the flavonoids exhibit a **higher solubility in organic solvent or when organic solvent are added to water** (Bertrand et al., 2006; Chebil et al., 2007; Peng et al., 2006). For example, naringenin and quercetin solubility reaches 21 g.L<sup>-1</sup> in acetonitrile and 25 g.L<sup>-1</sup> in acetone at room temperature, respectively (Chebil et al., 2007). These values are 40 and 250-fold higher than those observed in water.

However, **it remains difficult to predict the solubility of a flavonoid in a given solvent**. Indeed, the number, localization and nature of the many possible substituents on the flavonoid skeleton, such as hydroxyl, methoxy, osidic, or acyl groups, are critical for its solubility (Plaza et al., 2014). Nevertheless, it is well established that **glycosylation of flavonoids globally improve their solubility in aqueous solutions**, as it will be latter exemplified in the next parts of this literature study.

## I.5. Bioavailability of flavonoids

The high reactivity of flavonoids renders them **sensitive to several environmental parameters**, such as light exposure, pH, temperature, solvent nature, presence or absence of metallic ions and oxidants (Jackman et al., 1987). **Flavonoid stability is highly related to their chemical nature (subclass and substituents)** (Plaza et al., 2014). Indeed, the stability of flavones to light exposure is higher than that of flavonol, due to the absence of the C3 hydroxyl group in the flavone structure (Smith et al., 2000). Accordingly, **their degradation processes may vary importantly**, depending on their structure. This was shown for the degradation of daidzein and genistein (isoflavones), which exhibit different degradation profiles at high temperature (Ungar et al., 2003). In addition, we have seen that many flavonoids are poorly soluble in aqueous medium. All these parameters influence their bioavailability.

Numerous studies attempted to identify the cellular targets that could be involved in the health promoting actions of flavonoids. However, the molecular interactions between polyphenols and these cellular targets remain mostly speculative, as most of the studies on flavonoids activities were performed *in vitro* (Clere et al., 2011; Faria et al., 2013; Hertog et al., 1993). Moreover, a very important problem that often compromises *in vitro* studies is the dose applied. The dose used should reflect real life but the **tested concentrations commonly range from  $\mu\text{mol/L}$  to  $\text{mmol/L}$** , while bioavailability (proportion of a substance which reaches systemic circulation unchanged, after a particular route of administration (glossary of terms: <http://cot.food.gov.uk/moreinfo/cotglossary>), data showed that the **concentration in plasma**

and organs following oral ingestion rarely exceed the nM range (Williamson and Manach, 2005).

The two main routes of flavonoids administration in the organism are oral and percutaneous. The percutaneous route mostly concern cosmetic applications of flavonoids. Bonina et al. (1996) have shown that naringenin, hesperetin and, at a much lower degree quercetin, were able to permeate through the *stratum corneum* (main barrier against exogenous substances penetration through skin) and penetrate into deeper skin layers. But, **intradermal delivery, as for most of the flavonoids, remains very low, due to the skin impermeability as well as the low solubility of flavonoids** in both aqueous and lipophilic environments (Kitagawa et al., 2009). This inefficient incorporation of flavonoids is often **overcome in cosmetic preparation by addition of enhancers or co-surfactants**, such as deoxycholic acid, dicetyl phosphate, diethylene glycol monoethylether and propylene glycol (Fang et al., 2006; Marti-Mestres et al., 2007).

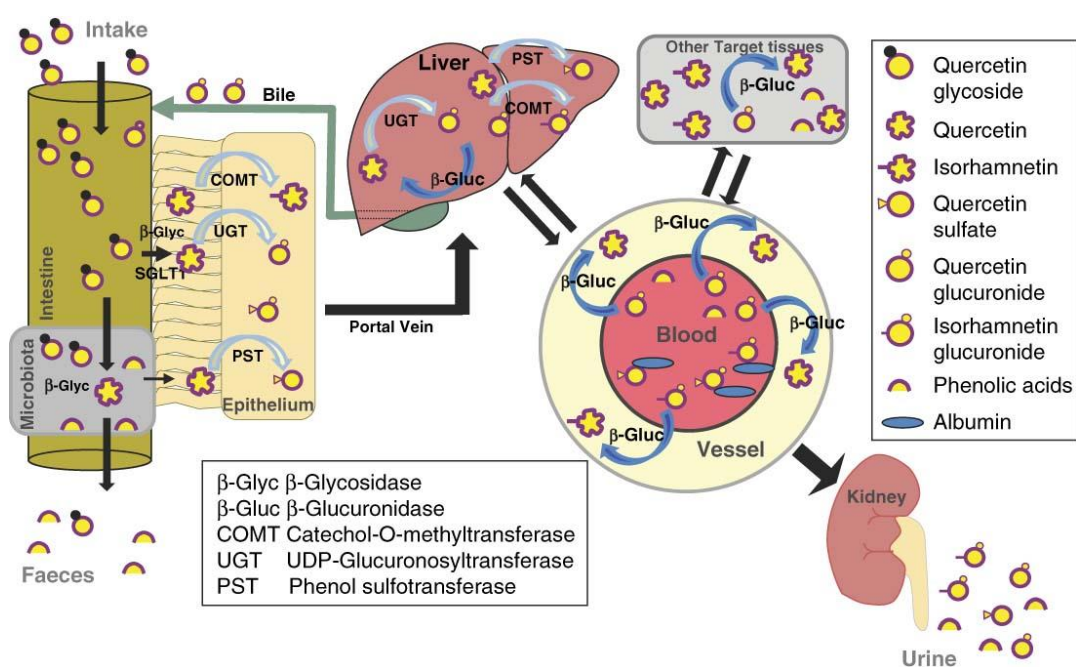
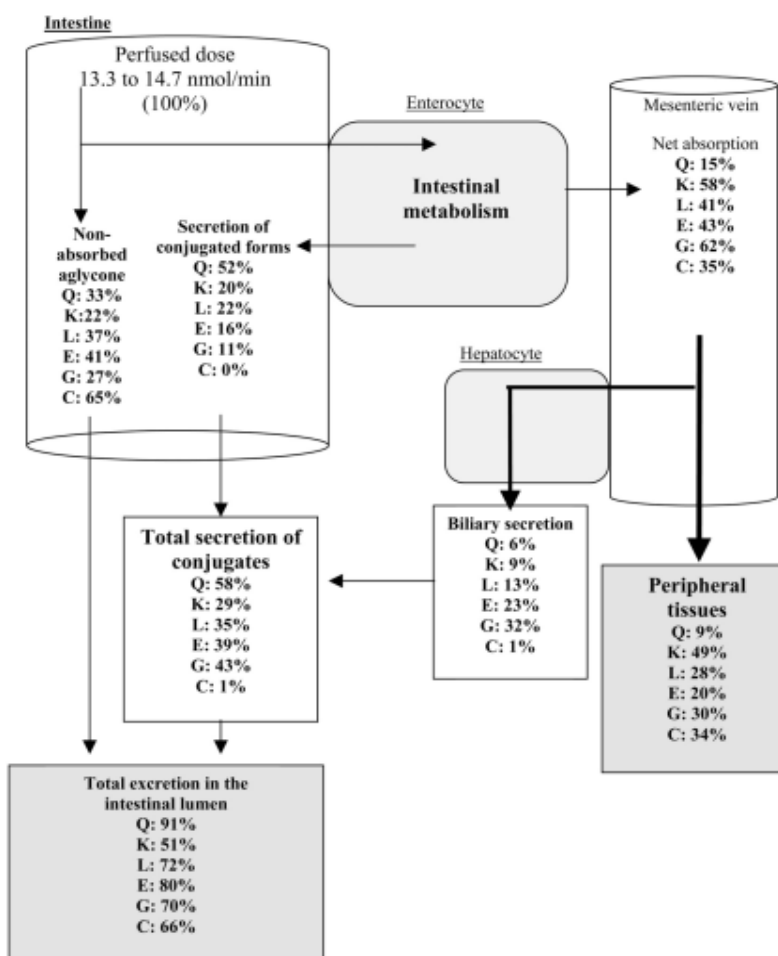


Figure 8. Schematic representation of the pharmacokinetics of quercetin (Perez-Vizcaino et al., 2012).

Regarding the oral route, transport across **the gastro-intestinal epithelium constitutes a barrier for flavonoid diffusion into peripheral tissues**. It is generally assumed that aglycon forms of flavonoids are absorbed only after **prior hydrolysis of the glycosides** along the aerodigestive tract, mainly due to digestive tractus microflora (Németh et al., 2003; Walle, 2004),

as shown for quercetin on Figure 8. Oral quercetin glycosides are deglycosylated before partial absorption in the intestine and **metabolization into glucurono- and sulfo-conjugated metabolites**, which are the circulating forms in plasma (Perez-Vizcaino et al., 2012).

A difficulty encountered when looking at the bioavailability of ingested flavonoids, is the **high variation of their behavior in the splanchnic metabolism**. Indeed, when flavonoids absorption was investigated in rats after an *in situ* perfusion of *jejunum plus ileum*, results showed that their repartitions in organism depend on the nature of the compound considered, which highly influences its availability for peripheral tissues (Crespy et al., 2003) (Figure 9).



**Figure 9. Distribution of flavonoids in the splanchnic area** (Crespy et al., 2003). For each compound, the data presented are expressed as the percentage of the initial perfused flux. Q, quercetin; K, kaempferol; L, luteolin; E, eriodictyol; G, genistein; C, catechin.

Most of the studies on the bioavailability of orally administrated flavonoids have been performed through the intake of functional food. But, it has been suggested that clinical trials, as used in pharmaceutical research, may not be appropriate for functional foods. Indeed, the



participants generally have varied diets before, during and after the trial. These differences may significantly affect the outcomes (Scheepens et al., 2010). Concerning humans, only few studies on the bioavailability of orally administrated purified flavonoids have been conducted. One of them concern the disposition of [ $^{14}\text{C}$ ]quercetin after oral dose administration to humans. It established that the oral absorption was comprised between 36 and 54% of the diet among the six subjects studied. However, the bioavailability was close to zero (Walle et al., 2001a). In a second study, the behavior of chrysin was reported to be very similar to that of quercetin. After recommended oral doses of 400 mg of chrysin, there were only trace amounts in plasma, corresponding to an estimated bioavailability of 0.003 – 0.02% (Walle et al., 2001b).

Thus, different **formulations have been proposed to enhance flavonoids bioavailability after oral intake**. The strategies for enhancing polyphenol bioavailability reported in the literature are freeze drying or emulsion (Fang and Bhandari, 2010), phospholipid–polyphenol complexes, formation of inclusion complexes with cyclodextrins or dendrimers, use of structural analogues, formation of derivatives (e.g., esterification or glycosylation), transdermal delivery systems, polymeric implants, nanodisks and nanofibers, crystal engineering (e.g., cocrystallization) and use of adjuvants as absorption enhancers (Lewandowska et al., 2013). **Concerning adjuvants, combinations of flavonoids with sugar fractions** such as  $\alpha$ -1,6-glucooligosaccharide (Shinoki et al., 2013), fructooligosaccharides (Matsukawa et al., 2009; Phuamongkolwivat et al.), chito-oligosaccharide (Zhou et al., 2013), inulin (Piazza et al., 2007; Zduńczyk et al., 2006) were developed and shown to improve their bioavailability.

Despite that, **generating flavonoid structures with intrinsic enhanced bioavailability stay an attractive research field** as it can allow readily using only one active compound and avoiding having need of much more complicated formulations.

## I.6. Modifications of flavonoids to improve their properties

As seen before, the industrial use of flavonoids is often limited by their low stability, solubility and bioavailability. Thus, there is a strong demand for improving their physicochemical properties by introducing selective modifications (Chebil et al., 2006). **In that way, chemical, enzymatic, or chemo-enzymatic modifications of a number of flavonoids have been performed**. Biocatalytic modifications of flavonoids are generally seen as a good alternative to conventional chemical methods because of the mild reaction conditions used, the high catalytic



efficiency, and the regio- and stereo-selectivity of the biocatalysts that result in purer products (Bresson et al., 1999; Gao et al., 2000; Katsoura et al., 2007).

**Glycosylation and acylation of flavonoids are the two biocatalytic reactions that have received particular attention to enhance their hydrophilic or hydrophobic character, respectively, as well as their stability** (Plaza et al., 2014). Concerning flavonoids glycosylation, a stability study of several flavonoids in honey samples during the ultrasonic extraction and microwave-assisted extraction showed that glycosides (naringin, rutin, hesperidin and quercetrin) exhibited highest stability than the tested aglycons (Biesaga and Pyrzyńska, 2013). Moreover, (Raab et al., 2010) reported that, in general, the (+)-catechin glucosides were at least twice more stable than (+)-catechin, between pH 4 and 8. It is also well established that glycosylated flavonoids often present higher water solubility than their aglycons, due to the hydrophilic nature of the linked sugar(s).

Many examples of flavonoid glycosides improved water solubility will be given in the following parts of this chapter which will focus on the enzymatic glycosylation of flavonoids.

Thus, these **enhanced solubility and stability of flavonoid glycosides may improve their bioavailability**, as some studies on hesperetin, hesperidin and quercetin metabolic absorption showed that their glycoside derivatives exhibited enhanced bioavailability in rats, compared to aglycons (Makino et al., 2009; Murota et al., 2010; Nielsen et al., 2006; Yamada et al., 2006). Nevertheless, as seen before in this part, flavonoids physicochemical properties, biological activities, stability, and solubility **may greatly vary according to the nature and localization of the substituent linked to their backbone** (Plaza et al., 2014). Moreover predicting the bioavailability and the tissues distribution of flavonoid intake appear to be an intricate issue, even if much attention is paid to the development of computational pharmacokinetics predicting models (Mostrag-Szlichtyng and Worth, 2010; Singh et al., 2009).

Thus, these considerations point out the **need to generate a large panel of different flavonoids with multiple patterns of substitutions, such as glycosylation**, to allow finding a new compound that exhibits all the desired properties for its use as health promoter.

## II. Glycosylation of flavonoids

In nature, the large number of different and often species-specific flavonoid structures is mainly due to the variations of their substitution patterns with hydroxyl, glycosyl, methyl or acyl groups. In this part, we will focus our study on flavonoid glycosylation. Flavonoids substituted by glycosyl moieties are named flavonoid glycosides. When no sugar is attached to their backbone, they are referred to as aglycon.

### II.1. Natural glycosylation of flavonoids in plants

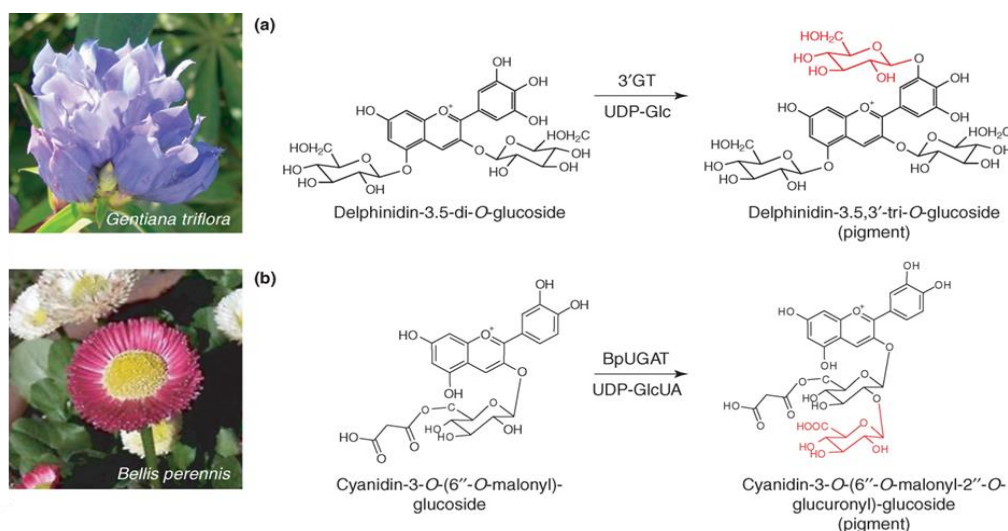
#### II.1.1. General Considerations

In a single plant species, dozens of different flavonoids may be present and most of them are linked to various sugar moieties (Forkmann and Heller, 1999). They are mainly found in the form of their O- and C-glycosides. Here, we will **focus our study on the O-glycosylation of flavonoids**. In plants, the position of the O-glycosylation is variable. However, some trends can be drawn depending on the flavonoid subclass. For example, glycosylation of flavones mainly occurs at C7 position while C3 position is preferred for flavonols. In both cases, glycosylation at C4', C3' and C5 carbons is a rare phenomenon (Hollman and Arts, 2000). **Glycosylation usually occurs in the last steps of the flavonoid metabolic pathway** (Harborne and Williams, 2000; Harborne et al., 1975; Jones and Vogt, 2001).

The most often encountered sugar moieties include D-glucose, D-galactose, L-rhamnose, D-glucuronic acid, D-galacturonic acid, L-arabinose, D-xylose and rutinose. In addition, D-allose, D-apiose, and D-mannose can be occasionally found (Bohm, 1999). D-glucosides usually display a  $\beta$ -anomeric configuration while L-sugars adopt a  $\alpha$ -anomeric configuration. **The most frequent natural form of flavonoid glycosides correspond to  $\beta$ -glycoside derivatives** (Harborne et al., 1975; Havsteen, 1983).

The presence and the position of one or several sugar moieties substituted on the flavonoid backbones account for their variety and also significantly impact their properties, in particular their solubility, stability, bioavailability and also their biological activities such as phytotoxicity (De Martino et al., 2012; Vogt and Jones, 2000). For example, flavonoids and anthocyanidins with a free hydroxyl group at the C3 position of the heterocyclic C-ring are

unstable under physiological conditions and, therefore, rarely found in nature (Forkmann and Heller, 1999).



**Figure 10.** Examples of plant pigment formation through anthocyanin glycosylation (adapted from Bowles et al., 2005). (a) In *Gentiana triflora*, formation of the blue pigment delphinidin-3,5,3'-tri-O- $\beta$ -glucoside. (b) Cyanidin-3-O-(6''-O-malonyl-2''-O-glucuronyl)- $\beta$ -glucoside is a major red pigment in *Bellis* flowers.

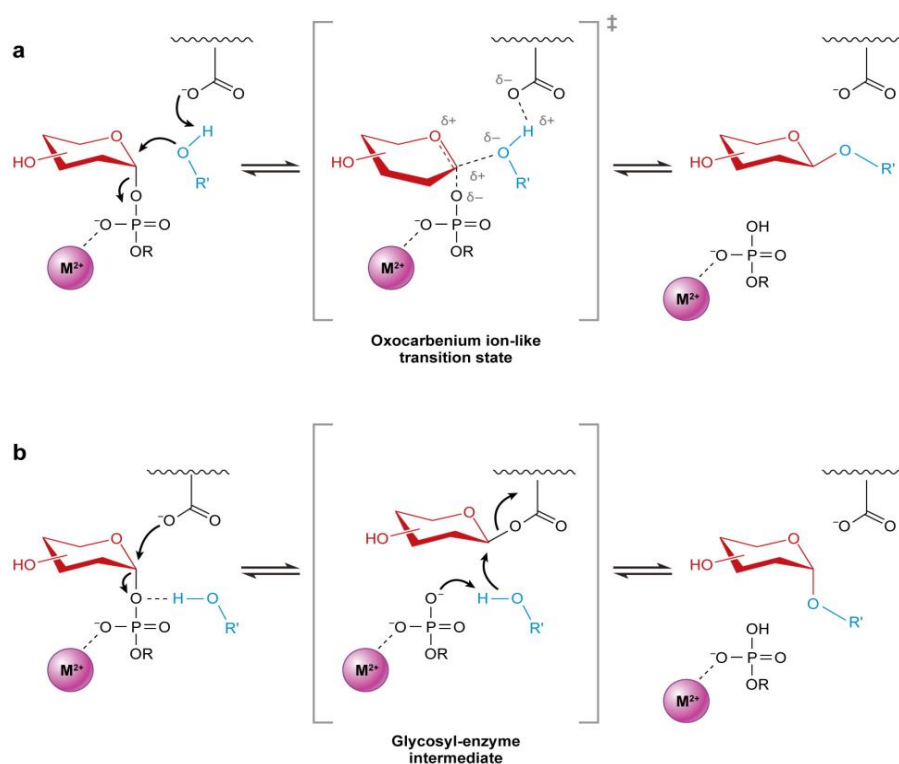
Sugar moieties linked to the flavonoid backbones of flavonoids were also reported to play a significant role in their physiological transport and accumulation in the plant cell vacuoles (Zhao and Dixon, 2010). A crucial role of glycosylation reactions in plants is the formation of glycosylated anthocyanins, these glycosides are water soluble pigments that contribute to color formation in many plants tissues (Tanaka et al., 2008). Figure 10 illustrates the role of anthocyanin glycosides in flower pigmentation. **In plants, glycosylation of flavonoids is catalyzed by a family of ubiquitous enzymes called glycosyltransferases (GTs).**

### II.1.2. Glycosyltransferases (GTs) overview

Glycosyltransferases (GTs) are found in all living organisms, including plants, animals, and bacteria (Keegstra and Raikhel, 2001; Ross et al., 2001). They are most accurately defined as **enzymes utilizing an activated sugar donor substituted by a leaving group at the anomeric carbon**. Sugar donors are most commonly activated in the form of nucleoside diphosphate sugars such as UDP-glucose, for example. **Nucleotide sugar-dependent glycosyltransferases are referred to as Leloir enzymes**, in honor of Luis F. Leloir who discovered the first nucleotide sugar and was awarded the Nobel Prize in chemistry, in 1970.

Leloir GTs catalyze glycosylation reactions involve the transfer **of a glycosyl group from an activated sugar moiety (NDP-donor) onto a broad variety of acceptor molecules** (Lairson et al., 2008). Two stereochemical outcomes are possible. The **anomeric configuration of the product can be either retained or inverted**, compared to that of the donor substrate (Figure 11).

**The mechanistic strategy employed by inverting glycosyltransferases** is that of a direct displacement  $S_N2$ -like reaction. A divalent cation interacts with the donor to facilitate binding and cleavage of the glycosidic linkage, by neutralizing the negative charges of the pyrophosphate group. A catalytic amino acid (usually an aspartic acid or glutamic acid) acts as a base catalyst to deprotonate the hydroxyl group of the acceptor and allows the nucleophilic attack at C1 carbon of the donor. The transition state is believed to possess substantial oxocarbenium ion character (Figure X11a).

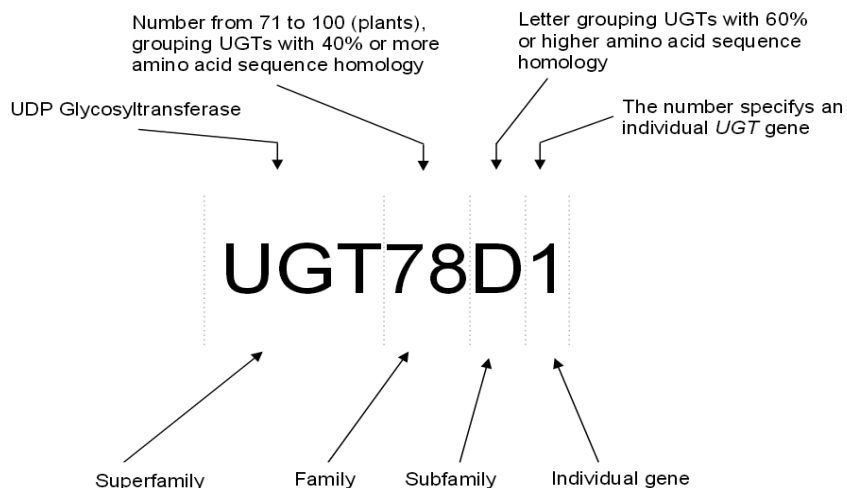


**Figure 11. GT reaction mechanisms that lead to the inversion or retention of the anomeric stereocenter upon glycosyl transfer** (Lairson et al., 2008) (a) *Inverting glycosyltransferases utilize a direct-displacement  $S_N2$ -like reaction that results in an inverted anomeric configuration via a single oxocarbenium ion-like transition state.* (b) *A proposed double-displacement mechanism for retaining glycosyltransferases involves the formation of a covalently bound glycosyl-enzyme intermediate.* Abbreviations:  $R$ , a nucleoside;  $R'OH$ , an acceptor group.

**The catalytic process of retaining glycosyltransferases** results in retention of the anomeric carbon configuration. Mechanistic evidences remain unclear, but the suspected mechanism, by comparison with the classical Koshland retaining mechanism for Glycoside Hydrolases, involves a two step reaction (Koshland, 1953). The enzymatic nucleophile attacks the glycosyl donor at the opposite side of the anomeric carbon to generate a glycosyl enzyme intermediate with inversion of stereochemistry. The subsequent reaction with the glycosyl acceptor provokes a second inversion of stereochemistry, to yield a product glycoside with a net retention of anomeric stereochemistry. However, no structure of the glycosyl-enzymes could be obtained (Figure 11b). The most common alternative mechanism proposed involves an S<sub>N</sub>i-like process (Lee et al., 2011). In this process, the pyrophosphate group on the donor departs, and the nucleophile attacks from the same face, while the other face of the donor is blocked by the enzyme.

Since 1955, all experimentally **characterized** enzymes, including GTs have been **categorized according to the rules from International Union of Biochemistry and Molecular Biology (IUBMB)** that attribute an Enzyme Commission number (EC number), based on the chemical reactions they catalyze (<http://www.chem.qmul.ac.uk/iubmb/enzyme/>). In this universal classification, glycosyltransferases form the **EC 2.4** group.

In 1997, the **UGT Nomenclature Committee proposed a nomenclature to code all glycosyltransferases** (Mackenzie et al., 1997), characterized or not. They suggested a name for the encoding gene, according to the universal guidelines established by the Human Gene Nomenclature (<http://www.gene.acl.ac.uk/nomenclature/guidelines.html>). For naming an enzyme or a gene, they recommended to **use the symbol XGT** denoting 'XDP-sugar Leloir glycosyltransferase' where X can be A, C, T or U according to the glycosyl-attached nucleotide phosphate. When the donor specificity remains unknown, X can also be the initials of the organism from which the enzyme have been isolated. It is followed by an Arabic number representing the family (1-200; 71-100 for plant enzymes), a letter designating the subfamily (A-Z), and an Arabic numeral denoting the gene within the family or subfamily (Figure 12).



**Figure 12. Nomenclature system of the UGT superfamily** (adapted from Ross et al., 2001).

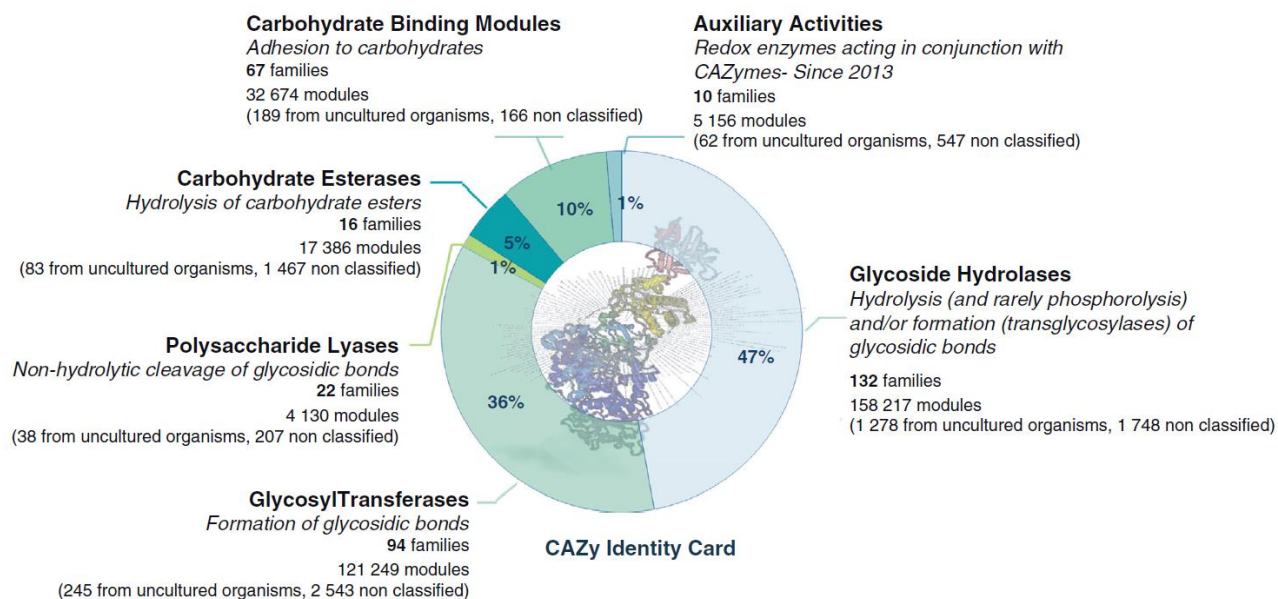
This nomenclature is also **hoped to help identify potential substrates for putative XGTs**. However, it can be difficult to predict their function and substrate specificity by solely based the analysis on primary sequence homology with known and characterized XGTs. For example, two UGTs identified in *Dorotheanthus bellidiformis*, exhibiting below 20% sequence similarity, were shown to use the same substrates and the only difference was the specific transfer of the glucosyl moiety (from UDP-glucose) onto one or the other vicinal hydroxyl groups of the flavonoid, with the same reaction mechanism (Vogt, 2002).

In 1998, the **CAZy (Carbohydrate-Active Enzyme) database** dedicated to the display and analysis of genomic, structural and biochemical **information on Carbohydrate-Active Enzymes (CAZymes)** was developed to improve their classification (<http://www.cazy.org/>) (Cantarel et al., 2009; Lombard et al., 2013). The classification method is based on significant amino acid sequence similarity with at least one biochemically characterized founding member, and also **takes into account the characteristic features identified among the characterized CAZymes members**. Thus, proteins with no biochemical evidence, such as the thousands of uncharacterized sequences of carbohydrate-active enzymes, also enter the classification. This sequence-based method is **rather different but complementary to the Enzyme Commission classification** scheme.

In the CAZy classification, families are designated using a number that reflects the order of family creation within the same group (Glycosyltransferase Family 1: GT1). Subfamilies are subgroups found within a family that share a more recent ancestor and, that are usually more uniform in molecular function. Currently, subfamilies are only described for family GH13 and

Polysaccharide Lyase Families. **Glycosyltransferases have been classified on the basis of their amino acid sequence, structural similarities, and catalytic mechanism** (inversion or retention of the anomeric configuration of the sugar moiety). At present, **94 distinct families** of glycosyltransferases are referenced. GTs represent the second group of enzymes the most represented in the CAZy classification with more than **36% of the enzymes** of the database and more than 121 000 modules reported (Figure 13).

The **GTs catalyzing flavonoid glycosylation** are found in both prokaryotic and eukaryotic organisms but the **vast majority** of them have been **identified in plant species**. Indeed, they are expressed in all plant tissues and are often found as soluble enzymes in cells cytosol. They can form a temporary structural-functional complex with other enzymes of the metabolic pathway, through non-covalent interactions, named metabolon (Jørgensen et al., 2005). **Most of the GTs naturally involved in flavonoids glucosylation belongs to glycosyltransferase family 1 (GT1)**, which gather 6157 putative inverting glycosyltransferases. As the enzymes of this family usually transfer sugar units from uridine-diphosphate activated monosaccharides (UDP-sugar), they are **often referred to as UGTs** (Bowles et al., 2006).



**Figure 13. Diversity of CAZymes and auxiliary enzymes, according to the CAZy database, 2013 release (André et al., 2014).**

The GTs family 1 is defined by the presence of a consensus sequence named PSPG-box (Plant Secondary Product Glycosyltransferase motif) located near the C-terminal end of the protein (Coutinho et al., 2003; Offen et al., 2006). This box consists of a short 44 amino acid

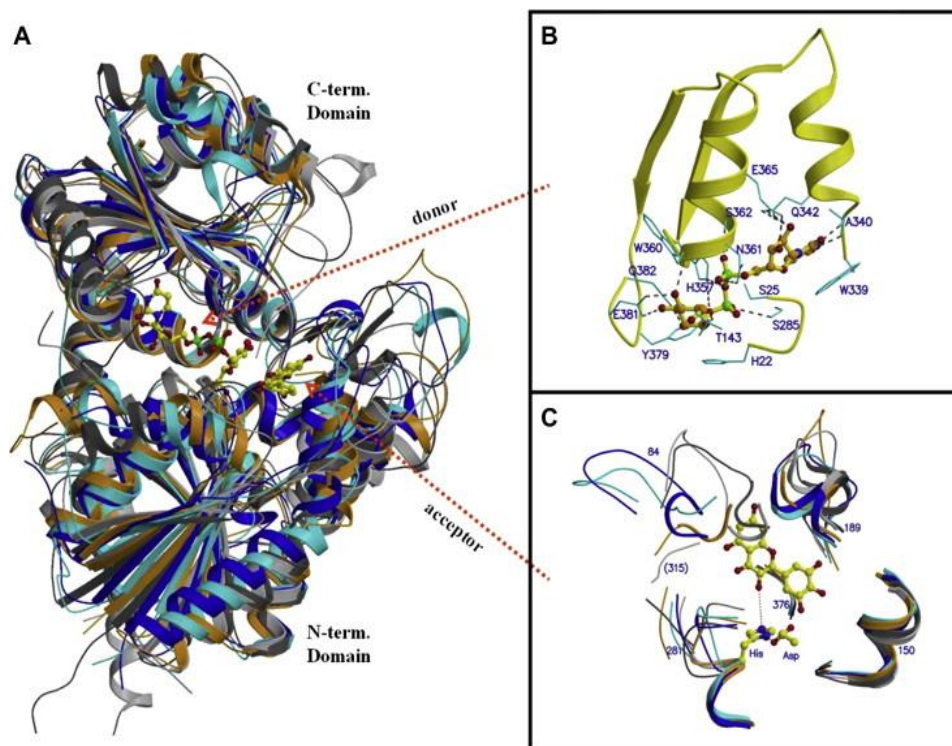


fragment that is involved in the binding of the UDP moiety of the sugar nucleotide (Lorenc-Kukuła et al., 2004). Using this PSPG box consensus sequence as a search tool, 117 genes encoding for putative or characterized UGTs of family 1 were identified in the *Arabidopsis thaliana* genome, 187 in *Medicago truncatula* and 202 in *Oryza sativa* (Gachon et al., 2005; Kim et al., 2006a; Yonekura-Sakakibara et al., 2007).

At present, in the Protein Database Bank (PDB, <http://www.rcsb.org/pdb/>) **six 3D-structures of plant UGTs from GT family 1 that naturally glycosylate flavonoids have been reported** : 3 UGTs from *Medicago truncatula*, *i.e.* UGT71G1 (saponin biosynthesis, recognizing (iso)flavonoids) (Achnine et al., 2005; Shao et al., 2005), UGT85H2 ((iso)flavonoid glycosyltransferase) (Li et al., 2007) and UGT78G1 (cyanidin glycosyltransferase) (Modolo et al., 2009a); 1 UGT from *Vitis vinifera*, VvGT1 (cyanidin 3-O-glycoside formation in red grape) (Offen et al., 2006); 1 UGT from *A. thaliana*, UGT72B1 (bifunctional N- and O-glucosyltransferase of xenobiotic metabolism) (Brazier-Hicks et al., 2007); 1 UGT from *Clitoria ternatea*, Ct3GT\_A (UDP-glucose:anthocyanidin 3-O-glucosyltransferase) (Hiromoto et al., 2013).

Although they share relatively low sequence identities, these plant UGTs show **highly similar structures, as the GT-B fold**, which is one of the two general folds found in the GT superfamily (Bourne and Henrissat, 2001; Coutinho et al., 2003). It consists of two N- and C-terminal domains with similar Rossmann-like folds (Figure 14 A). Each domain contains a central  $\beta$  sheet flanked by  $\alpha$  helices on both sides.





**Figure 14. Structures of plant UDP glycosyltransferases**, adapted from Wang (2009). (A) Comparison of structures of plant UGT78G1 (cyan, PDBID 3HBF), VvGT1 (blue, 2C1Z), UGT85H2 (orange, 2PQ6), UGT71G1 (dim grey, 2ACW) and UGT72B1 (light grey, 2VCE). The UDP-2-fluoroglucose (upper left) and acceptor kaempferol (lower right) in grape VvGT1 structure are shown as ball-and-stick models. (B) Donor binding site and interaction between the donor molecule UDP-glucose and the enzyme UGT71G1. The PSPG motif is shown as a ribbon model in yellow. The structure of UDP-glucose is shown as a ball-and-stick model. Some protein residues interacting with the donor are labeled and shown in cyan as bond models. (C) Comparison of the acceptor binding pockets of five UGTs. Catalytic residues and acceptor are in the UGT78G1 structure. Residue numbers in UGT78G1 are labeled, and the UGT72B1 unique long loop (~ residue 315) is also labeled.

The N- and C-terminal domains of the UGTs enzymes are organized in a cleft in which are found the donor binding pocket containing the PSPG motif and the acceptor binding site (Figure 14 A and B). The nucleotide sugar donor mainly interacts with the C-terminal domain whereas the acceptor mainly binds to the N-terminal one (Figure 14 B and C). Similarity between UGTs C-terminal domains is higher than that observed between their N-terminal domains, presumably because UGTs exhibit a **narrow selectivity toward donor, conversely to their higher acceptor promiscuity** (Figure 14 A and C).

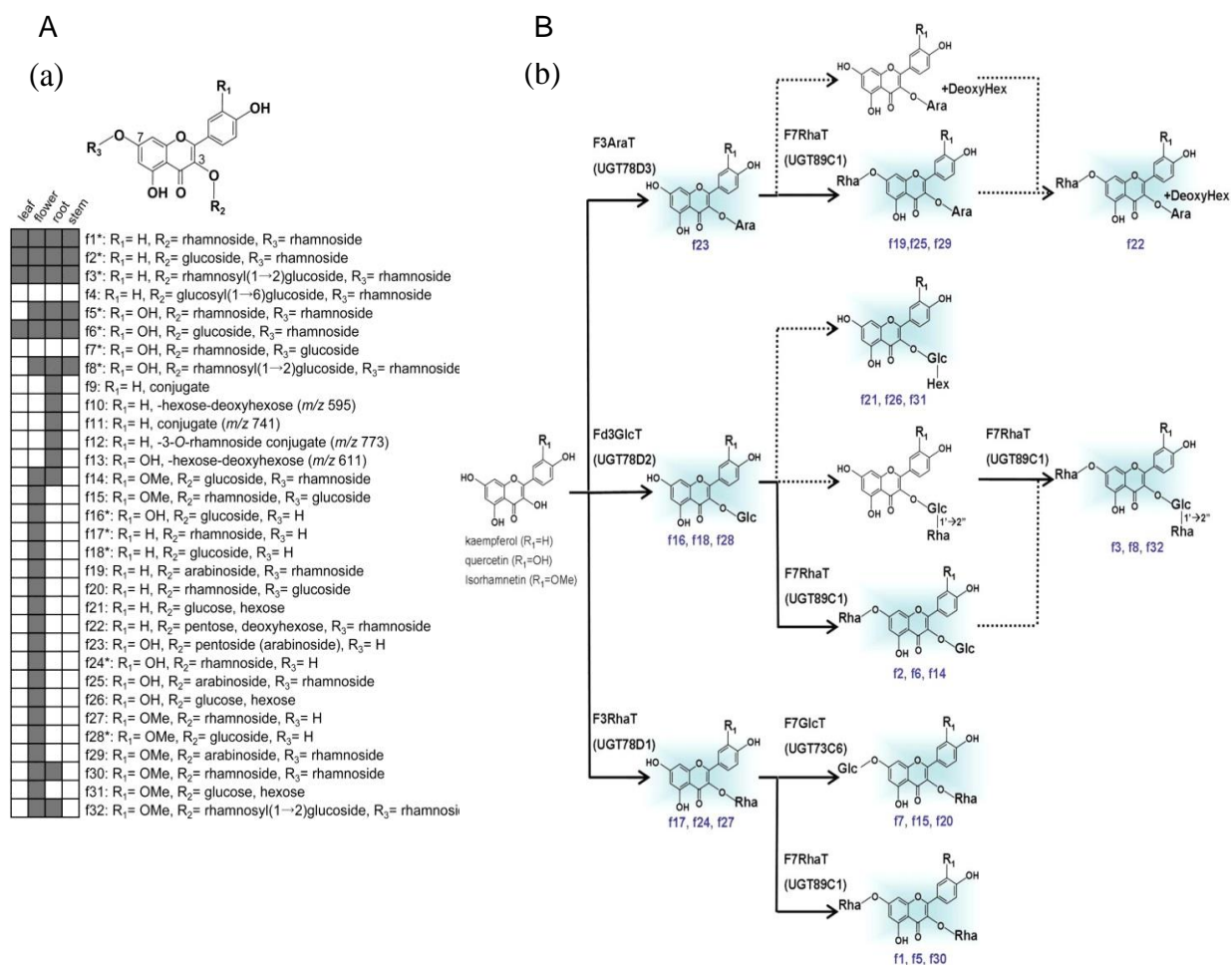
### II.1.3. Glycosylation of flavonoids in plants

Flavonoid glycosylation performed by UGTs are responsible for the physicochemical and biological properties of the flavonoid molecules which accumulated in cells. Thus, **sugar**

**decorations vary between plant species but also among tissues of a same species.** For example, in *Arabidopsis thaliana*, the abundance of each flavonoid glycosides follows a strict spatial regulation (Tohge et al., 2007; Yonekura-Sakakibara et al., 2008). Kaempferol and quercetin di-glycosides (di-rhamnosides or glucoside/rhamnoside in positions 3 and 7) and tri-glycosides (di-rhamnosides and mono-glucoside at positions 3, 2" and 7) are commonly distributed in all tissues (Yonekura-Sakakibara et al., 2008). But flavonol-3-O-glucoside and arabinosyl flavonols specifically accumulate in flowers. Moreover, in seeds, quercetin-3-O-rhamnoside, biflavonols (dimers of quercetin-rhamnoside), epicatechin and procyanidins accumulated in the seed coat, whereas diglycosylated flavonols were essentially detected in the embryo (Routaboul et al., 2006).

Consequently, it is **difficult to establish a general pathway for flavonoid glycosylation in plants.** However, in most of the plants, 3-O-glycosylations of flavonoids occur prior to glycosylation at other positions. These particular glycosylations were shown to be catalyzed by regiospecific UGTs named UDP-sugar:flavonoid-3-O-glucosyltransferases as those characterized from *Gentiana triflora*, (Tanaka et al., 1996), *Perilla frutescens* (Gong et al., 1997), *Vitis vinifera* (Ford et al., 1998), *Petunia hybrid* (Yamazaki et al., 2002) or *Glycine max (L.) Merr.* (Kovinich et al., 2010).

A **functional genomics approach** was developed in 2008 to establish flavonol's glycosylation pathway in *A. thaliana* (Yonekura-Sakakibara et al., 2008). First of all, a **flavonol profiling** in the flowers, leaves, stems, and roots of *A. thaliana* was conducted to identify all of the flavonol derivatives (Figure 15). Then, candidate genes potentially encoding enzymes involved were listed to achieve a more comprehensive annotation of the flavonol pathway in *Arabidopsis thaliana*. Thus, a **transcriptome coexpression analysis** was conducted. Generally, genes involved in certain metabolic processes are coordinately regulated. Genes with identified metabolic functions were selected as "guide". The coexpression analysis provides a network of the relationships between "guide" and candidate genes that serves to create the criteria by which gene functions can be predicted. **Together with metabolic profiling analyses,** this strategy efficiently narrowed the pool of listed candidate genes.



**Figure 15. Flavonol glycosylation pathway in *Arabidopsis thaliana*, adapted from (Yonekura-Sakakibara et al., 2008).** (a) Flavonol glycosides identified in *Arabidopsis thaliana*. Asterisks indicate that the compounds were identified based on a comparison of retention times and UV/mass spectra of the standards used in this study.  $R_1=H$ , kaempferol;  $R_1=OH$ , quercetin;  $R_1=OMe$ , isorhamnetin. The presence of a compound in the indicated tissue is denoted by a gray box. (b) Proposed comprehensive flavonol glycosylation pathway in *Arabidopsis thaliana*. Each number in blue corresponds to the compounds shown in Figure 3. The compounds detected in *Arabidopsis* are shown in a blue background. Dotted lines indicate proposed but unidentified pathways. F3AraT, flavonol 3-O-arabinosyltransferase; Fd3GlcT, flavonoid 3-O-glucosyltransferase; F3RhaT, flavonol 3-O-rhamnosyltransferase; F7GlcT, flavonol 7-O-glucosyltransferase; F7RhaT, flavonol 7-O-rhamnosyltransferase; Ara, arabinose; DeoxyHex, deoxyhexose; Glc, glucose; Hex, hexose; Rha, rhamnose.

The **specific identification of genes functions** was accomplished using traditional reverse genetics, gene knockout and/or biochemical characterization of the recombinant proteins (Figure 15). As metabolic engineering of flavonoid pathways in plants and microorganisms gain in interest, such a combined strategies is hoped to increase the **knowledge on flavonoid glycosylation pathways, required for engineering approaches** (Wang et al., 2011).

## II.2. Production of flavonoid glycosides using glycosyltransferases (GTs)

An increasing attention has been paid to plant glycosyltransferases, especially to access to specific structures of glycoconjugates. **Wild-type and engineered UGTs** have been employed to achieve **tailored productions of flavonoid glycosides** using recombinant and purified enzymes as well as whole-cell systems (Lim, 2005).

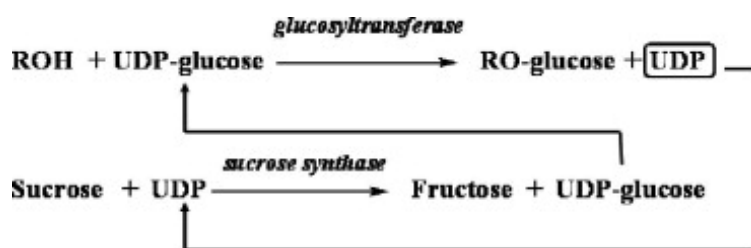
### II.2.1. Recombinant GTs for in vitro production of flavonoid glycosides

**Only few plant UGTs have been heterologously expressed in microbial cells for flavonoid glycosylation reactions.** Noguchi et al., in 2008b, isolated a glycosyltransferases from *Lycium barbarum* (GT73A10). The enzyme was shown to catalyze the regiospecific glucosylation of (+)-catechin (2.2 mM), using UDP-glucose (5.4 mM) as donor. A single transfer reaction product was obtained, (+)-catechin-4'-O- $\beta$ -D-glucopyranoside, with an overall conversion yield of 83%. A flavonol-specific glycosyltransferase from *Arabidopsis thaliana* (UGT78D1), was also expressed in *E. coli*, purified and tested to produce various flavonol glucosides (Ren et al., 2012). Quercetin, myricetin and kaempferol were totally converted into their 3-O-glucosylated forms in reactions (100 mL) containing UDP-glucose (4 mM), and the flavonols (0.5 mM).

**Microbial glycosyltransferases were also tested for in vitro production of flavonoid glycosides** (Gurung et al., 2013; Hyung et al., 2006; Ko et al., 2006a). For example, Hyung Ko et al. (2006) used the recombinant and purified His-tagged BcGT1 from *Bacillus cereus* to glycosylate of apigenin, genistein, luteolin, naringenin, kaempferol, and quercetin (70  $\mu$ M) using of UDP-glucose (500  $\mu$ M) as glucosyl donor. In contrast to the previous examples, a mix of monoglucosylated products was obtained, mainly consisting of 7-O, 3-O or 4'-O-glucosides, depending on the aglycon considered.

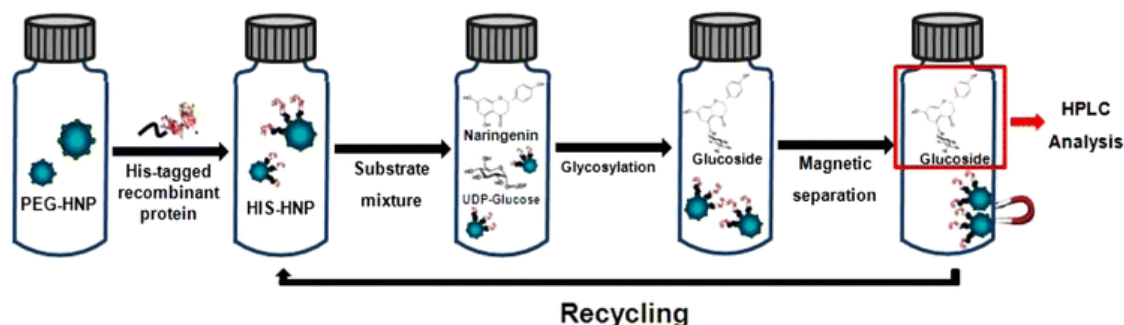
*In vitro* enzymatic productions of flavonoid glycosides using recombinant UGTs require activated sugar donors that must be supplied in the reaction mix. These **activated donors are usually expensive and difficult to obtain in sufficient quantity for large-scale productions.** Furthermore, UDP produced during the reaction course, may act as a potent inhibitor of UGT activity.

To avoid these problems, a **one-pot reaction system** coupling CaUGT2 from *Catharanthus roseus* and the sucrose synthase from *Arabidopsis thaliana* (AtSUS1), was proposed to ensure the **UDP removal and the regeneration of UDP-glucose** (Masada et al., 2007). This system was used to produce quercetin glycosides, with a recombinant flavonoid 7-O-glucosyltransferase from *S. baicalensis* (F7GT). Production of apigenin 7-O-glucoside catalyzed by the recombinant F7GT was drastically enhanced with the addition of AtSUS1 in the reaction mixture (Figure 16). After 30 min incubation, apigenin had been completely converted to apigenin 7-O-glucoside whereas only 50% remained non glucosylated when no AtSUS1 was added. Moreover, prolonged incubation of apigenin with F7GT plus AtSUS1 yielded a second minor product identified as apigenin 7, 4'-O-diglucoside, never obtained before with F7GT. Thus, this method is **efficient for the synthesis of flavonoid glucosides but is still limited for large-scale productions**. Indeed, **NDP-sugars must be initially supplied in reaction mixture**. In addition, the **often expensive production and purification of the enzymes** required for each reaction also impact the economic balance of such system.



**Figure 16. General concept of the one-pot system involving two enzymes for efficient synthesis of glucosides from flavonoids** (Masada et al., 2007).

To **limit enzyme cost, recycling of immobilized enzymes** was investigated. The 6xHis-tagged recombinant oleandomycin glycosyltransferase (OleD UGT) from *Streptomyces antibioticus* was purified and immobilized on **magnetic nanoparticles** for the preparation of flavonoid glycosides (Choi et al., 2012; Lee et al., 2010). Apigenin, luteolin, kaempferol, naringenin, daidzein, and genistein were tested for glycosylation (50  $\mu$ M) with UDP-glucose donor (500 $\mu$ M) and the immobilized enzyme (50–400  $\mu$ g). The reactions led to 35%, 33%, 100%, 42.5%, 0.9% and 6.6% of conversion, respectively. The immobilized OleD UGT was successfully re-used 7 times for naringenin glycosylation (Figure 17) without drastic loss of activity. Indeed, the relative conversion rates (compared to the first use) remain above 60%.



**Figure 17. Re-use of the nano-immobilized OleD GT for the synthesis of the glycosylated naringenin.** The reaction scheme represents the nano-immobilization of recombinant OleD GT on magnetic nanoparticles on the basis of the affinity between his-tag and Fe<sub>3</sub>O<sub>4</sub>/silica/NiO nanoparticles (Choi et al., 2012)

**Combination of multi-enzymatic system and enzyme immobilization may appear quite promising** to overcome the unavailability of NDP-sugars together with expansive use of pure enzymes. But progresses in metabolic engineering and pathways knowledge have led to the development and use of promising whole-cell systems for the production of flavonoid glycosides.

### II.2.2. Metabolic engineering of whole-cell production systems

Productions in **living cells combine the advantages of cofactor regeneration while avoiding expensive and laborious enzyme isolation** (Křen and Thiem, 1997). For this purpose whole-cell systems have been developed to produce some flavonoid glycosides, using either plant cells or microorganisms as biological factory (often also referred to as ‘combinatorial biosynthesis’).

**Concerning plant whole-cell biocatalysis**, exogenous aglycons are incubated with plant cell for glycosylation by endogenous UGTs that are either present as **constitutively expressed enzymes or induced by the exogenous aglycon**. The major advantage of such a bioconversion is that living cells provide the activated sugar donors necessary for glycosylation. But, **very few examples of flavonoid glycosylation using plant cells were reported**. Incubation of *Nicotiana tabacum* cells with quercetin yielded quercetin-3-O- $\beta$ -D-glucoside (51%), quercetin-3-O-(6-O-malonyl)- $\beta$ -D-glucoside (10%), quercetin-3-O-[6-O-( $\alpha$ -L-rhamnosyl)]- $\beta$ -D-glucoside (3%), quercetin-3,4'-O- $\beta$ -D-diglucoside (1%), and 3,7-O- $\beta$ -D-diglucoside (1%) (Hamada, 2007). With (-)-epicatechin, three glycosylation products were isolated : (-)-epicatechin-3'-O- $\beta$ -D-glucoside (38%), (-)-epicatechin-5-O- $\beta$ -D-glucoside (7%), and (-)-epicatechin-7-O- $\beta$ -D-glucoside (15%). Similar glycosides and yields were obtained when (+)-



catechin was used as acceptor instead of the epicatechin (Hamada, 2007). Biocatalytic glycosylation of daidzein with cultured cells of *Catharanthus roseus* was attempted by Shimoda et al., 2011. Incubation of *C. roseus* cells with daidzein for five days produced daidzein-7-O- $\beta$ -glucoside (30%), daidzein-7-O- $\beta$ -primeveroside (5%), and daidzein 4'-O- $\beta$ -glucoside (2%).

In recent years, an **increasing number of reports describe the use of microbial cells expressing heterologous plant UGTs to produce flavonoid glycosides**. Engineered microbial species such as *Escherichia coli* and *Saccharomyces cerevisiae* are useful because of their fast doubling times compared to plant species (minutes vs. days), inexpensive carbon sources, easy genetic modification and well established scale-up technologies (Chang and Keasling, 2006; Chang et al., 2007; Roberts, 2007).

In 2004, 91 different **recombinant *E. coli* strains producing UGTs** from *Arabidopsis thaliana* were generated to **synthesize glucosides of quercetin by taking advantage of the endogenous UDP-glucose** (Lim et al., 2004). 29 UGT-expressing strains were shown to produce quercetin-3-O-, 7-O-, 3'-O-, and/or 4'-O-monoglucosides. Among them, 14 strains expressed highly regiospecific UGTs glucosylating only one position on quercetin (C3 or C7); 14 strains glucosylated two or three positions; and one strain was able to glucosylate all four positions, revealing a far less constrained regiospecificity of the recombinant UGTs toward quercetin. Seven of the recombinant strains that were active on quercetin, were used in larger scale culture. From 2 g of quercetin and in the absence of supplied activated sugar, the quercetin glucosides were produced at concentration levels ranging from 0.19 to 10.9  $\mu\text{g} \cdot \text{mL}^{-1}$  in supernatant (Table 4).

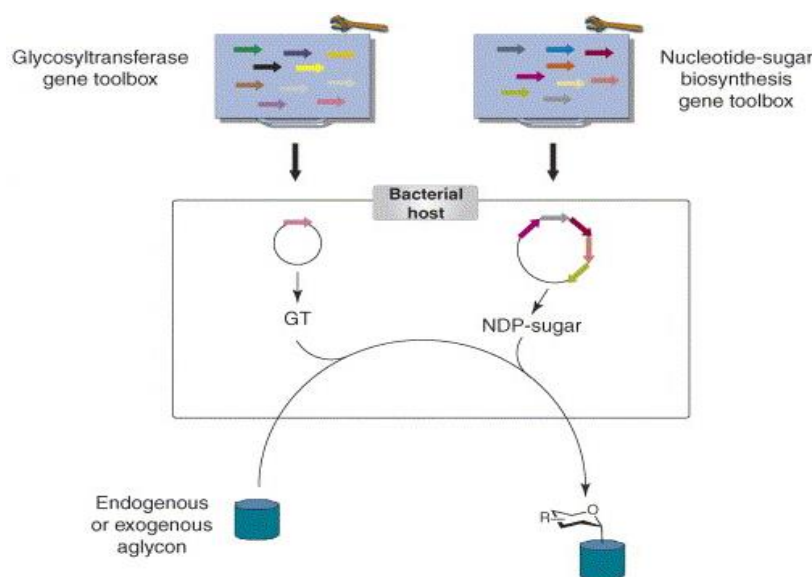
**Table 4. Glucosyltransferases tested as whole-cell biocatalysts for quercetin glucosylation (Lim et al., 2004).**

UGT	Number of products formed	Culture medium ( $\mu\text{g/mL}$ in supernatant)					
		3-O-glc	7-O-glc	3'-O-glc	4'-O-glc	3,7-di-O-glc	7,3'-di-O-glc
73B3	1	8.96	—	—	—	—	—
78D2	1	6.83	—	—	—	—	—
84B1	1	—	4.83	—	—	—	—
74F1	2	—	0.53	—	0.19	—	—
76E12	3	1.41	4.74	—	—	0.71	—
89B1	3	1.08	2.36	—	0.94	—	—
71C1	3	—	2.12	8.00	—	—	10.90

In the same way, Willits et al. (2004) conducted the culture (500 mL) of a recombinant *E. coli* strain expressing the UGT73B2 from *Arabidopsis thaliana*, and supplemented with 150 mg quercetin. They achieve the production of quercetin 7-O-glucoside as major product, and

quercetin 3,7-O-diglucoside, as a minor product with an overall conversion rates of 52%. When using a recombinant *E. coli* strain producing the RF5 glucosyltransferase from *Oryza sativa*, all the quercetin (100  $\mu$ M) was converted to its corresponding 3-O-glucoside (Kim et al., 2006b). A recombinant strain producing the glucosyltransferase from *Xanthomonas campestris* (XcGT-2) was also reported to glucosylate 70% of the quercetin supplied in the culture at 30  $\mu$ M through a single product formation, suspected to be a quercetin 3'-O-glucoside (Kim et al., 2007).

Naturally, UGTs can act with a broad variety of glycosyl unit donors, but, **when using *E. coli* as a recombinant biocatalytic host, only UDP-glucose is naturally available as a glycosyl donor** in the bacteria. Thus, **metabolic engineering approaches** allowing the co-expression of genes involved in the **formation of various NDP-sugars** and GTs have been proposed (Figure 18; Blanchard and Thorson, 2006).



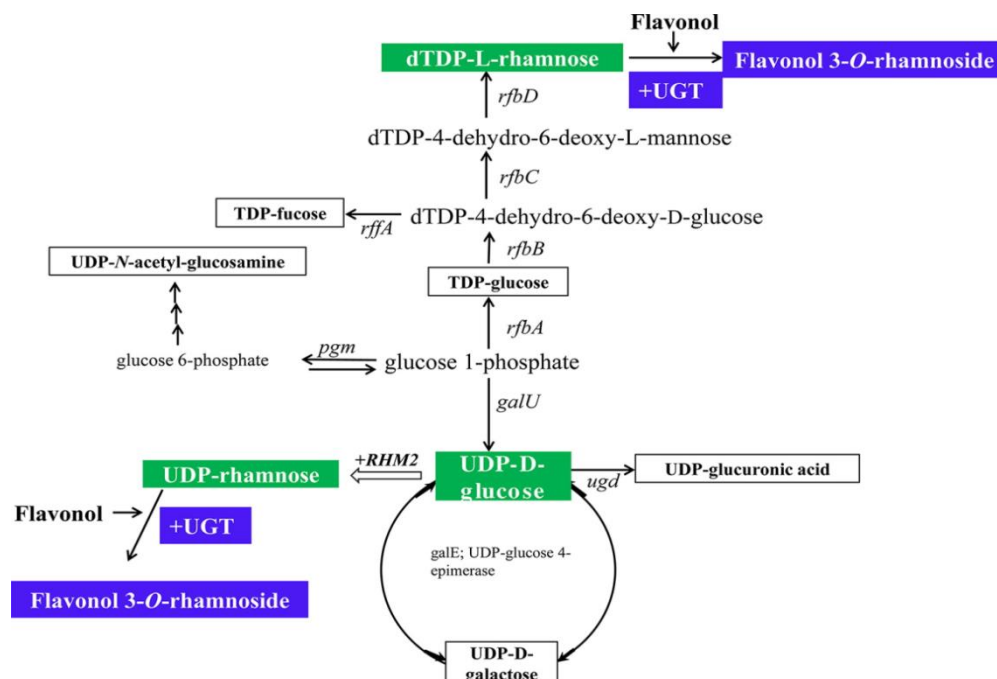
**Figure 18. A glycosyltransferase gene toolbox and a nucleotide-sugar biosynthesis gene toolbox provide the basis for a general in vivo pathway engineering strategy to generate novel glycosylated natural products (Blanchard and Thorson, 2006).**

Pandey et al. (2013) reported a metabolic engineering approach for the xylosylation of quercetin. The *E. coli* strain was first engineered to produce a cytoplasmic pool of UDP-xylose. Genes involved in the metabolic use of glucose-6-phosphate, an initial precursor of NDP-sugar biosynthetic pathway, were knocked out. Then, the four required UDP-xylose biosynthetic genes were inserted in bacterial genome. The gene encoding the glycosyltransferase ARGt-3 from *A. thaliana* was finally inserted in the genome. Thus, quercetin (100  $\mu$ M) was added to the production culture and nearly all the quercetin (98 %) was



transformed into its 3-O-xylosyl derivative to yield approximately 43 mg of product per liter of culture.

Another use of the *A. thaliana* glycosyltransferase ARGt-3 is reported for the production of 3-O-rhamnosyl quercetin using an engineered *Escherichia coli* strain (Simkhada et al., 2010a). In this study, the **strain was modified to express an effective biosynthetic pathway of a TDP-L-rhamnose donor** (Figure 19). The production yield, corresponding to an average conversion rate of 30%, reached 24 mg.L<sup>-1</sup> of 3-O-rhamnosyl quercetin, from a feeding of 0.2 mM (60 mg) of the aglycon.



**Figure 19. Nucleotide sugar biosynthesis pathway and production of UDP-rhamnose in engineered *E. coli*** (Kim et al., 2012). *galU*, UTP-glucose 1-phosphate uridylyltransferase; *pgm*, phosphoglucomutase; *ugd*, UDP-glucose 6-dehydrogenase; *rffA*, dTDP-4-oxO-6-deoxy-d-glucose transaminase; *rfbA*, dTDP-glucose pyrophosphorylase; *rfbB*, dTDP-glucose 4,6-dehydratase; *rfbC*, dTDP-4-dehydrorhamnose 3,5-epimerase; *rfbD*, dTDP-4-dehydrorhamnose reductase; and *RHM2*, rhamnose synthase 2.

Kim et al. (2012) also reported the synthesis of quercetin 3-O-rhamnoside. They engineered *E. coli* to favor the synthesis of UDP-rhamnose donor (Figure 19) used as substrate of *Arabidopsis thaliana* glycosyltransferase AtUGT78D2. **To avoid cell lysis due to high concentration of quercetin, gradually feeding of the aglycon was carried out up to 500  $\mu$ M (143 mg.L<sup>-1</sup>).** Quercetin 3-O-rhamnoside reached 150 mg.L<sup>-1</sup> with a quercetin conversion corresponding to 65%. In a subsequent study, the authors attempted the synthesis of 3,7-O-dirhamnoside quercetin, using the same **engineered *E. coli* strain expressing an additional**

**UGT** from *Arabidopsis thaliana* (AtUGT89C1) displaying a different regioselectivity for quercetin, compared to AtUGT78D2 (Kim et al., 2013a). They achieved the biosynthesis of 67.4 mg.L<sup>-1</sup> of quercetin 3,7-O-dirhamnoside from 30 mg.L<sup>-1</sup> (100 µM) of the aglycon, corresponding to its total conversion.

In addition to glucose, rhamnose or xylose, other sugar moieties can be linked to flavonoids. To illustrate, a 7-O-glucuronyl quercetin was produced, by overexpressing the *atGt-5* gene coding for an *Arabidopsis thaliana* glycosyltransferase, in a **metabolically engineered *E. coli* strain producing UDP-glucuronic acid** (Simkhada et al., 2010b). Simkhada et al. (2010a) also produced a 3-O-allosyl quercetin from endogenous TDP-6-deoxy-D-allose, as sugar donor, using an engineered *E. coli* strain expressing recombinant glycosyltransferase from *A. thaliana* (ARGt-3). ***E. coli* strains were also engineered to synthesize a pool of unnatural sugar donor**, such as dTDP-6-deoxytalose. The recombinant glycosyltransferases from *A. thaliana* (AtUGT78D1) expressed in the engineered *E. coli* strain produced around 98 mg.L<sup>-1</sup> of quercetin 3-O-(6-deoxytalose) from the 75.5 mg.L<sup>-1</sup> of quercetin (250 µM) supplied to the culture (Yoon et al., 2012).

Besides these studies mainly focused on *in vivo* quercetin glycosylation, bioconversion involving engineered *E. coli* strains have been reported for the glycosylation of other flavonoids, like kaempferol (He et al., 2008; Kim et al., 2006b; Simkhada et al., 2010a), apigenin (Thuan et al., 2013a), baicalein (Thuan et al., 2013a), myricetin (Thuan et al., 2013b), genistein (He et al., 2008), luteolin (He et al., 2008) and naringenin (Simkhada et al., 2009, 2010b).

**Nowadays, *Escherichia coli* still remains the preferred host for *in vivo* biosynthesis of flavonoid glucosides, but yeast strains, like *saccharomyces cerevisiae* also present great interests.** Indeed, *S. cerevisiae* is an **eukaryotic GRAS microorganism** (Generally Regarded As Safe), which gives yeast an advantage compared to *E. coli* with respect to production of flavonoid glycosides, using plant UGT, for their use as dietary supplements and pharmaceuticals (Limem et al., 2008).

A *Saccharomyces cerevisiae* yeast strain expressing a recombinant *Dianthus caryophyllus* flavonoid glucosyltransferase (DicGT4) was used as a whole-cell biocatalyst for the production of naringenin glycosides from endogenous UDP-glucose (Werner and Morgan, 2009). Culture medium fed with 136 mg.L<sup>-1</sup> of naringenin yielded the synthesis of around 40 mg.L<sup>-1</sup> of two products: naringenin-7-O-glycoside and naringenin-4'-O-glycoside. However, naringenin-7-O-glycoside concentration decreased during the fermentation, due to its **hydrolysis to naringenin by *S. cerevisiae* cells**, making naringenin-4'-O-glycoside the major

product. In a following study, the authors optimized culture conditions (10 mM orotic acid supplied) and achieved to produce 155 mg.L<sup>-1</sup> of these naringenin glycosides, which represented a 71% conversion of the aglycon (Werner and Morgan, 2010).

The use of **wild-type glycosyltransferases to produce diversified flavonoid glycosides takes advantage of the GTs natural promiscuity for both donors and acceptors**. In general, UGTs exhibit activity with several UDP-sugars, but often with a preference for one specific UDP-sugar. Concerning aglycon acceptors, some UGTs are highly specific and catalyze glycosylation of only one or a few acceptors, whereas other UGTs are more promiscuous and glycosylate a broad range of acceptors (Osmani et al., 2009). According to these considerations, **engineering of UGTs has gained interest to design GTs with controlled and/or original regioselectivity and/or improved efficiency**.

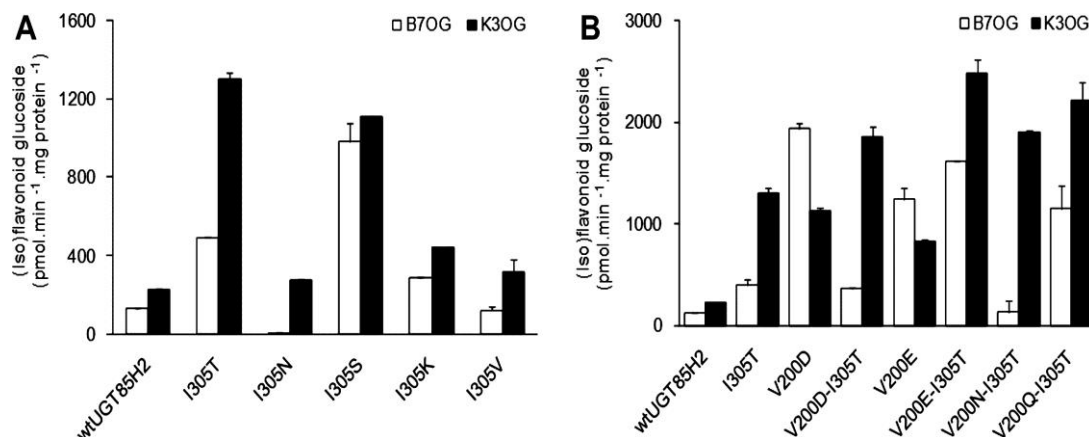
### II.2.3. Enzymatic engineering of GTs

To allow GTs engineering, High-Throughput Screening (HTS) assays were developed (Lee and Thorson, 2011; Park et al., 2009; Williams and Thorson, 2008; Williams et al., 2007) and led to patented applications for the glycodiversification of natural products (Thorson et al., 2009). However, only few examples of applications surprisingly concern flavonoid glycosylation.

#### II.2.3.a. Site-directed mutagenesis

In 2006, He et al. demonstrated that engineered glycosyltransferases with altered regioselectivity for flavonoids could be obtained through site-directed mutagenesis. They engineered the glycosyltransferase UGT71G1 from *Medicago truncatula* that catalyzes quercetin and genistein glycosylation from UDP-glucose. Incubation of UGT71G1 with quercetin led to the formation of the five possible monoglucosides, with the 3-O-glucoside predominating, whereas genistein's glucosylation led to 7-O-glucoside formation. **Based on the 3D structure** of UGT71G1, the authors generated a series of **single mutants** targeting few amino acids of the sugar donor or phenolic acceptor pockets. Of this library, Y202A mutant, harboring a **mutation in the acceptor binding pocket showed an altered acceptor regioselectivity**, compared to wild-type and exhibited a 15-fold higher activity toward the 3'-hydroxyl group of quercetin. Moreover, when genistein was used as acceptor, genistein 5-O-glucoside was formed in addition to the 7-O-glucoside. But, considering UGT71G1 **catalytic efficiency towards both flavonoids, no improvements** were obtained although the mutant kept 96% of the wild-type activity.

Another glycosyltransferase from *Medicago truncatula*, UGT85H2, was engineered using the same **structurally-guided site-directed mutagenesis approach** (Modolo et al., 2009b). This enzyme naturally catalyzes the regioselective formation of 3-O-kaempferol and 7-O-biochanin A glucosides, using UDP-glucose. Two positions, I305 and V200, identified as potentially crucial for UGT85H2 activity were separately and simultaneously submitted to mutagenesis, to generate **libraries of mono and double mutants**.



**Figure 20. Specific activity of UGT85H2 and its mutants toward kaempferol and biochanin A** (Modolo et al., 2009b). (A) Effect of I305 mutations on activity; (B) Effect of V200 mutations on activity. B7OG, biochanin A 7-O-glucoside; K3OG, kaempferol 3-O-glucoside.

Some mutants with **highly improved activities and different specific activities** toward biochanin A or/and kaempferol were obtained (Figure 20). Nevertheless, in this study, **no changes were observed in the glucosylation patterns**, as only 3-O-kaempferol and 7-O-biochanin A glucosides were detected.

Similarly, the BpUGT94B1a glucosyltransferase from *Bellis perennis* (red daisy) was engineered (Osmani et al., 2008). Wild-type BpUGT94B1 was shown to use UDP-glucuronic acid to specifically glycosylate the acceptor cyanidin 3-O-glucoside. BpUGT94B1 can also alternatively use UDP-glucose, but with a catalytic efficiency of less than 0.5% compared to that observed with glucuronic acid. As **no 3D structure BpUGT94B1 was available**, the authors constructed a **model using coordinates of two other plant GT crystal structures as scaffolds** to target residues for saturation mutagenesis. Three of the constructed mutants (R25S, R25G, R25K) presented a **modified donor specificity**. Indeed, they exhibited 0.5% to 2.5% of wild-type activity with UDP-glucuronic acid, but showed a 3-fold increase in activity with UDP-glucose and cyanidin 3-O-glucoside.

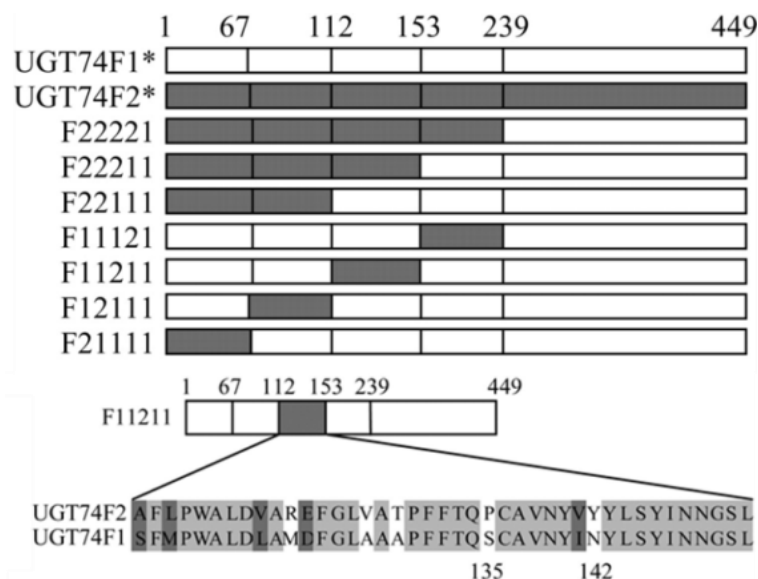
These three studies highlights the **contribution of site-directed mutagenesis to generate UGTs mutants with improved activity, modified regioselectivity** (He et al., 2006) **or altered sugar donor specificity** (Modolo et al., 2009b; Osmani et al., 2008). Nevertheless, **no changes in regioselectivity concomitant with significant catalytic improvements** were reported.

### *II.2.3.b. Domain swapping*

To improve both catalytic efficiency and selectivity, **domain swapping combined to molecular modeling and site-directed mutagenesis** has been experimented. Domain swapping is derived from a natural mechanism where one or more structural elements of enzymes, defined as "domains" and showing high amino acid sequence homology and folding similarity, are interchanged between partners to generate chimerical enzymes.

Cartwright et al. (2008) used a domain-swapping approach to **explore the determinants of regiospecific glycosylation** of quercetin, using two *Arabidopsis thaliana* UGTs, UGT74F1 and UGT74F2. The two enzymes share 76% amino acid sequence identity (90% similarity) and were predicted to share an identical secondary structure. UGT74F1 was reported to form monoglucosides targeting the C7, C3', and C4' hydroxyl groups of quercetin, whereas UGT74F2 synthesizes only quercetin 7-O- and 4'-O-monoglucosides. Both enzymes naturally use UDP-glucose as sugar donor. First, **four "shuffle points"** were identified at amino acid residues 67, 112, and 153 (located within the N-terminal Rossmann fold-like domain) and 239 (located within the linker connecting the two Rossmann fold-like domains) of the enzyme. **Seven chimeras** with sequence switched at these amino-acid positions were constructed using chimerical PCR (Figure 21)

Of these chimeras, F11211 (Figure 21) showed the greatest positive shift for quercetin 4'-O-glucoside formation. Then, **site-directed mutagenesis** further showed that **only one mutation (N142Y), not located in the active site, was sufficient to enhance regiospecificity** toward the C4' hydroxyl group of quercetin, luteolin, kaempferol, and apigenin. UGT74F1 N142Y mutant was further expressed in **engineered *E. coli* strain for a whole-cell preparative scale production** of quercetin 4'-O-glucoside. The final purified amount reached 9.9 mg.L<sup>-1</sup> of culture (conversion rate of 17%). Other glucosides were also produced in lesser amounts (representing less than 5% of the glucosylated products) indicating that **in vivo and in vitro specificity were similar**.



**Figure 21. Schematic representation of parent UGT74F1\*, UGT74F2\*, and the generated chimeras** (adapted from (Cartwright et al., 2008). Shuffle points are given as their equivalent amino acid position, partitioning the primary sequence into five segments. The amino acid alignment of UGT74F1 and UGT74F2 from position 112 to 153 is given. Amino acid identity is shown in light gray, similarity in dark gray.

A second study using **combination between domain swapping and site-directed mutagenesis** was conducted with UGT78D2 and UGT78D3 from *Arabidopsis thaliana* (75 % identity) to **modify sugar donor selectivity** (Kim et al., 2013b). Indeed, the two enzymes glycosylate the C3 hydroxyl group of flavonols using UDP-glucose and UDP-arabinose, respectively. Finally, **a single mutation (M288A)** was showed to be **sufficient to create a new activity** for UDP-arabinose, with no effect on the use of UDP-glucose. Moreover, the **mutation effect** was reported to be **indirect** and mediated through the opening access of residue Q381 to form a hydrogen bond stabilizing UDP-arabinose in the active site of AtUGT78D2.

Those two reports set out the combination between domain swapping, molecular modeling and site-directed mutagenesis as a **promising engineering method to identify amino acid residues not necessarily directly involved in the catalysis but essential for the creation of new specificities towards flavonoid acceptors or sugar donors**.

### II.2.3.c. Random directed evolution

Recent developments of efficient **high throughput screening assays** to discriminate larger UGTs libraries have opened the way to directed evolution strategies.

The glycosyltransferase OleD, from *Streptomyces antibioticus*, was submitted to directed evolution through an **error-prone PCR strategy to improve its catalytic efficiency toward a panel of acceptors** (Williams et al., 2007). An OleD library mutant containing one or two amino acid mutations per variant was constructed. A relatively small library (~1,000 colonies) of variants was initially screened using a fluorescence based GT assay with UDP-glucose as sugar-donor and fluorescent 4-Methylumbelliferone as acceptor. Selected mutants were then assayed with a panel of acceptors, including kaempferol, daidzein and genistein (200  $\mu$ M). One OleD monomutant (P67T) exhibited 3.8-, 8.5- and 10-fold higher conversion degree of daidzein and genistein, and kaempferol respectively, compared to that obtained with the wild-type enzyme.

The glycosyltransferase BcGT-1 from *Bacillus cereus* was also submitted to **directed evolution to create modified enzymes with new regioselectivity** (Jung et al., 2010a). The authors wanted to produce kaempferol monoglucosides, whereas wild-type catalyzed the formation of kaempferol-3,7,4'-O-triglucosides. They used an error-prone PCR strategy to create a random library in which a total of 150 clones were screened. One monomutant (P343L) and one triple mutant (M161T-Q193G-M361T) were shown to be specific of kaempferol 3-O-glucoside formation. Interestingly, none of the mutated residues were close to the substrate binding sites. This indicates that a **long distance effect of the mutations** might be involved in the mutant selectivity.

**In conclusion, glycosyltransferases main physiological role in the flavonoid metabolism was exploited to produce existing or new flavonoids structures with success but also limitations. To overcome donor low availability and high cost issues, efforts were paid in production, metabolic and enzyme engineering. This led to development of relatively efficient productions of flavonoids. Nevertheless, the flavonoid glycosides produced, using GTs, were essentially  $\beta$ -glucosides.**

**To expand the available flavonoids with original glycosylation patterns and access to bioactive compounds with new properties and biological effects, Glycoside Hydrolases have been studied as biocatalysts.**

## II.3. Glycoside-Hydrolases for flavonoid glycosylation

As glycosyltransferase, Glycoside Hydrolases represent a widespread group of pro- and eukaryotic enzymes that play a pivotal role in many biological processes. The vast majority of Glycoside Hydrolases is **involved in the metabolism of oligosaccharides or the degradation of endogenous and exogenous glycosides**. Thus, they are often referred as **glucosidases**. In **plant** metabolism the majority of GH enzymes are involved in cell wall polysaccharide metabolism (Minic and Jouanin, 2006). Some plant GHs were also shown to be responsible of the flavonoids osidic units hydrolysis (Minic, 2008). However, they **can also achieve the reverse reaction and transfer glycosyl moieties onto flavonoids**, such as reported for one of the  $\beta$ -D-glucosidase of *Impatiens balsamina* L flowers (Harborne and Mabry, 1982; Koes et al., 1994). Despite their essential role the number of GHs expressed in plant and their substrate specificities still largely remain unknown (Minic, 2008). **Microbial** GHs are naturally involved in nutrient acquisition and were mostly studied for their natural role in the hydrolysis plant cells wall glycosidic fraction (Gilbert, 2010). As for plants, some specific hydrolases can also catalyze the transfert of a glycosidic moiety from one glycoside to another (or another nucleophile), rather than water (Vocadlo and Withers, 2000). **Flavonoids were largely reported to inhibit many GHs**, as seen in the previous part (Table 3, p32). However, many studies also reported the important role of GHs expressed by mammalian gut microflora in the metabolization of flavonoids. Thus, GHs exhibiting a transglycosylase activity were assayed as natural or engineered biocatalysts for flavonoid glycosylation, in alternative to glycosyltransferases.

### II.3.1. Glycoside-Hydrolases (GH) short overview

The term Glycoside Hydrolases brings together numerous enzymes and represents the **largest group in the CAZy classification** (47%, Figure 13). It is subdivided in **132 families** that can also be brought together in clans, which represent groups of families sharing a same fold and catalytic machinery (Table 5).

GH enzymes may adopt **two different main mechanisms for the cleavage of the osidic bond**. They were initially described by Koshland and referred to as **retaining or inverting GHs** (Koshland, 1953), depending on the variation of the anomeric carbon configuration during the reaction (Figure 22).



**Table 5. The established Glycoside Hydrolases clans of related families in the CAZy database** (adapted from <http://www.cazy.org/>)

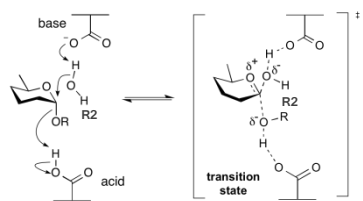
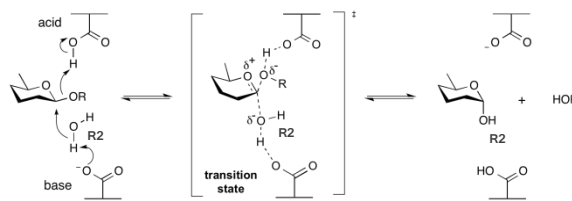
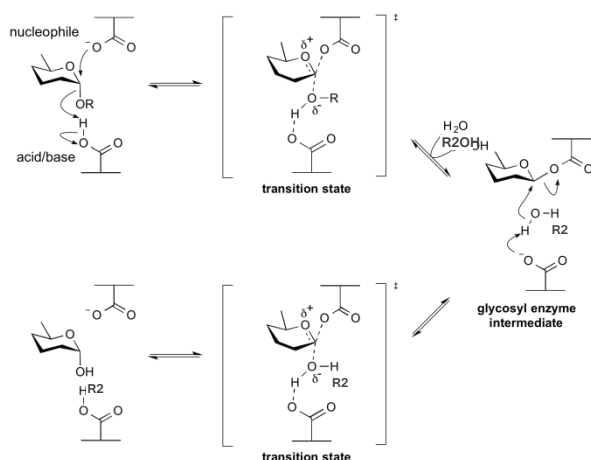
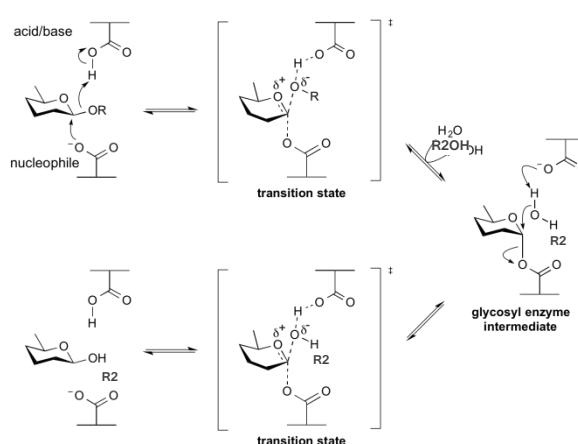
GH clans	Shared structural features	Related GH families
GH-A	( $\beta/\alpha$ ) <sub>8</sub>	1, 2, 5, 10, 17, 26, 30, 35, 39, 42, 50, 51, 53, 59, 72, 79, 86, 113, 128
GH-B	$\beta$ -jelly roll	7, 16
GH-C	$\beta$ -jelly roll	11, 12
GH-D	( $\beta/\alpha$ ) <sub>8</sub>	27, 31, 36
GH-E	6-fold $\beta$ -propeller	33, 34, 83, 93
GH-F	5-fold $\beta$ -propeller	43, 62
GH-G	( $\alpha/\alpha$ ) <sub>6</sub>	37, 63
GH-H	( $\beta/\alpha$ ) <sub>8</sub>	13, 70, 77
GH-I	$\alpha+\beta$	24, 46, 80
GH-J	5-fold $\beta$ -propeller	32, 68
GH-K	( $\beta/\alpha$ ) <sub>8</sub>	18, 20, 85
GH-L	( $\alpha/\alpha$ ) <sub>6</sub>	15, 65, 125
GH-M	( $\alpha/\alpha$ ) <sub>6</sub>	8, 48
GH-N	$\beta$ -helix	28, 49

Hydrolysis or synthesis of a glycoside with **net inversion** of anomeric configuration is generally achieved via a one step single-displacement mechanism involving oxocarbenium ion-like transition states (Figure 22 A). The reaction typically involves acid/base assistance from two amino acid side chains, normally glutamic or aspartic acids.

Reactions with **net retention** of the anomeric configuration are most commonly achieved via a two-step, double-displacement mechanism involving the formation of a covalent glycosyl-enzyme intermediate (Figure 22 B).

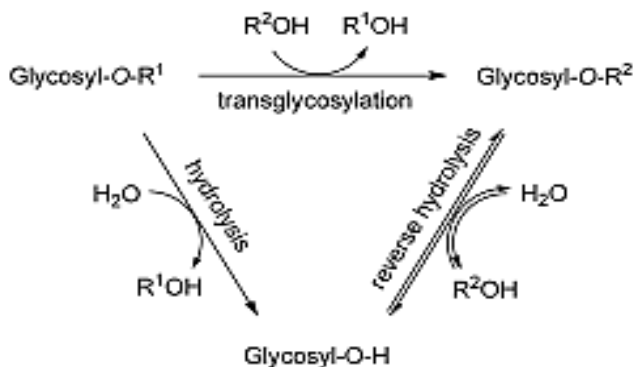
**In the first step**, the nucleophile residue (usually an aspartic or glutamic acid) attacks the anomeric center to form a glycosyl enzyme intermediate. At the same time, the second residue (acid/base) acts as an acid catalyst and protonates the glycosidic oxygen to release the aglycon.

**In the second step**, the acid/base residue acts as a base catalyst, deprotonating the hydroxyl group of the second acceptor molecule (water for hydrolysis) to allow the attack of the glycosyl enzyme intermediate and the product release.

**(A) Inverting mechanism for glycoside hydrolases****Inverting  $\alpha$ -glycoside hydrolases****Inverting  $\beta$ -glycoside hydrolases****(B) Retaining mechanism for glycoside hydrolases****Retaining  $\alpha$ -glycoside hydrolases****Inverting  $\beta$ -glycoside hydrolases**

**Figure 22. General Glycoside Hydrolases mechanisms** depending on (A) inverting  $\beta$  or  $\alpha$ -Glycoside Hydrolase (B) retaining  $\beta$  or  $\alpha$ -glycoside Glycoside Hydrolase proceeding through an intermediate glycosyl-enzyme. R and R2 may be either hydrogen or substituted groups (adapted from Davies and Henrissat, 1995; Rye and Withers, 2000; Withers, 2001; Vasella et al., 2002; Vuong and Wilson, 2010).

According to these mechanisms **Glycoside Hydrolases can transfer the sugar moiety onto different acceptor molecules** (Figure 23). When the transfer is performed onto a water molecule, it corresponds to a hydrolysis activity and GH involved is often referred to as transglycosidase. The sugar moiety may also be transferred onto a hydroxylated acceptor (simple sugar, oligosaccharides or aglycon) through a **transglycosylation reaction**. In this case, GH is often referred to as **transglycosylase**. Transglycosylation is kinetically controlled, and during the reaction it is assumed that there is competition between the nucleophilic water and the acceptor substrate at the glycosyl-enzyme intermediate (Hancock et al., 2006; Nakatani, 2001).



**Figure 23. Synthetic and hydrolytic reactions catalyzed by Glycoside Hydrolases** (Desmet et al., 2012)

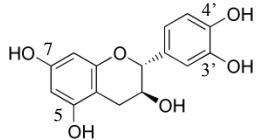
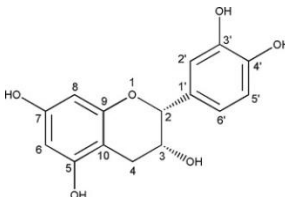
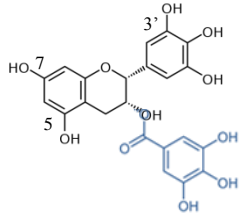
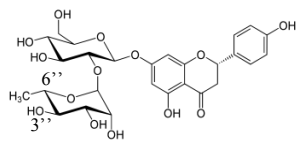
It is to note that other few represented reaction mechanisms were reported for GHs, such as the requirement to neighboring group participation as an intramolecular nucleophile (Terwisscha van Scheltinga et al., 1995), or to a NAD cofactor (Yip et al., 2004).

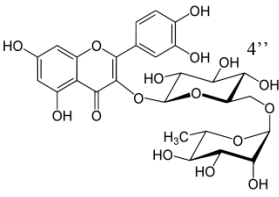
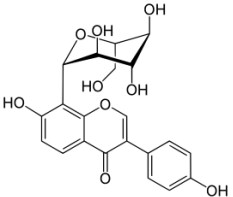
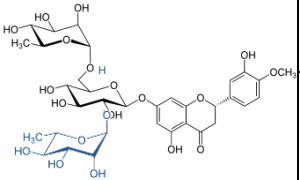
### II.3.2. Flavonoid Glycosylation

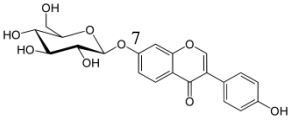
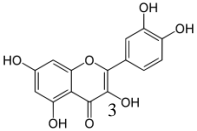
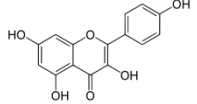
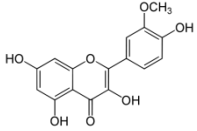
The use of GHs as specific catalysts to modify the substitution pattern of flavonoids is currently gaining attention. Indeed, in contrast to glycosyltransferases, Glycoside Hydrolases use **cheaper and renewable agro-resources as sugar donor and do not require cofactors**.

In plants, most of the secondary metabolites occur as  $\beta$ -glycosides. Thus, in nature, hydrolysis of these glycosylated secondary metabolites are mainly catalyzed by GHs referred to  $\beta$ -glycosidases (Walle, 2004). Their natural hydrolytic specificity is often used to help in the characterization, analysis or pretreatment of flavonoid mixtures. Isolation of the polyphenolic compounds using different extraction methods can also be coupled with enzymatic hydrolysis. Indeed, the GHs can hydrolyze polysaccharide fibers, favoring the release of the secondary metabolites and also change the glycosylation pattern to facilitate the extraction (Ara et al., 2013; Desmet et al., 2012).

**GHs exhibiting a transglycosylase activity have also been widely exploited for the *in vitro* synthesis of glycosylated flavonoids** (Desmet et al., 2012). Table 6 provides a summary of flavonoid glycosylation reaction carried out with GHs and report the effect of glycosylation on glycoside water solubility. We did not include in this table the results obtained with glucanases. A specific part will be devoted to these enzymes.

Flavonoid acceptors	Enzymes	Sugar donors	Flavonoid glycosides	Water Solubility Improvement	References
<b>(+)-catéchine</b> <i>Water solubility: ~22 µM</i>  <b>(-)-epicatechin</b> <i>Water solubility: ~22 µM</i> 	Cyclodextrin glucanotransferase from <i>Bacillus macerans</i> (GH13)	Starch, maltotriose	(+)-catechin-3'- <i>O</i> - $\alpha$ -D-glucopyranoside And other non-characterized (+)-catechin-glucosides	100-fold	Funayama et al., 1993; Sato et al., 2000
	$\alpha$ -amylase from <i>Bacillus subtilis</i> (GH13)	Maltotriose	Non-characterized (+)-catechin-glucosides	Not available	Nishimura et al., 1994
	$\alpha$ -glucosidase from <i>Bacillus stearothermophilus</i> (GH13)	Maltose	(+)-catechin-5- <i>O</i> - $\alpha$ -D-glucopyranoside (+)-catechin-7- <i>O</i> - $\alpha$ -D-glucopyranoside	Not available	Gao et al., 2000
	$\alpha$ -amylase from <i>Bacillus sp.</i> (GH13)	Maltodextrin	(+)-catechin-5- <i>O</i> - $\alpha$ -D-maltopyranoside (+)-catechin-7- <i>O</i> - $\alpha$ -D-maltopyranoside	Not available	
	GH with transglycosylase activity, derived from the genus <i>Trichoderma</i>	Carbohydrates containing a maltotriose residue	(+)-catechin-5- <i>O</i> - $\alpha$ -D-glucopyranoside * (+)-catechin-7- <i>O</i> - $\alpha$ -D-glucopyranoside (+)-catechin-4'- <i>O</i> - $\alpha$ -D-glucopyranoside (+)-catechin-4'- <i>O</i> - $\alpha$ -D-maltopyranoside (+)-catechin-3'- <i>O</i> - $\alpha$ -D-maltopyranoside	* 40-fold > 10-fold for all other glucosides	Ochiai et al., 2010
	Cyclodextrin glycosyltransferase from <i>Paenibacillus sp.</i> RB01 (GH13)	Starch, cyclodextrin, maltoheptaose	(-)-epicatechin-3'- <i>O</i> - $\alpha$ -D-glucopyranoside * (-)-epicatechin-3'- <i>O</i> - $\alpha$ -D-maltopyranoside (-)-epicatechin-3'- <i>O</i> - $\alpha$ -D-maltotriopyranoside (-)-epicatechin-4'- <i>O</i> - $\alpha$ -D-glucopyranoside	* 44-fold	Aramsangtienchai et al., 2011
<b>(-)-epigallocatechin gallate</b> <i>Water solubility: ~3mM; 10 mM</i> 	Cyclomaltodextrin glucanotransferase from <i>Bacillus stearothermophilus</i> (GH13)	Starch, cyclodextrin, maltodextrin	(-)-epigallocatechin-3',7-di- <i>O</i> - $\alpha$ -D-glucopyranoside (-)-epigallocatechin-3',5-di- <i>O</i> - $\alpha$ -D-glucopyranoside (-)-epigallocatechin-3'- <i>O</i> - $\alpha$ -D-glucopyranoside Gallate-3- <i>O</i> -( $\alpha$ )-epigallocatechin-3'- <i>O</i> - $\alpha$ -D-glucopyranoside Gallate-3- <i>O</i> -( $\alpha$ )-epigallocatechin-3',7-di- <i>O</i> - $\alpha$ -D-glucopyranoside (-)-gallocatechin-3'- <i>O</i> - $\alpha$ -D-glucopyranoside	> 10-fold	Ochiai et al., 2010
<b>Naringin</b> <i>Water solubility: ~3 µM</i> 	Cyclodextrin glucanotransferase from <i>alkalophilic Bacillus sp.</i> BL-31 (GH13)	Maltodextrin	Three non-characterized naringin monoglucosides One non-characterized naringin diglucoside	Not available	Go et al., 2007
	Maltogenic amylase from <i>Bacillus stearothermophilus</i> (GH13)	Maltotriose	Naringin-6''- $\alpha$ (1 $\rightarrow$ 6)-D-glucopyranoside Naringin-6''- $\alpha$ (1 $\rightarrow$ 6)-D-isomaltopyranoside * Naringin-6''- $\alpha$ (1 $\rightarrow$ 6)-D-isomaltotriopyranoside	* 250-fold	Lee et al., 1999
	Cyclodextrin glucanotransferase from <i>alkalophilic Bacillus</i> (GH13)	Starch, cyclodextrin,	Naringin-3''- <i>O</i> - $\alpha$ (1 $\rightarrow$ 6)-D-glucopyranoside	1000-fold	Kometani et al., 1994, 1996

<b>Rutin</b> <i>Water solubility: ~200 <math>\mu</math>M</i> 	Cyclodextrin glucanotransferase from <i>Bacillus stearothermophilus</i> , <i>macerans</i> and <i>circulans</i> (GH13)	Maltodextrin	Rutin-4''- <i>O</i> - $\alpha$ (1 $\rightarrow$ 4)-D-glucopyranoside * Rutin-4''- <i>O</i> - $\alpha$ (1 $\rightarrow$ 4)-D-maltopyranoside	* 30000-fold	Suzuki and Suzuki, 1991; Kren and Martínková, 2001
	$\alpha$ -glucosidase derived from animals, plants, and microorganisms (GH13)	Maltose, maltotriose, maltotetraose	Rutin $\alpha$ -glucoside, rutin $\alpha$ -maltoside, rutin $\alpha$ -maltotriose, rutin $\alpha$ -maltotetraose, rutin $\alpha$ -maltopentaose, etc.	Not available	Hijiya et al., 1992
	Cyclomaltodextrine glucanotransferase from <i>Bacillus stearothermophilus</i> and <i>Klebsiella sp.</i> (GH13)	Maltodextrin, starch	Rutin $\alpha$ -glucoside, rutin $\alpha$ -maltoside, rutin $\alpha$ -maltotriose, rutin $\alpha$ -maltotetraose, rutin $\alpha$ -maltopentaose, etc.	Not available	
<b>Puerarin</b> <i>Water solubility: ~10 mM</i> 	Maltogenic amylase from <i>Bacillus stearothermophilus</i> and 4- $\alpha$ -glucanotransferase from <i>Thermus scotoductus</i> (GH13)	Maltodextrin and amylose	Puerarin- $\alpha$ -(1 $\rightarrow$ 6)-D-glucopyranoside * Puerarin- $\alpha$ -(1 $\rightarrow$ 6)-D-maltopyranoside ** Puerarin- $\alpha$ -(1 $\rightarrow$ 3)-D-glucopyranoside Puerarin- $\alpha$ -(1 $\rightarrow$ 6)-D-glucopyranoside-cycloamyloside (with DP=6–14)	* 15-fold ** 167-fold Not available Not available	Choi et al., 2010 a, 2010 b
	$\beta$ -fructosidase from <i>Arthrobacter nicotianae</i> XM6	Sucrose	Puerarin-[(2 $\rightarrow$ 6)- $\beta$ -D-fructofuranoside] <sub>n</sub> (1 $\leq$ n $\leq$ 4)	Not available	Wu et al., 2013
	Maltogenic amylase from <i>Bacillus stearothermophilus</i> (GH13)	Maltotriose, starch, cyclodextrin	Puerarine-7- <i>O</i> - $\alpha$ -D-glucopyranoside * Puerarine-7- <i>O</i> - $\alpha$ -D-isomaltopyranoside **	* 14-fold ** 168-fold	Li et al., 2004 b
	Maltogenic amylase from <i>Microbacterium oxydans</i> (GH13)	Sucrose, maltose, maltotriose	Puerarine-7- <i>O</i> - $\alpha$ -D-glucopyranoside * Puerarine-7- <i>O</i> - $\alpha$ -D-isomaltopyranoside **	* 18-fold ** 100-fold	Jiang et al., 2008
	Maltogenic amylase from archaeon <i>Thermofilum pendens</i> (GH13)	Cyclodextrin	Puerarin-[(1 $\rightarrow$ 3)- $\beta$ -D-glucopyranoside] <sub>n</sub> (1 $\leq$ n $\leq$ 8)	Not available	Li et al., 2011
<b>Hesperidin / Neohesperidin</b> <i>Water solubility: ~32 <math>\mu</math>M</i> <i>Water solubility: ~820 <math>\mu</math>M</i> 	Maltogenic amylase from <i>Bacillus stearothermophilus</i> (GH13)	Maltotriose	Néohespéridine-6''- $\alpha$ -(1 $\rightarrow$ 6)-D-maltopyranoside	700-fold	Cho et al., 2000
	Cyclodextrin glucanotransferase from <i>Alkalophilic Bacillus</i> (GH13)	Starch, cyclodextrin	Neohesperidin-3''- $\alpha$ (1 $\rightarrow$ 3)-D-glucopyranoside	1500-fold	Kometani et al., 1996
	$\alpha$ -glucosidase derived from animals, plants, and microorganisms	Maltose, maltotriose, maltotetraose	Neohesperidin $\alpha$ -glucoside, neohesperidin $\alpha$ -maltoside, neohesperidin $\alpha$ -maltotriose, neohesperidin $\alpha$ -maltotetraose, neohesperidin $\alpha$ -maltopentaose, etc.	Not available	Hijiya and Miyake, 1990
	Cyclomaltodextrine glucanotransferase from <i>Bacillus stearothermophilus</i> and <i>Klebsiella sp.</i> (GH13)	Maltodextrin, starch	Neohesperidin $\alpha$ -glucoside, neohesperidin $\alpha$ -maltoside, neohesperidin $\alpha$ -maltotriose, neohesperidin $\alpha$ -maltotetraose, neohesperidin $\alpha$ -maltopentaose, etc.	Not available	

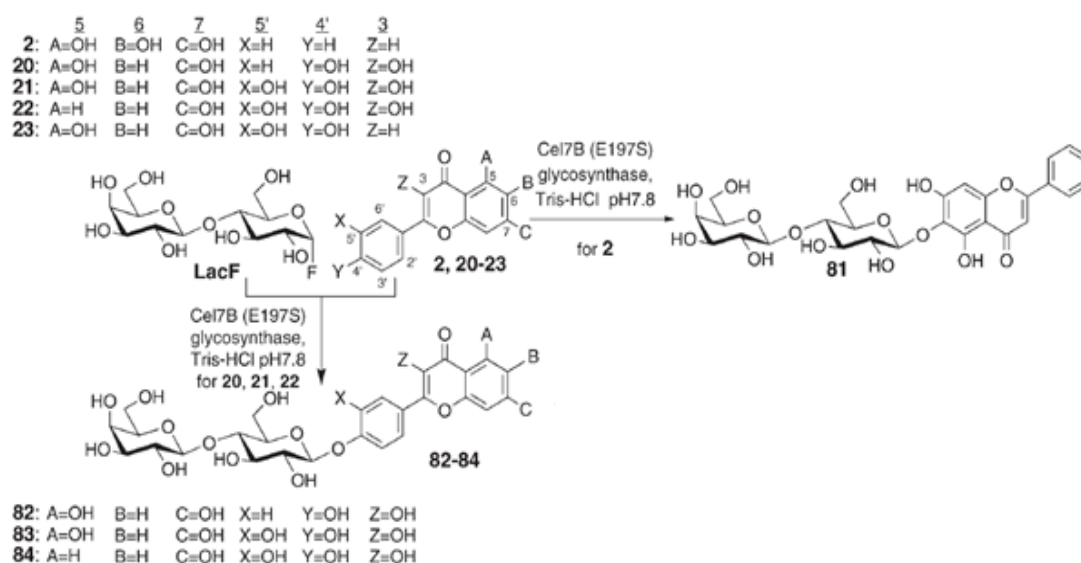
<b>Daidzin</b> <i>Water solubility: ~1.2 <math>\mu</math>M</i> 	$\beta$ -xylosidase from <i>Aspergillus</i> sp. (GH3)	Xylobiose	Daidzein-7- <i>O</i> -[6- <i>O</i> -(4- <i>O</i> -( $\beta$ -D-xylopyranosyl))- $\beta$ -D-xylopyranosyl]- $\beta$ -D-glucopyranoside	25-fold (850-fold compared to aglycon daidzein)	Shimoda et al., 2011
	Maltosyltransferase from <i>Thermotoga maritima</i> (GH13)		Daidzein-7- <i>O</i> - $\alpha$ -D-triglucoside * Daidzein-7- <i>O</i> - $\alpha$ -D-pentaglucoside Daidzein-7- <i>O</i> - $\alpha$ -D-heptaglucoside Daidzein-7- <i>O</i> - $\alpha$ -D-nonaglucoside	* 7500-fold	Li et al., 2004 a
<b>Quercetin</b> <i>Water solubility: ~7 <math>\mu</math>M</i> 	$\alpha$ -glucosidase, cyclodextrine glucanotransferase (GH13)	Glucose, maltose, amylose, starch, amylopectin, cyclodextrin	Quercetin-3- <i>O</i> -glucopyranosyl-[ $\alpha$ (1 $\rightarrow$ 4)-glucopyranosyl] <sub>n</sub> ; 1 $\leq$ n $\leq$ 30	Not available	Ono et al., 2009; Moriwaki et al., 2010
	Cellulase from <i>P. decumbens</i> (GH 5)	Soluble starch, maltose and dextrin	Quercetin monoglucoside	1.3-fold in water/ethanol (50/50)	Chen et al., 2011
<b>Kaempferol</b> 	Cellulase from <i>P. decumbens</i> (GH 5)	Starch, maltose, dextrin	Kaempferol monoglucoside	1.3-fold in water/ethanol (50/50)	
<b>Isorhamnetin</b> 	Cellulase from <i>P. decumbens</i> (GH 5)	Starch, maltose, dextrin	Isorhamnetin monoglucoside	1.3-fold in water/ethanol (50/50)	

**Table 6. Overview of flavonoids glycosylations catalyzed by Glycoside Hydrolases (except glucansucrases)**

As seen in Table 6, a vast panel of flavonoid glycosides showing improved water solubility can be synthesized with GHs, **mostly belonging to GH family 13**, from cheap and abundant sugar donor. Nevertheless, most of these enzymes also show a reverse **hydrolytic activity that may diminish the production yields**. To reduce hydrolytic activity, glycosidase variants named **glycosynthases were engineered** (Mackenzie et al., 1998). These enzymes are mutated on the nucleophile so that they can catalyze the formation of the glycosidic linkages starting from a donor **substrate harboring a good leaving group, usually a fluor atom**. They are **incapable of hydrolyzing the product**. In retaining enzymes, the nucleophile is usually replaced by an alanine, serine or glycine residues whereas a cysteine mutation of the general base catalyst is preferred in inverting ones (Figure 23) (Park, 2008).

Yang et al. (2007) have illustrated that the glycosynthase corresponding to the nucleophile mutant E197S of *Humicola insolens* Cel7B cellulase can transfer glycosyl fluorides onto flavonoids. The wild-type Cel7B is a retaining endoglucanase that hydrolyzes the  $\beta$ -(1 $\rightarrow$ 4) glucosidic bonds of cellulose. Using a MS-based method, the authors have screened 80 different acceptors and more than 20 glycosyl donors. The Cel7B–E197S mutant **glycosylated a large panel of flavonoids using  $\alpha$ -lactosyl-fluoride (LacF) as donor** (Figure 24). Then, it was assessed through preparative-scale syntheses of bacalein (2), kaempferol (20), quercetin (21) and fisetin (22) glycosides with 100  $\mu$ M of flavonoid acceptor (Figure 24). Production yield of flavonol-4'-O- $\beta$ -D-lactoside (82-84) and bacalein-6-O- $\beta$ -D-lactoside (81) were ranging from 72% to 95% and the catalytic efficiency was shown to be comparable with those of the natural glycosyltransferases.

In a recent report, two thermostable  $\beta$ -glucosidases from *T. neapolitana*, TnBgl1A and TnBgl3B, plus a cellulase from *R. marinus* were engineered to be turned into glycosynthases (Pozzo, 2012). Quercetin-3-O-glucoside was used as acceptor molecule with fluorinated sugar as donor, in reactions at high temperature (70 °C). The main product synthesized was characterized as quercetin-3,4'-O-diglucoside.



**Figure 24. Biocatalysis by Cel7B–E197S glycosynthase:** stereo and regioselective transfer of lactosyl from lactosyl fluoride (LacF) to flavonoids (adapted from Yang et al., 2007).

In conclusion, **GHs** appears interesting candidates for flavonoid glycosylation, to produce a panel of new glycosides exhibiting in many cases improved water solubility compared to their aglycons. Moreover, GHs are able to glycosylate flavonoids using cheapest and more abundant glycosyl donors (maltodextrin, starch, amylase, etc.) than the NDP-sugar for GTs. But, together with the glycosylation of flavonoids, most of them also exhibit a hydrolytic activity. It could be overcome by enzyme engineering, as reported for glycosynthases, but these GHs variants requires expensive fluorinated sugar as glycosyl donors.

**Glucansucrases**, a particular class of GH enzymes, has been deliberately left untackled here, as they constitute the focus of the next part of this bibliographic study. These particular GHs belong to the GH-H clan (GH families 70, 13 and 77), and were shown to use sucrose as a preferred D-glucopyranosyl donor to naturally synthesize  $\alpha$ -D-glucans of high molecular mass. At the difference of other GHs, they exhibit a highly predominant transglucosylase activity. Thus they are also known as glucosyltransferase and present a high potential for the synthesis of new flavonoid  $\alpha$ -D-glucoside structures, using the cheap and readily available sucrose as glucosyl donor.



### III. Glucansucrases (GS)

Glucansucrases (GSs) are **bacterial  $\alpha$ -transglucosylases** naturally involved in the **biosynthesis of  $\alpha$ -glucan polymers and glucooligosaccharides**. When suitable exogenous hydroxylated acceptors -osidic or not- are added to the reaction medium, these enzymes catalyze the **synthesis of a wide variety of oligosaccharides or glucoconjugates**. In contrast to Leloir glycosyltransferases that require rare and expensive activated sugars, glucansucrases only use the energy of sucrose osidic linkage rupture ( $\sim 27.6 \text{ kJ.mol}^{-1}$ ) to synthesize their glucosidic products. As **sucrose is a cheap and readily available resource**, the use of glucansucrases to catalyze the synthesis of new active compounds, such as flavonoid glucosides, has recently gained in interest (André et al., 2010).

#### III.1. Few generalities

##### III.1.1. Origins and Classification

Most of the glucansucrases (57 characterized enzymes to date) are produced by **Gram+ lactic acid bacteria** and were isolated from *Lactobacillus*, *Leuconostoc*, *Streptococcus*, and *Weissella* genera. However, five glucansucrases were also isolated from the **Gram- bacteria** belonging to the *Deinococcus*, *Neisseria*, *Alteromonas* and *Arthrobacter* genera (André et al., 2010; van Hijum et al., 2006; Monchois, 1999). Paralogous genes are often found in a same genome, and GSs are not an exception to this rule. To illustrate, 3 glucansucrases (DSR-E, DSR-A and DSR-B) were isolated from the unique *L. citreum* NRRL B-1299 strain (Bozonnet et al., 2002; Monchois et al., 1996, 1997), for instance.

With the advances in next generation sequencing, the number of genes encoding putative glucansucrases has rapidly grown to reach about **250 sequences reported in CAZy databases**. To date, all the putative sequences are issued from bacterial genomic sequences. According to the Enzyme Commission (E.C.), glucansucrases are classified in the **E.C. 2.4.1.x.** category. They are transferases (class 2) which transfer glycosyl units (sub-class 4), and more precisely hexosyl units which are glucosyl units in the case of glucansucrases (sub-sub-class 1). The last number of the E.C. classification refers the polymer structure that they produced. Indeed, glucansucrases are often named according to the linkage specificity of their polymerization activity. As seen in Figure 25, a **wide variety of  $\alpha$ -glucoosidic bonds such as  $\alpha$ -(1 $\rightarrow$ 2),  $\alpha$ -(1 $\rightarrow$ 3),  $\alpha$ -(1 $\rightarrow$ 4) and  $\alpha$ -(1 $\rightarrow$ 6) can be formed** (Leemhuis et al., 2012).

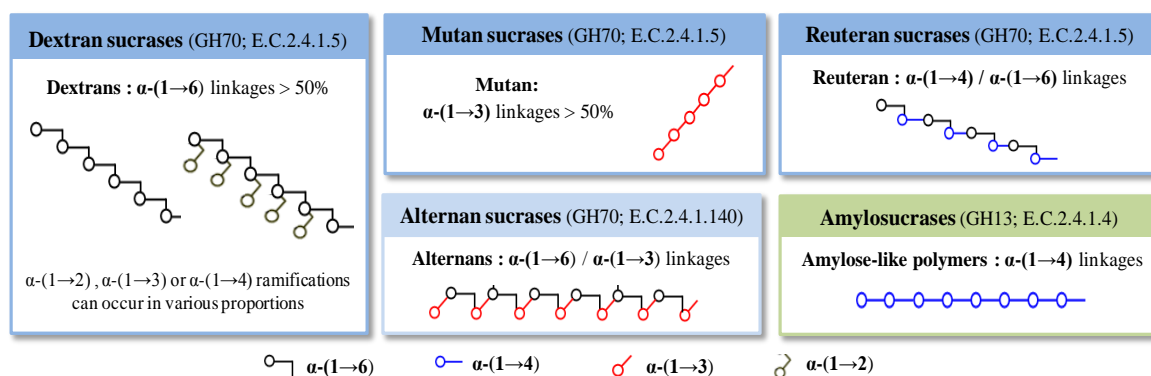


Figure 25. GS categorization according to the structure of the polymer produced (Leemhuis et al., 2013a)

However, there are severe limitations to the utility of this classification system. Indeed, it requires the full characterization of the enzyme donor, acceptor(s) and product(s) before an Enzyme Commission (EC) number can be assigned, whereas only about 25% of the putative GSs identified are characterized. Moreover, EC numbers do not reflect the structural features and evolutionary relationships of these enzymes.

Thus, the alternative CAZy classification, based on amino acid sequence similarities was proposed. According to the CAZy classification, **glucansucrases are classified in two Glycoside-Hydrolases families: GH13 and GH70** (<http://www.cazy.org>; Lombard et al., 2014). These two families, with the GH77 family (amylomaltases and 4- $\alpha$ -glucanotransferases), form the **GH-H clan** (Cantarel et al., 2009; Davies et al., 1997).

### III.1.1.a. GH family 13

The glucansucrases produced by the Gram<sup>-</sup> bacteria *Neisseria*, *Deinococcus* and *Alteromonas* sp. are found in the Glycoside-Hydrolase family 13, which contains more than 15,000 enzymes, and is the **largest GH family** (Lombard et al., 2014). It is often referred to as the superfamily of  $\alpha$ -amylases, which brings together most of the enzymes implicated in starch biotransformation, including hydrolases (E.C.3.x), transferases (E.C.2.x) and isomerases (E.C.5.x). In addition, some GH13 enzymes use sucrose as substrate, such as sucrose-phosphorylases, sucrose-isomerases, sucrose-hydrolases and amylosucrases. The GH family 13 was subdivided in 35 sub-families GH13.1 to GH13.35 (CAZy) that account for these various substrate and reaction (Stam et al., 2006).

To date, the glucansucrases found in GH family 13 are transglucosylases, all named amylosucrases, as they **naturally synthesize, from sucrose, amylose-like polymers by polymerization of the D-glucosyl units exclusively via the formation of  $\alpha$ -(1 $\rightarrow$ 4) linkages**. They represent a very small part of the GH family 13, as only **8 amylosucrases** were reported (3 putative enzymes, and 5 characterized). Amylosucrases (E.C.2.4.1.4)

belong to **GH13.4 subfamily**, which also contain the sucrose-hydrolases (E.C.3.2.1.x). Five amylosucrases have been characterized from *N. polysaccharea* (*NpAS*) (Potocki De Montalk et al., 2000), *D. radiodurans* (*DrAS*) (Pizzut-Serin et al., 2005), *D. geothermalis* (*DgAS*) (Emond et al., 2008a), *A. macleodii* (*AmAS*) (Ha et al., 2009), *A. chlorophenolicus* (*AcAS*) (Seo, 2012). Very recently, a new amylosucrase has been identified in the marine cyanobacterium *Synechococcus* sp. PCC 7002 and characterized, but is not yet incorporated in the CAZy database (Perez-Cenci and Salerno, 2014). These amylosucrases are relatively large enzymes as they exhibit a **molecular weight of about 70 kDa**, but they remain smaller than the GH70 enzymes.

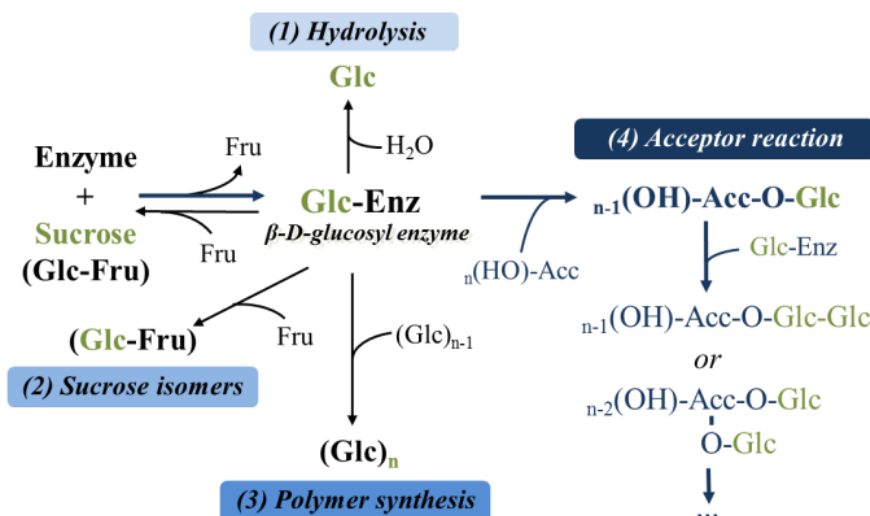
### III.1.1.b. GH family 70

The GH70 family groups together **enzymes produced by Gram + lactic acid bacteria**. They are extracellular transglucosylases with a molecular mass ranging from **120 up to 200 kDa** with few of them even larger. This corresponds to an average primary structure of 1,500 amino acids (Leemhuis et al., 2013a; Remaud-Simeon et al., 2000). This family currently comprises only **231 enzymes** showing either glucansucrases,  $\alpha$ -(1→2) branching sucrase activity, or 4,6- $\alpha$ -glucanotransferases (Mai 2014). Only 57 members have been characterized up to now. Most of them use **sucrose to synthesize  $\alpha$ -D-glucans of high molar mass (>10<sup>6</sup> Da)**. GH70 glucansucrases may show different osidic linkage specificities, as they can produce  **$\alpha$ -D-glucans containing  $\alpha$ -(1→2),  $\alpha$ -(1→3),  $\alpha$ -(1→4) or  $\alpha$ -(1→6)** (Leemhuis et al., 2012). Accordingly, they are often referred to the type of polymer they naturally synthesize, as dextransucrases, mutansucrases, alternansucrases or reuteransucrases (Figure 25). Finally in this family, 4,6- $\alpha$ -glucanotransferases were recently discovered (Kralj et al., 2011). They are related to glucansucrases but essentially catalyze transglucosylation reaction from malto-oligosaccharide donor through  $\alpha$ -(1→6) linkage formation (Leemhuis et al., 2013b).

### III.1.2. Main reactions and product applications

From sucrose only, glucansucrases catalyze different reactions (Figure 26). Sucrose is first cleaved by the enzyme through the formation of a  $\beta$ -D-glucosyl enzyme covalent intermediate with a concomitant release of fructose. Then, the **D-glucosyl unit can be transferred onto a broad variety of acceptors**, through the formation of an  **$\alpha$ -glucosidic linkage**. The main products synthesized are soluble oligosaccharides and  $\alpha$ -glucan polymers, resulting from the stepwise transfers of the glucosyl units onto growing  $\alpha$ -glucan chains (Figure 26). The active site architecture determines how the catalytic pocket

accommodates the growing  $\alpha$ -glucan chains and, thus, the type of glucosidic linkage synthesized.



**Figure 26.** Main reactions catalyzed by sucrose-utilizing glucansucrases. *Enz*: enzyme, *Fru*: fructose, *Glc*: glucose, *Acc*: acceptor.

Usually, one GS does not catalyze the formation of more than two types of glucosidic linkages, in specific proportions determined by the enzyme specificity (Leemhuis et al., 2012). **Highly branched glucans** are also produced by some enzymes. For example, **DSR-E** glucansucrase catalyze the synthesis of a  **$\alpha$ -(1 $\rightarrow$ 6) dextran branched via  $\alpha$ -(1 $\rightarrow$ 2) glucosidic linkages** (Bozonnet et al., 2002). So, depending on the enzyme specificity, a wide range of  $\alpha$ -glucans can be produced, varying in terms of size, structure, degree of branches and spatial arrangements (Moulis et al., 2006). To illustrate, **amylosucrase** from *N. polysaccharea* (*NpAS*) catalyzes the formation of  **$\alpha$ -(1 $\rightarrow$ 4) amylose-like polymers** with molecular weights about  **$10^4$  g/mol** (~60 units) and **insoluble in water** (Potocki-Veronese et al., 2005), whereas dextran produced by the **dextransucrase DSR-S** from *L. mesenteroides* B512F mainly contains an  **$\alpha$ -(1 $\rightarrow$ 6) linked glucosyl units** and is far much larger with a polymerization degree reaching up to  **$10^8$  g/mol** (~61,000 units) and is **soluble in water** (Irague et al., 2012).

Polymerization is often accompanied **by side-product formation** such as glucose resulting from sucrose hydrolysis, and sucrose isomers (turanose, trehalulose, isomaltulose, or leucrose) which result from the glucosyl transfer onto fructose released during the reaction (Figure 26).

In addition, GSs can also transfer the glucosyl unit to a wide variety of hydroxyl-group containing molecules in a so called "**acceptor reaction**" (Figure 26). Indeed, when a hydroxylated acceptor, such as saccharide, alcohol, polyol, or polyphenol is added to the

reaction mixture, the enzymatic activity of glucansucrases can be re-directed from its natural reaction toward the glucosylation of these exogenous acceptors (Daudé et al., 2014; Monsan et al., 2010a). **The efficiency of such reactions is also related to the enzyme specificity.**

In the seventies, glucansucrases from pathogenic *Streptococcus mutans* species found on tooth surface were largely studied to elucidate their role in the formation of dental caries. They have also attracted most attention due to the **potential industrial applications of the polymers or oligosaccharides** they produce. These products can be used as viscous and water-binding additives for food and non-food applications (Badel et al., 2011), blood plasma substitutes (Groenwall and Ingelman, 1948); prebiotics (Kim et al., 2014; Monsan et al., 2000, 2010b; Sarbini et al., 2011, 2013), cosmetic additives (Dijkhuizen et al., 2010), anti-diabetic treatment (André et al., 2010), biomaterials (Arthur et al., 2014; Irague et al., 2012; Latham et al., 2013; Sousa et al., 2012), material for dental health (Mattos-Graner et al., 2000; Yamashita et al., 1993), etc. As glucansucrases are also able to transfer the glucosyl units from sucrose onto a large panel of exogenous hydroxylated acceptor molecules, they **also stand as very attractive biocatalysts for the production of novel glucoderivatives** (Monsan et al., 2010b).

### III.1.3. *NpAs* and GBD-CD2, the two GS templates of this study

In this thesis, the focus was placed on two flagship glucansucrases from GH families 13 and 70, namely the **amylosucrase from *N. polysaccharea* (*NpAS*)** and the  **$\alpha$ -(1→2) branching sucrase (GBD-CD2), a truncated form of DSR-E dextransucrase** produced by *L. citreum* NRRL B-1299. Indeed, both enzymes have been well studied at the laboratory (LISBP), notably due to their remarkable and distinct specificities and their potential ability to perform unusual acceptor reactions.

#### III.1.3.a. Amylosucrase from *N. polysaccharea* (*NpAS*)

Amylosucrase activity was first identified and characterized from supernatant culture of *Neisseria perflava* (Hehre, 1949). It was only in 1995 that the gene coding for this amylosucrase was first isolated from *Neisseria polysaccharea* (Remaud-Simeon et al., 1995) and further cloned in *E. coli* (Büttcher et al., 1997; Potocki De Montalk et al., 1999). Sequence homology of *NpAS* with  $\alpha$ -amylase allowed its classification in **GH 13 family**. The *NpAS* encoding **gene was fused with that encoding Glutathion S Tranférase (GST)** to allow affinity purification to homogeneity (Potocki De Montalk et al., 1999). The recombinant *NpAS* was shown to catalyze, the formation of an amylose-like polymer from sucrose and has been rapidly considered as a very promising biocatalyst for the development of valuable

products (Potocki-Veronese et al., 2005). Amylose, often known as one of the two component of starch, is indeed largely used as biomaterial or in food industry.

*NpAS* was also shown to catalyze very efficiently the elongation of glycogen branches to produce new original glycogen structures (Potocki de Montalk et al., 2000; Potocki De Montalk et al., 1999). Glycogen exerted an activated effect that decreased at high sucrose concentrations. It was attributed to the binding of glycogen external chains at the enzyme surface, as confirmed by Albenne et al. (2007). *NpAs* was shown to display a **narrow specificity toward glucosyl donor** as its activity appeared highly affected toward other donors than sucrose. To get better insight in the donor and acceptor natural promiscuity of *NpAS*, eleven different glycosyl donors and twenty hydroxylated acceptors including monosaccharides and polyols but no flavonoids were tested for glucosylation (Daudé et al., 2013a). *NpAS* showed a **high promiscuity towards the tested acceptors**, as 19 of the 20 tested were recognized by the enzyme.

### III.1.3.b. The $\alpha$ -(1→2) branching sucrose (GBD-CD2)

In 2005, the full length  $\alpha$ -(1→2) branching sucrose, GBD-CD2 (192 kDa), was obtained by **truncation of the very large DSR-E enzyme** (313 kDa), isolated from *L. citreum* NRRL B-1299. DSR-E possesses two catalytic domains, CD1 and CD2, separated by a glucan-binding domain (GBD) (Bozonnet et al., 2002; Fabre et al., 2005) as shown on Figure 27. The recombinant protein catalyses the synthesis of both  $\alpha$ -(1→6) and  $\alpha$ -(1→2) glucosidic linkages, using sucrose as the donor. It is classified in **Glycoside Hydrolase family 70**, where it is the unique enzyme to harbor two catalytic domains.



**Figure 27. Schematic structure of DSR-E from *L. citreum* B1299 and GBD-CD2 truncated variant.** VR, Variable Region; CD, Catalytic Domain; GBD, Glucan Binding Domain.

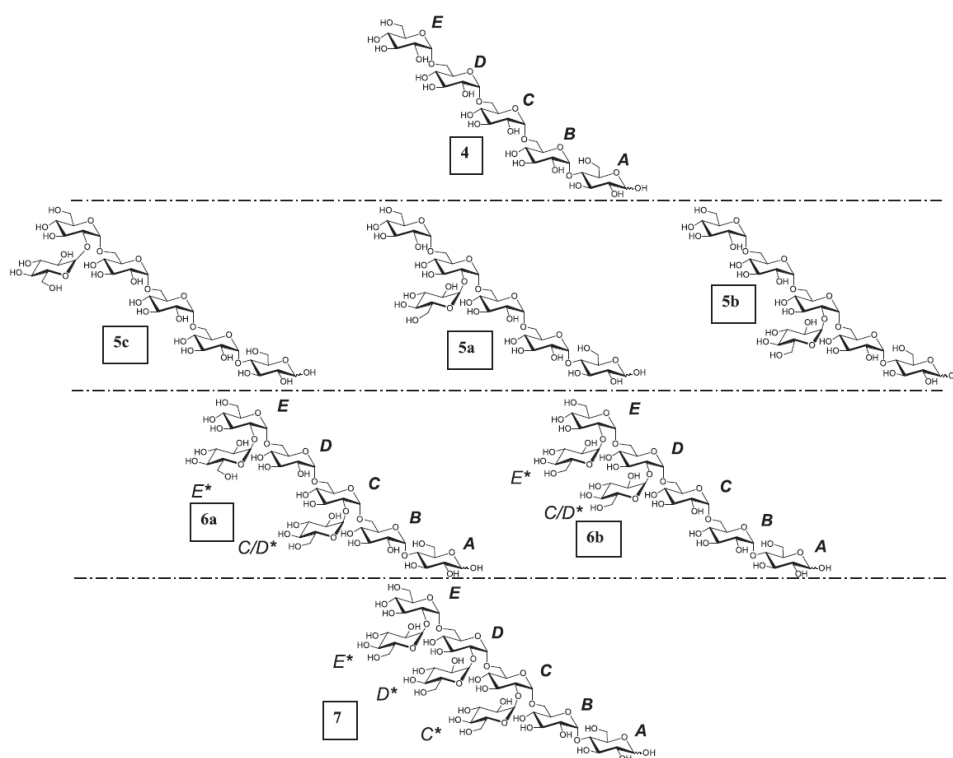
To better characterize the role of each domain, two truncated mutants, CD1-GBD and GBD-CD2 (192 kDa) were generated. CD1-GBD was shown to be responsible of the  $\alpha$ -(1→6) polymerization reaction that generates dextran chains from sucrose.

In contrast, **GBD-CD2 was shown to act as a regiospecific  $\alpha$ -(1→2) transglucosylase in the presence of sucrose, and linear  $\alpha$ -(1→6) glucans (dextrans) acceptors. However, from sucrose only, it is unable to synthesize glucans, and only**



**acts as a sucrose hydrolase** (Fabre et al., 2005). This makes this recombinant enzyme **unique** compared to all other reported glucansucrases. Moreover, the  **$\alpha$ -(1 $\rightarrow$ 2) branched glucooligosaccharides are rarely present in nature**, and were shown to exhibit **prebiotic properties**. A mix of glucooligosaccharides containing isomaltooligosaccharides and  $\alpha$ -(1 $\rightarrow$ 2) branched ones has been commercialized by Solabia since about twenty years as prebiotics for dermo-cosmetics. Pure and highly  $\alpha$ -(1 $\rightarrow$ 2) branched isomaltooligosaccharides are thought to be more effective (Sarbin 2013).

An important point is that the **degree of  $\alpha$ -(1 $\rightarrow$ 2) linkages** incorporated in dextrans was shown to be **controlled by the initial molar [Sucrose]/[Dextran] ratio** used for the acceptor reaction (Brison et al., 2009). By varying this ratio from 0.92 to 4.74, the degree of  $\alpha$ -(1 $\rightarrow$ 2) branching varies from 10% to 40%, respectively. This value corresponds to the highest degree of  $\alpha$ -(1 $\rightarrow$ 2) linkages reported for branched dextrans and led to patented applications (Einerhand et al., 2010). The  $\alpha$ -(1 $\rightarrow$ 2) transglucosylase activity was shown to be independent of the dextran polymerization degree (DP). Investigation on the kinetics of the branching event further revealed that the branching process of pure glucooligosaccharides is stochastic. They also confirmed that the **dextran can bind the active site in various ways** and that a wide variety of branched oligosaccharides can be formed (Brison et al., 2013) (Figure 28).



**Figure 28. Structures of the  $\alpha$ -(1 $\rightarrow$ 2) GOS 5a, 5b, 5c, 6a, 6b and 7 produced by GBD-CD2 using sucrose as donor and pentasaccharide 4 as acceptor (Brison et al., 2013).**

**No other exogenous acceptor reactions were reported to date.** However, several features make up this second transglycosylase an interesting candidate for this thesis: i) sole GH70 enzyme **specific for acceptor reaction** compared to the others, which naturally synthesize concomitant polymers; ii) **remarkable catalytic efficiencies**; iii) **3D structure recently solved**, allowing enzyme semi-rational design

## III.2. Structure/function relationships studies

### III.2.1. Catalytic mechanism

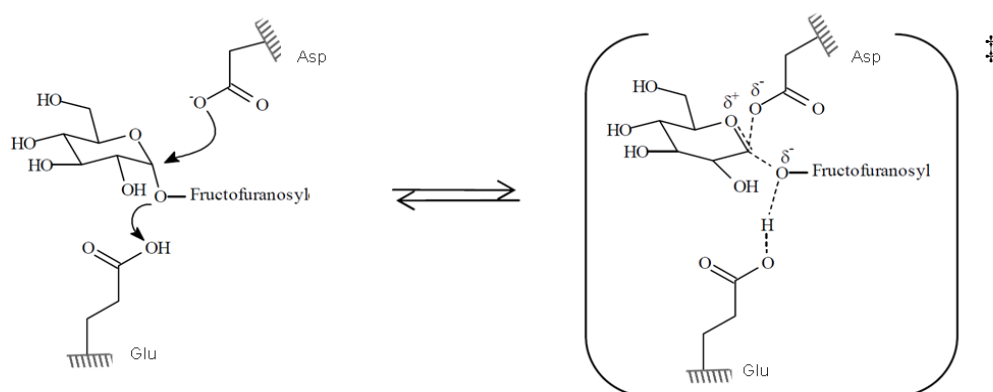
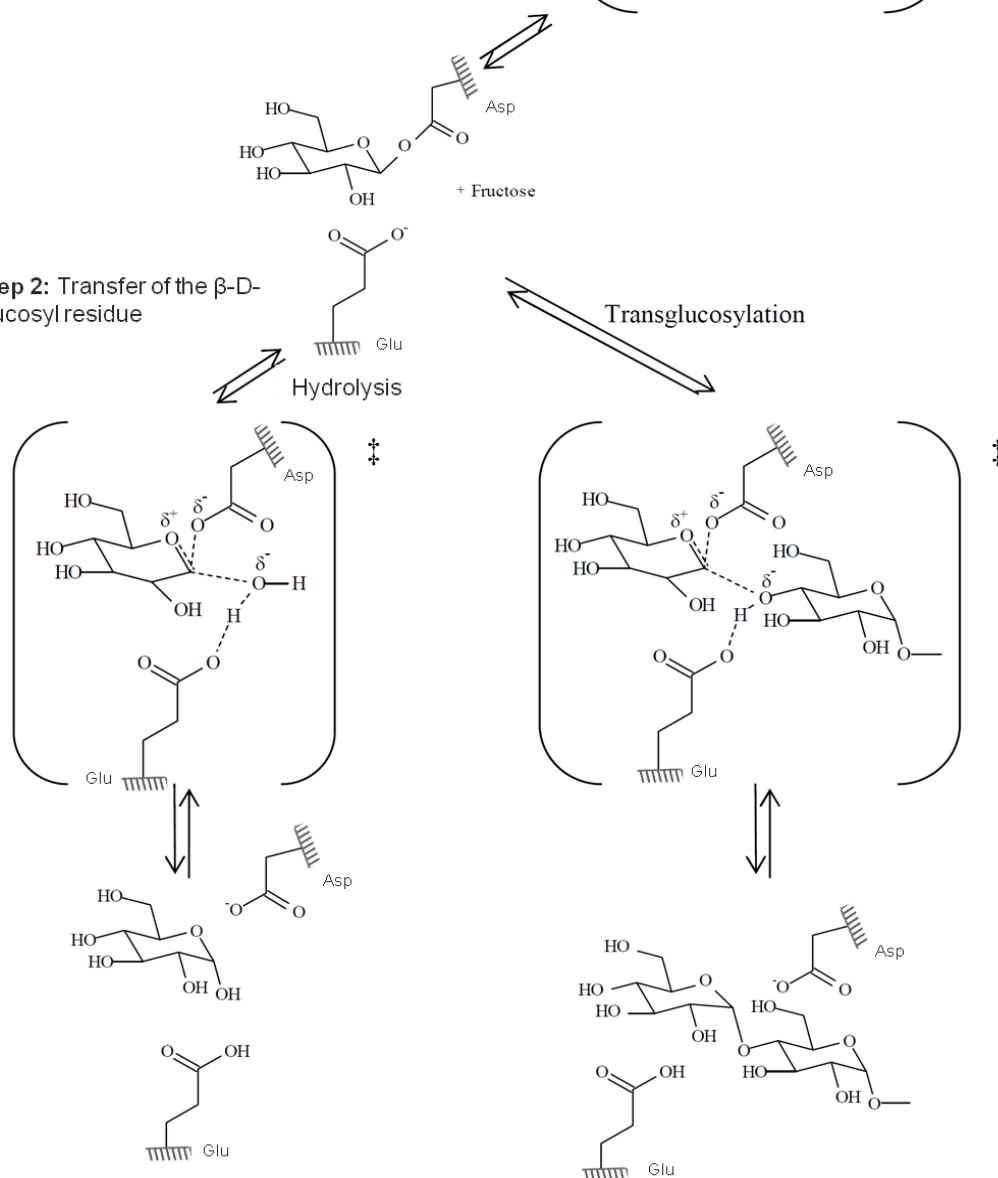
**GH13 and GH70 enzymes** belong to the same GH-H clan and possess the **same two step  $\alpha$ -retaining-type mechanism** (Figure 29). The reaction starts with the binding and the cleavage of the sucrose substrate. In a first step (glucosylation step), the catalytic nucleophile residue (aspartate) attacks the anomeric carbon of the sucrose glucosyl ring. Simultaneously, the general acid/base catalyst (glutamic acid) protonates the glycosidic oxygen of the sucrose donor. This leads to the **formation of a  $\beta$ -D-glucosyl enzyme covalent intermediate** through the release of fructose (Jensen et al., 2004).

In a second step (de-glucosylation step), the glucosyl moiety is transferred to an acceptor activated by the deprotonated glutamic acid lateral chain to attack the covalent bond formed between the anomeric C1 of the glucosyl ring and the aspartate. This reaction provokes a second inversion of the anomeric configuration of the D-glucosyl. A third aspartate residue (not shown on Figure 29) is assumed to be involved in the mechanism by ensuring the correct positioning of the glucosidic bond via the distortion of the glucosyl ring (Leemhuis et al., 2013a).

### III.2.2. Polymerization process

GH13 and GH70 glucansucrases share also a close polymerization mechanism. It consists in a **glucosyl unit transfer at the non-reducing end of the glucosidic acceptor**, as first proposed by Mooser et al. (1991, 1992), followed by Albenne et al. (2004) and Moulis et al. (2006). It was confirmed by 3D-structures of glucansucrases in complex with acceptors (Ito et al., 2011; Skov et al., 2002; Vujičić-Žagar et al., 2010).



**Step 1: Formation of the  $\beta$ -D-glucosyl-enzyme intermediate****Step 2: Transfer of the  $\beta$ -D-glucosyl residue**

**Figure 29. The two-step  $\alpha$ -retaining-type mechanism of glucansucrases** (adapted from Skov et al., 2001 and Withers, 2001). Brackets and ‡ point out the transition states. Nucleophile (Asp) and acid/base (Glu) catalysts are shown.

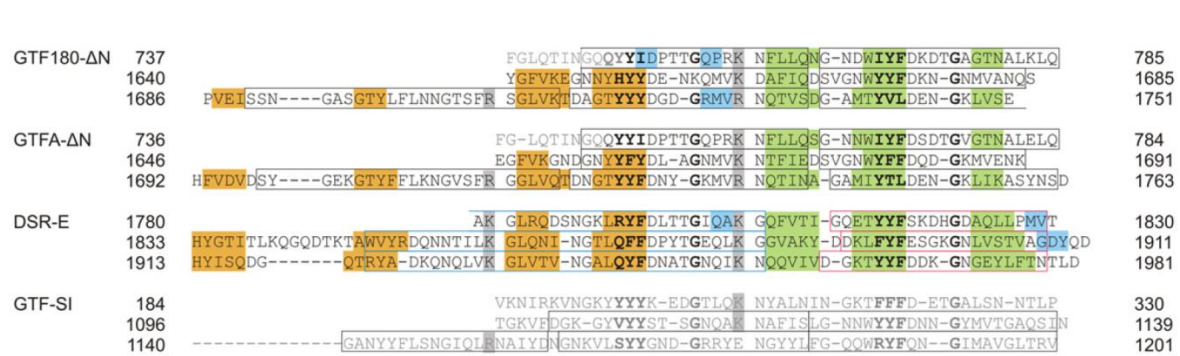
Generally, two modes of polymerization can occur:

- **Non-processive:** A growing glucosidic chain binds at the active site, is elongated by transfer of one glucosyl unit and released. Then, the same chain or another one can randomly accommodate the active site in order to be elongated.
- **Processive:** Only the growing glucosidic chain is elongated and remains bond to the enzyme during successive glucosyl residue transfers.

Concerning GH13 glucansucrases, Albenne et al. (2004) showed that **NpAS naturally catalyze amylose-like polymer formation via a non-processive elongation**. They clearly demonstrated that the insoluble fraction contains polydisperse maltooligosaccharides (MOS). Indeed, the MOS produced are randomly elongated until they reach a critical size and concentration responsible for chain precipitation highly limiting the size of the amylose chains formed. The authors also reported that the **amylose-like chain length can be controlled by the initial substrate conditions**. However, presence of **anchoring sites at the NpAS surface** was suggested by Albenne et al. (2007). These acceptor binding sites were proposed to capture the polymer and direct the branches of glycogen toward the catalytic site for their efficient elongation. The authors proposed that **glycogen elongation might occur via a semi-processive mechanism**.

Concerning GH70 family members, **polymerization from sucrose only** was proposed to **start with a non-processive mechanism** via successive transfers of glucosyl residues, first onto water molecules, and then onto glucose units (Moulis et al., 2006). Short oligosaccharides with degrees of polymerization comprised between 7 and 10 were identified to support this hypothesis. However, as the concentration of these products remained very low along the course of the reaction, they were proposed to strongly interact with the enzyme **after sufficient elongation** (DP 7 to 10) due to the presence of **glucan binding sites** suggested to be involved in a **processive mechanism** enabling very efficient elongation until formation of a high molecular weight  $\alpha$ -glucan chain. These investigations allowed to suggest that a **“semi-processive” mechanism of polymer elongation for this class of enzymes** (Moulis et al., 2006).

From sequence analyses, it was proposed that **these glucan binding sites** are based on a **consensus Y(or F)G-repeat** (Figure 30), which has homology with the cell-wall binding motif found in choline binding proteins, toxins and surface-associated proteins (Giffard and Jacques, 1994).



**Figure 30. Structural/sequence alignment of proposed GBD domains of GTF180-ΔN, GTFA-ΔN and DSR-E ΔN123-GBD-CD2 and GTFI-SI (Leemhuis et al., 2013a). Open boxes correspond to the YG-repeats sequences (black lines). YG motifs are shown in bold. Grey boxes correspond to the basic residues between motifs.**

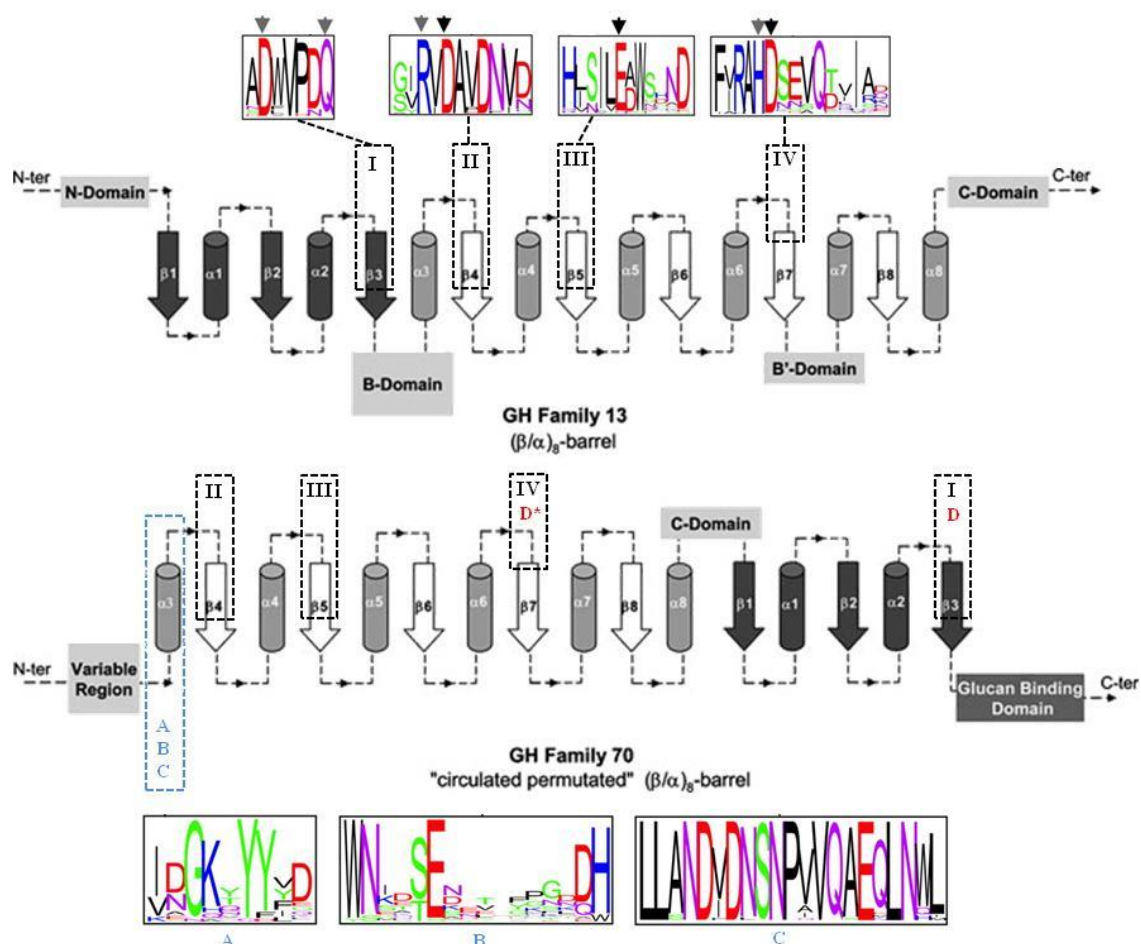
These repeating elements were shown to be **specific of GH70 glucansucrases** (not found in GH13) and located at their N- and/or C-termini sequences. They are thought to constitute a functional domain named **Glucan Binding Domain (GBD)**. The dextransucrase DSR-E from *L. citreum* NRRL B-1299 appeared as an exception to the rule as its GBD is located between its two catalytic domains (CD1 and CD2) (Bozonnet et al., 2002). The role of these GBDs has been investigated by generation of truncated mutants for which the repeating elements were progressively deleted (Moulis et al., 2006). The mutants gradually exhibited a clear non-processive mechanism of polymerization, as the GBD was reduced, confirming its **involvement in the processive mechanism**. Moreover, mutations of conserved aromatic and basic residues in repeating elements of the GBD of *S. downei* Gtfl also confirmed that they were important for dextran binding (Shah et al., 2004). However, their **precise role at the molecular level is still not fully understood**.

Finally, **after total sucrose depletion, glucansucrases from both GH13 and GH70 families disproportionate α-glucans chains** (Albenne et al., 2002a; Binder et al., 1983) by cleaving the glucosyl unit at the non-reducing end of one α-glucan chain and transferring it to the non-reducing end of another chain. This reaction highlights the evolutionary relationship between glucansucrases and the 4,6-α-glucanotransferases also found in GH70 family, which are only specialized in the disproportionation of α-(1→4)-glucan chains to synthesize linear α-(1→6) chains branched to the size-reduced α-(1→4)-glucans. They can use as an acceptor but do not use it as a glucosyl donor (Dobrurowska et al., 2012; Kralj et al., 2011; Leemhuis et al., 2013b).

### III.2.3. Sequences and 3D-structures

Structural and mechanistic analogies between GS from families 13 and 70 were revealed by Devulapalle et al. (2008), MacGregor et al. (1996), Mooser et al. (1991), Ferretti et al. (1987), who suggested that **GH13 and GH70 glucansucrases both adopts a  $(\beta/\alpha)_8$  barrel structure.**

The  $(\beta/\alpha)_8$  barrel consists of eight parallel  $\beta$ -strands forming the inner  $\beta$ -barrel surrounded by eight  $\alpha$ -helices so that the individual  $\beta$ -strands and  $\alpha$ -helices alternate and are interconnected by loops (Figure 31). The structural analyses of the catalytic domains of GH70 GS confirmed that the **catalytic core of the GH70 is related to GH13  $(\beta/\alpha)_8$  barrel but results from a circular permutation.** In GH70,  $\beta 1$ ,  $\alpha 1$ ,  $\beta 2$ ,  $\alpha 2$ ,  $\beta 3$  and  $\alpha 3$  constituents of the  $(\beta/\alpha)_8$  barrel are situated downstream the  $\alpha 3$ ,  $\beta 4$ ,  $\alpha 4$ ,  $\beta 5$ ,  $\alpha 5$ ,  $\beta 6$ ,  $\alpha 6$ ,  $\beta 7$ ,  $\alpha 7$ ,  $\beta 8$  and  $\alpha 8$ , in contrast to the sequential  $\beta 1 \rightarrow \alpha 8$  arrangement of GH13 (Figure 31).



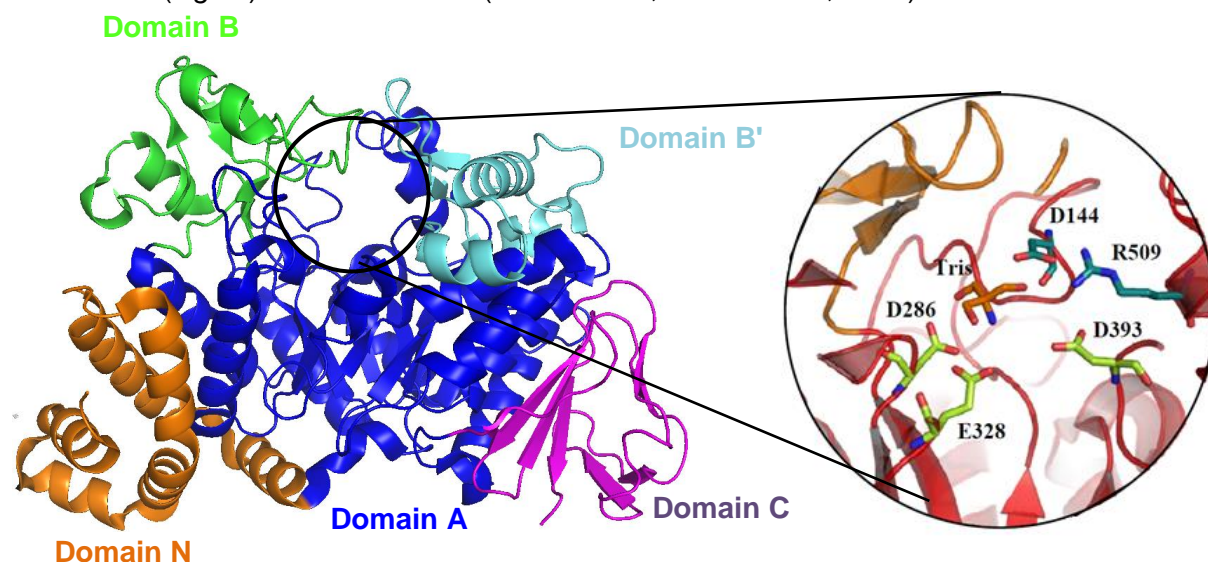
**Figure 31. Topology diagrams of GH13 and GH70 sucrose-utilizing transglucosylases** (adapted from André et al., 2010 and van Hijum et al., 2006). Cylinders represent  $\alpha$ -helices and arrows  $\beta$ -sheets constituting the  $(\beta/\alpha)_8$  barrel. The permuted elements are shown in dark grey. The four conserved regions between GH13 and GH70 are indicated in black (I to IV). The seven residues that are fully conserved in both families are indicated with arrows, and black arrows indicate the three catalytic residues. The three conserved regions between GH70 members only are indicated in blue (A to C).

Sequence alignment revealed that **four regions of the  $(\beta/\alpha)_8$  barrel are highly conserved among the two families** (Figure 31). In these four regions, **seven residues comprising the catalytic triad residues are strictly conserved** (van Hijum et al., 2006; Jespersen et al., 1993; Sarçabal et al., 2000). For **GH70 glucansucrases, three additional conserved regions** among the family have also been identified. They are located in the catalytic domain of the enzymes, upstream the four others (Figure 31). They were shown to contain some **important residues for the enzymes activity and specificity** (van Hijum et al., 2006).

If the structure of the amylosucrase (*NpAS*) from GH13 family has been available since 2001, the first 3D structure of a GH70 member was recently solved in 2010. Nowadays, structures of several glucansucrases of both families GH 13 and 70 are available in *apo* or *holo* forms, and have confirmed the previously described structural features deduced from sequences and mechanisms comparisons.

### III.2.3.a. General structural organization

The amylosucrase from *Neisseria polysaccharea* (*NpAS*) was the first structurally characterized glucansucrase from GH-13 family (PDB: 1G5A; Skov et al., 2001). Very recently, the structure of the thermostable amylosucrase from *Deinococcus geothermalis* (*DgAS*) was also solved (PDB: 3UCQ; Guérin et al., 2012).



**Figure 32. Schematic representation of the *NpAS* structure with labeling and color-coding of the five domains (PDB code: 1G5A). Domain caption: N (orange), A (blue), B' (cyan), B (green) and C (magenta). Black circle defines the catalytic site region, also presented as a zoomed view showing the catalytic triad (green), salt bridge (blue) and TRIS molecule (orange).**

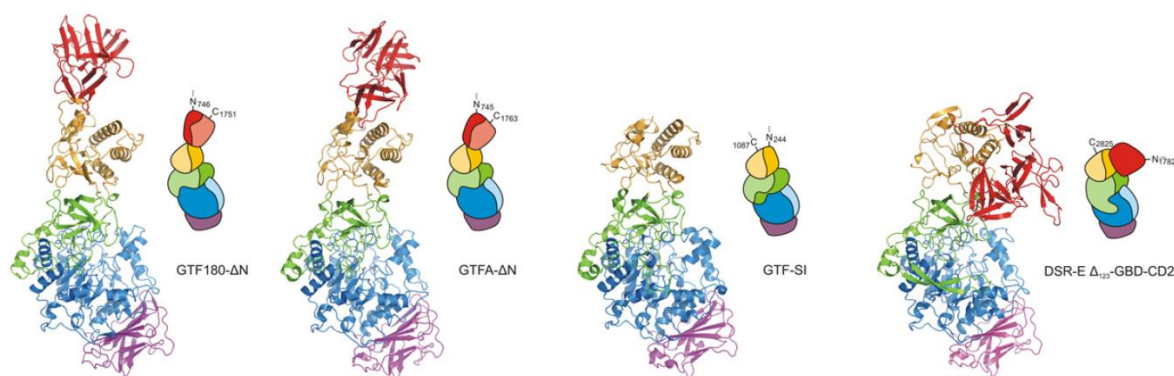
*NpAS* 3-dimensionnal structure revealed that a single polypeptidic chain is fold in **five domains named N, A, B, B' and C** (Figure 32). Three of them (A, B and C) are also encountered in the  $\alpha$ -amylases. Domain A architecture adopts the  $(\beta/\alpha)_8$ -barrel folding, which

includes residues 98 to 184, 261 to 395 and 461 to 550, and contain the amino acids of the enzyme **catalytic triad: D286 (nucleophile), E328 (acid/base catalyst) and D393 (stabilizer)** (Figure 32). Loop 3, composed by residues 185 to 260, is particularly long and form an entire module named domain B. Residues 395 to 460 form the B' domain, which has no equivalent in  $\alpha$ -amylases. The N-terminal domain (residues 1 to 90) is only composed of six  $\alpha$ -helices. Finally, domain C (residues 55 to 628) is a succession of anti-parallel  $\beta$ -strands forming a Greek-key pattern. **No calcium ions** were found in the structure.

***NpAS* (GH13) active site is located at the bottom of a narrow pocket** (Figure 32), supported by the  $(\beta/\alpha)_8$ -barrel (domain A) and covered with domains B and B' in which both the sucrose binding site (SB1) and the maltooligosaccharide binding site (OB1) were mapped (Skov et al., 2002). Two additional maltooligosaccharide binding sites were described at the *NpAS* surface (Skov et al., 2002) and named OB2 and OB3. OB2 is located in domain B' and OB3 in domain C. Albenne et al. (2007) suggested that the OB2 surface site provides an anchoring platform at the enzyme surface to capture the glycogen polymer and direct the branches towards the catalytic pocket for elongation. OB3 was also suspected to facilitate and stabilize the accommodation of the long oligosaccharidic chains in the catalytic pocket of *NpAS*.

**In 2010, the first three-dimensional structure of a GH70 family member** was solved for the GTF180- $\Delta$ N glucansucrase from *L. reuteri* 180 (PDB: 3KLK; Vujičić-Žagar et al., 2010) (Figure 33). At present, three other GH70 structures have been solved including the structures of GTF-SI from *S. mutans* (PDB: 3AIE; Ito et al., 2011),  **$\Delta$ N<sub>123</sub>-GBD-CD2 truncated form of the  $\alpha$ -(1 $\rightarrow$ 2) branching dextranase DSR-E from *L. citreum* NRRL B-1299** (PDB: 3TTQ; Brison et al., 2012) and GTFA- $\Delta$ N from *L. reuteri* 121 (PDB: 4AMC; Pijning et al., 2012). In all cases, the crystals were obtained with truncated forms of the enzymes, exhibiting a reduced Glucan Binding Domain. No entire structures of a complete GH70 glucansucrase could be solved. This is, probably due to their large size and/or to the flexibility of their GBD (Pijning et al., 2014). Access to full structures will definitely be a challenge for the next years.

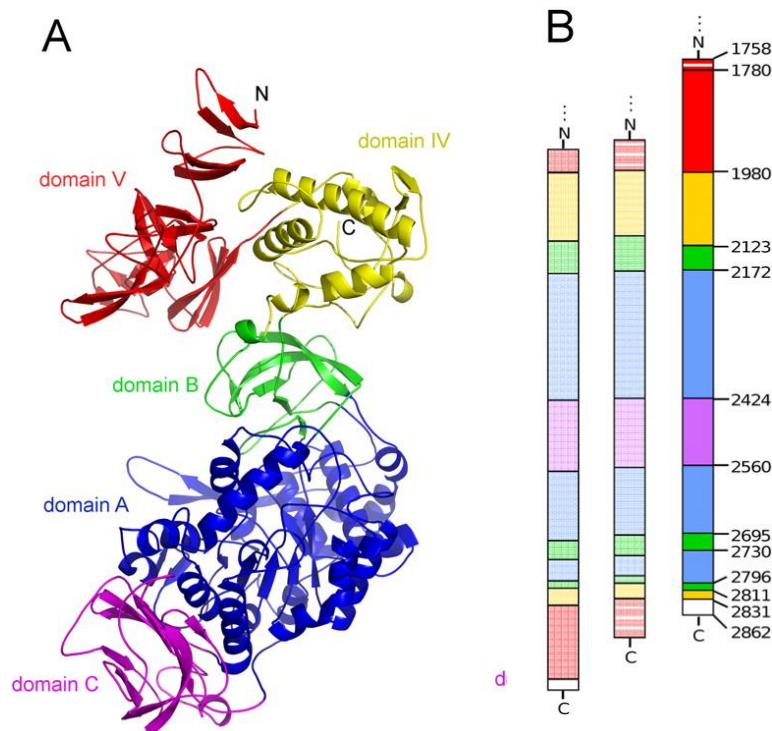




**Figure 33. Three-dimensional structures of glucansucrases** (Leemhuis et al., 2013a). Crystal structures of truncated glucansucrases in cartoon representation and schematic domain arrangement, colored by domain (A: blue, B: green, C: magenta, IV: yellow, V: red). From left to right: GTF180- $\Delta$ N (PDB: 3KLK), GTFA- $\Delta$ N (PDB: 4AMC), GTF-SI (PDB: 3AIE) and  $\Delta$ N<sub>123</sub>-GBD-CD2 (PDB: 3TTQ).

The three-dimensional structures of these four GSs revealed a very **different and original domain organization compared to GH13** glucansucrases (Figure 33). The catalytic core of the GH70 enzymes consists of three domains, which resemble the A, B and C domains also found in GH13 enzymes. But for GH70 glucansucrases, **two extra domains, called IV and V**, are attached to the core domains. The most surprisingly is that the **polypeptide chain successively contributes to domains V, IV, B, A, C, A, B, IV, and V from N- to C-terminus** (Figure 33). Thus, the A, B, IV and V domains are built up from two discontinuous segments of the polypeptide chain, such that the polypeptide adopts a **U-shaped fold**. Only the C domain consists of one contiguous polypeptide.

All the attempts to crystallize the full-length recombinant  $\alpha$ -(1 $\rightarrow$ 2) branching sucrase (GBD-CD2) failed, probably because the N-terminal GBD (849 amino acids with 41 consecutive repeat units rich in aromatic residues) did not favor crystallization. Thus, **truncations of the  $\alpha$ -(1 $\rightarrow$ 2) branching sucrase** enzyme were performed. The **shortest active truncated form,  $\Delta$ N123-GBD-CD2** (123 kDa; 1108 residues), harboring a **GBD reduced by 76%**, was successfully crystallized (Brison et al., 2012). The apo X-ray structure of this enzyme was solved at 1.90 Å resolution (Figure 34; Brison et al., 2012). An additional X-ray structure of the enzyme was also solved at a 3.3 Å resolution in a different space group. These are the **first 3D-structures of an  $\alpha$ -(1 $\rightarrow$ 2) branching sucrase** reported in literature (Brison et al., 2012).

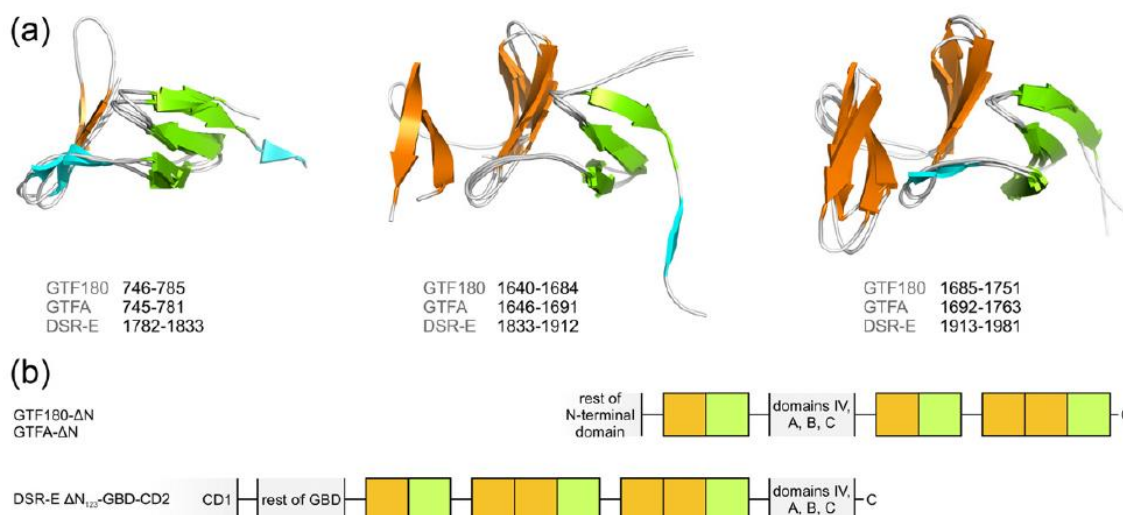


**Figure 34. (A) View of the  $\alpha$ -(1 $\rightarrow$ 2) branching sucrase domain organization** (Brison et al., 2012). Magenta, domain C; blue, domain A, which includes the  $(\beta/\alpha)$ -8 barrel; green, domain B; yellow, domain IV; red, domain V. (B) schematic representation of the domain arrangement along the polypeptide chains of crystallized GH70 glucansucrases, from left to right, GTF180- $\Delta$ N, GTF-SI, and  $\Delta$ N123-GBD-CD2. The color code is identical to that for A. Striped red, parts of domains V of GTF-SI and the  $\alpha$ -(1 $\rightarrow$ 2) branching sucrase, which are not visible in the electron density map; white, purification tag.

A comparison of the  $\alpha$ -(1 $\rightarrow$ 2) branching sucrase domains A, B, C and IV with other protein structures in the PDB using the DALI server (Holm et al., 2008) showed that the enzyme structure is mostly similar to that of GTF180- $\Delta$ N. However, **the  $\alpha$ -(1 $\rightarrow$ 2) branching sucrase domain V is entirely situated at the N-terminus** (A1758 to L1980). It corresponds to the GBD domain previously identified (Brison et al., 2012). It is divided in three subdomains; each of them being similarly organized (Figure 35). A three-stranded  $\beta$ -sheet is connected to two consecutive  $\beta$ -hairpins. Each  $\beta$ -hairpin is constituted of two  $\beta$ -strands of 3 to 6 residues separated by short loops of 2 to 11 amino acids. The two  $\beta$ -hairpins and the three stranded  $\beta$ -sheets in each sub-domain are approximately rotated by 120° with respect to each other.

**Domain IV** is built up of two polypeptide segments including residues D1981 $\rightarrow$ N2123 and residues M2812 $\rightarrow$ T2831 (Figure 35 a). It can be superimposed with that of GTF180- $\Delta$ N with secondary structure elements conserved but slightly shifted. A DALI analysis showed no significant structural similarities of this domain to other proteins in the PDB (Brison, 2010).





**Figure 35. Domain V organization observed in the structures of GTF180-ΔN, GTFA-ΔN and DSR-E and ΔN123-GBD-CD2** (Leemhuis et al., 2013a). Hairpins and β-sheets motifs forming the subdomains are shown in orange and green, respectively. (a) Superposition of structural modules. (b) Schematic organization of domain V; boxes correspond to motifs.

**Domain B** results from the assembly of three polypeptide segments, formed by residues →L2172, Y2696→V2711, and W2795→G2811. It is folded into a five-stranded β-sheet. A comparison with domain B of GTF180-ΔN showed that 94 Cα out of 99 Cα superimpose well. However, the loop inserted between R2157 and **F2163**, also located at the upper part of the catalytic gorge is eleven residues shorter than the equivalent loop in GTF180-ΔN.

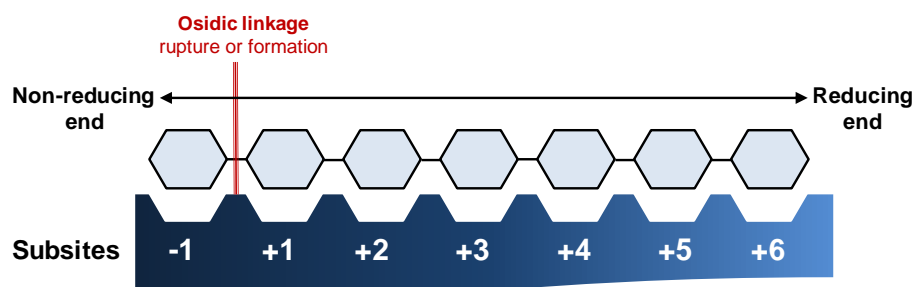
**Domain A** also consists of three polypeptide segments, which bring together residues L2173→S2424, S2561→V2695 and G2731→S2796. It represents the **largest domain of the enzyme**. It forms the catalytic core together with elements from domain B, and **supports the circularly permuted (α/β)<sub>8</sub> barrel** previously described, which is similar to that observed in GTF180-ΔN and GFT-SI (Ito et al., 2010, 2011; Vujičić-Žagar et al., 2010).

Finally, **domain C** consists of one polypeptide segment (G2425 to D2560) forming the bottom of the U-shape fold. It is composed of eight anti-parallel β-strands and includes a modified Greek-key motif. However, the function of this domain still remains unknown.

### III.2.3.b. Structure-based analysis of glucansucrase active sites

**Structural analysis of NpAS** in complex with maltoheptaose, sucrose and glucose allowed identification of the products binding site OB1 and its subdivision in seven subsites (Skov et al., 2002). The **subsites were numbered from -1 to +6**, according to the nomenclature proposed by Davies et al in 1997 for glycosyl hydrolases (Figure 36). Usually, subsites are denoted by numbers from -n (non-reducing end) to +n (reducing end), and

**glucosidic linkage cleavage occurs between -1 and +1 subsites** (Figure 36). The *NpAS* binding site OB1 is **closed by a salt bridge** (Figure 36; R509 and D144) which limits the negative binding subsites to -1 (Figure 37)



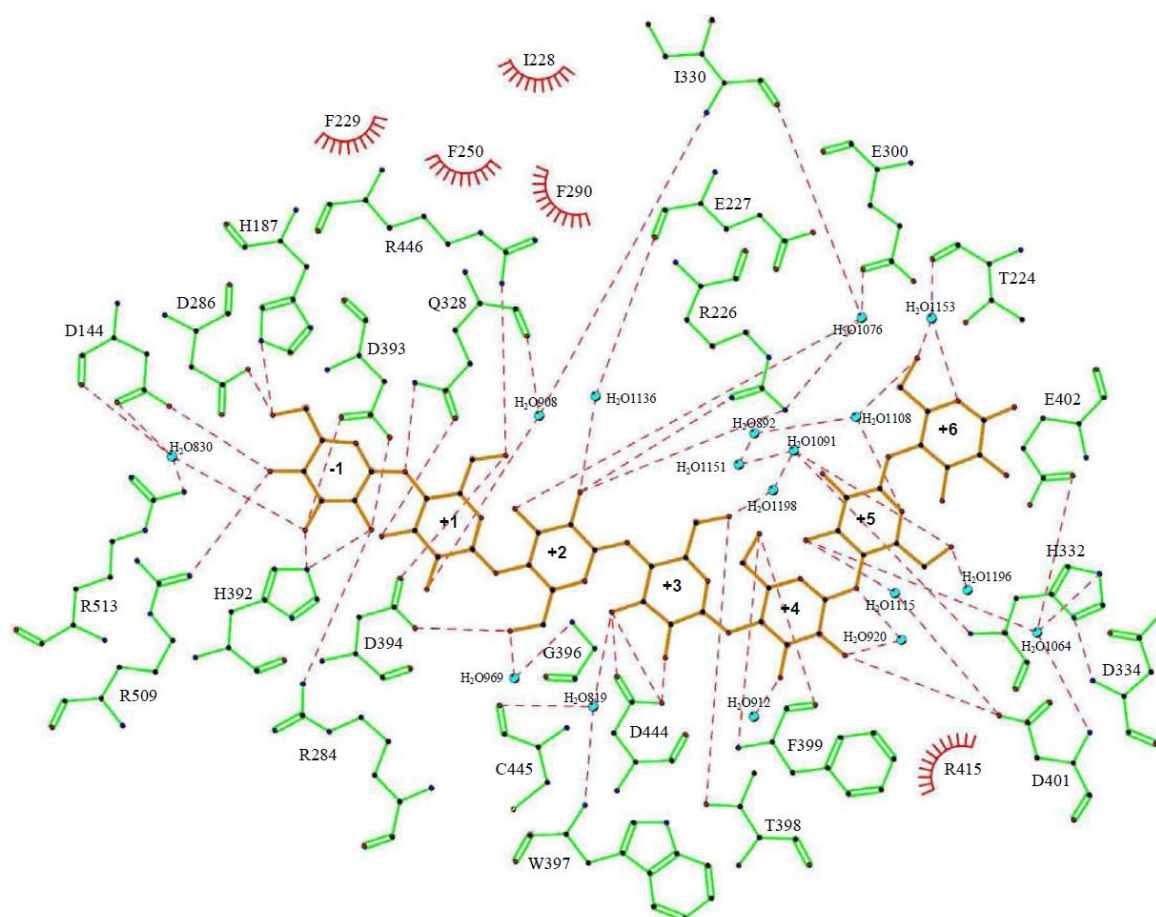
**Figure 36. Schematic representation of the acceptor binding subsites, according to Davies et al. proposed nomenclature (1997).**

**The glucosyl moiety of sucrose is bound in subsite -1 of *NpAS*** through dense network of hydrogen bonds strengthened by a double stacking involving the aromatic residue Y147 and F250 (Skov et al., 2002). Analysis of the  $\beta$ -D-glucosyl-enzyme intermediate structure confirmed the role of the catalytic residues involved in sucrose hydrolysis (Jensen et al., 2004; Skov et al., 2002).

Daudé et al. (2013b) conducted a site-directed mutagenesis strategy focused on *NpAS* subsite -1. They showed that none of the single mutations introduced at subsite -1 was sufficient to enable a productive binding of tested sucrose derivatives, confirming the **robustness of this *NpAS* catalytic region for sucrose binding only.**

**The fructosyl unit of sucrose is bound at subsite +1.** Binding is mostly ensured by **D394** and **R446** residues, as their mutations weaken the sucrose activity of *NpAS* (Albenne et al., 2002b). These two residues appeared especially important for the second step of the catalysis, corresponding to the acceptor accommodation in a productive manner, at the catalytic site of *NpAS* (Albenne et al., 2002b).

More recently, a complex of *NpAS* with turanose, a sucrose isomer, was structurally characterized (Guérin et al., 2012). Within the hydrogen bond network, the residue **R226**, at **subsite +2**, was shown to form **hydrogen bonds with** both O1' and O6' atoms of the **fructosyl unit**. It allows, **with I330**, a perfect positioning of the O3' for the nucleophilic attack of the  $\beta$ -glucosyl intermediate and the formation of turanose.

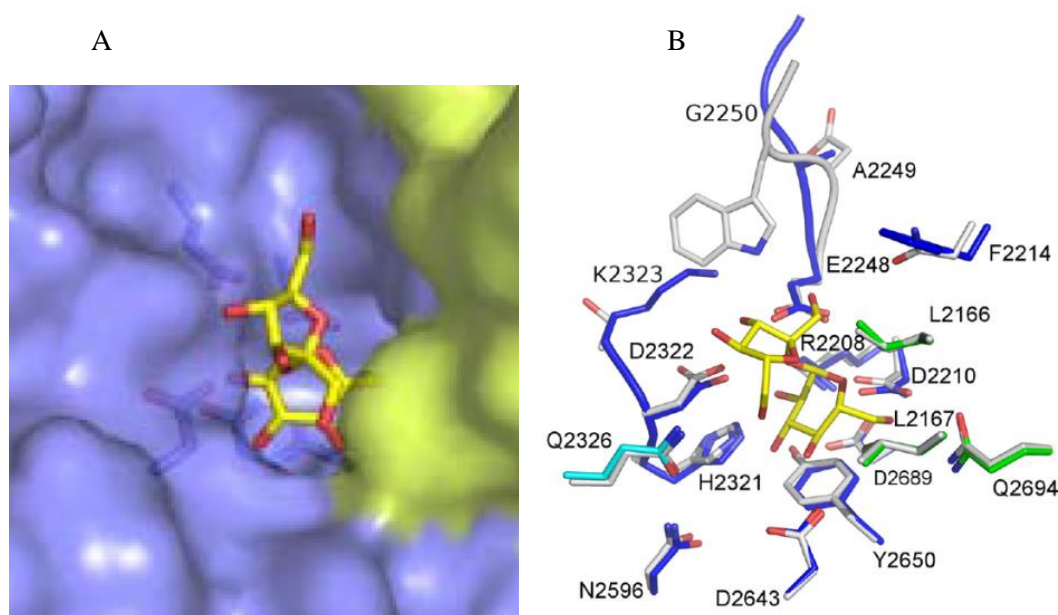


**Figure 37. Schematic planar view of NpAS catalytic pocket, in complex with maltoheptaose (PDB: 1MW0) with detailed maltoheptaose binding site and subsites (adapted from Guerin, 2012).**

To get better **insight in the structure-activity relationships of NpAS**, a study of OB1 (D394, R446, R226 and R415 residue) was conducted through an alanin scanning mutagenesis strategy (Albenne et al., 2004). Mutations **D394A** and **R446A**, at the NpAS subsite +1 (Figure 37) both induced a high sucrose hydrolytic activity of the mutants. Position **D394** was submitted to saturation mutagenesis and confirmed to be **crucial for amylase chains elongation** (Schneider et al., 2009). Indeed, most of the generated mutants were unable to synthesize the polymerization of more than two or three glucosyl units from sucrose. Structural models of the variants indicated that steric interference between the amino acids introduced at these sites and the growing oligosaccharide chain are mainly responsible for the limitation of glucosyl transfers. In contrast, **mutation of the residue R226 to a smaller alanin highly increases its polymerase activity**, probably by facilitating the binding of the maltotriose and longer chains in the active site of NpAS.

**The  $\alpha$ -(1→2) branching sucrase (GH70)** active site analysis was performed by superimposition of the structure of GTF180- $\Delta$ N inactive mutant in complex with sucrose (Brison et al., 2012; Figure38 A). **The active site of the  $\alpha$ -(1→2) branching sucrase forms a pocket included into a large gorge**, which is less constrained than that of the NpAS. As

seen in Figure 38 B, the residues defining the subsites -1 and +1 in GTF180- $\Delta$ N and  $\Delta$ N123-GBD-CD2 structure show high similarities. This allowed investigating the interactions between sucrose and the catalytic site of  $\alpha$ -(1 $\rightarrow$ 2) branching sucrose.



**Figure 38. (A) Active site of the  $\alpha$ -(1 $\rightarrow$ 2) branching sucrose superimposed with sucrose from GTF180- $\Delta$ N:sucrose complex structure (Brison, 2010). (B) View of subsites -1 and +1 of  $\Delta$ N123-GBD-CD2, with sucrose from the GTF180- $\Delta$ N-sucrose complex superimposed (Brison et al., 2012). The catalytic residues are D2210 (nucleophile), E2248 (acid/base), and D2322 (transition state stabilizer). Sucrose is shown with yellow carbons. Residues of the inactive GTF180- $\Delta$ N mutant (D1025N) that interact with sucrose are represented in gray. The carbon atoms of their structural equivalents in the  $\alpha$ -(1 $\rightarrow$ 2) branching sucrose structure are shown in blue (domain A) and green (domain B).**

In subsite -1, nine residues of GTF180- $\Delta$ N were reported to interact with the sucrose glucosyl ring and are conserved in  $\Delta$ N123-GBD-CD2 structure: R2208, **D2210 (nucleophile)**, **E2248 (acid/base)**, H2321, **D2322 (stabilizer)**, N2596, D2643, T2650 and Q2694. Among them, residues R2208, D2210, H2321, D2322, E2248 are also strictly conserved in subsite -1 of the GH family 13 members. In particular, **residues E2248 and D2210 coincide with the general acid base catalyst and nucleophile residues of the GH13 enzymes** (Jensen et al., 2004), and are assumed to play a similar role in the  $\alpha$ -(1 $\rightarrow$ 2) branching sucrose. Y2650 is also conserved in all glucansucrases from GH13 and GH70 families, and expected to provide a stacking platform for the glucosyl residue of sucrose.

In contrast, residues Q2694 and D2643 are all strictly conserved among GH70 glucansucrases but not in GH13 ones. Q2694 is equivalent to *Np*AS residue H187, which was proposed to stabilize of the  $\beta$ -D-glucosyl-enzyme transition state (Nakamura et al., 1993; Svensson, 1994). In the structure of GTF180- $\Delta$ N:sucrose complex, this residue is H-bonded to the glucosyl moiety C6 hydroxyl and its mutation strongly decreases GTF-I activity (Monchois et al., 2000a). By analogy to D1458 and N1411 residues of the GTF180-

$\Delta$ N:sucrose complex structure, **D2643 and N2596** could interact with C4 and C3 hydroxyl groups, ensuring a **strong binding of the sucrose glucosyl ring**. They are also suspected to be responsible of the active site pocket shape and the **closure of the subsites -2 and -3**, and could be the equivalent of residues D144 and R509 in *NpAS*.

At GTF180- $\Delta$ N **subsite +1**, Vujičić-Žagar et al. (2010) described seven residues **interacting with the fructosyl ring**. Five of them are conserved among the glucansucrases and correspond to  $\alpha$ -(1 $\rightarrow$ 2) branching sucrose residues **L2166, L2167, E2248, D2322 and Q2326**. The two leucine residues, located in the B domain, could interact through van der Waals interactions with the fructosyl moiety. E2248 and D2322 correspond to the catalytic residues, and Q2326 is H-bonded with the C6 hydroxyl of the fructosyl ring.

The fructosyl unit of sucrose also establishes weak interactions with N1029 and W1065, as shown in GTF180- $\Delta$ N:sucrose complex (Vujičić-Žagar et al., 2010). These two residues are strictly conserved among the characterized GH70 glucansucrases, but are substituted by a **phenylalanine (F2214)** and an **alanine (A2249)** respectively, in the  $\alpha$ -(1 $\rightarrow$ 2) branching sucrose structure, apparently unable to interact with the fructosyl unit of sucrose. Thus, **the fructosyl binding at subsite +1 may be weaker than that found other glucansucrases** and these residues may be responsible for the loss of polymerase activity. To investigate these hypotheses, the residues F2214, A2249 and G2250 were rationally mutated. Four mutants were constructed A2249W, G2250W, A2249D/G2250W and F2214N and showed to still consume sucrose. Thus, **the  $\alpha$ -(1 $\rightarrow$ 2) branching sucrose subsite +1 appeared tolerant to mutations with regard to sucrose utilization**, but all of the mutants showed reduced hydrolytic activity. No polymerization activities were detected, indicating that more complex changes have to be performed in the enzyme catalytic pocket to restore this activity. However, **residue F2214 appeared critical for dextran binding and branching**, as mutant F2214N was unable to use 1-kDa dextran as an acceptor, in contrast to the wild-type enzyme.

**To conclude, these structural investigations allowed deeply improving structure-function relationships studies of GS and have opened the way to GS semi-rational engineering aiming at improving or modifying their properties.** This will be described in the last part of this report, after a short presentation of *NpAs* and GBD-CD2 kinetic properties.



### III.3. Glucansucrases: Attractive tools for glucodiversification

#### III.3.1. *Np*As and $\alpha$ -(1→2) branching sucrase kinetic parameters

***Np*AS kinetics parameters** were determined by Potocki De Montalk et al., in 2000. *Np*AS initial velocity of sucrose consumption was modelled by **two Michaelis-Menten equations depending on sucrose concentrations** (Table 7).

**Table 7. Apparent kinetic values for sucrose consumption (ViS), sucrose hydrolysis(ViG) and polymerization (ViGx) (Potocki De Montalk et al., 2000)**

Initial [sucrose]	Apparent kinetic constants	ViS	ViG	ViGx
< 20 mM	$K_m$ $V_{max}$	1.9 mM 470 $\mu$ mol of sucrose consumed $\text{min}^{-1} \text{g}^{-1}$	1.7 mM 288 $\mu$ mol of glucose released $\text{min}^{-1} \text{g}^{-1}$	1.9 mM 147 $\mu$ mol of glucose incorporated into the $\alpha$ -glucan $\text{min}^{-1} \text{g}^{-1}$
> 20 mM	$K_{cat}$ $K_m$ $V_{max}$	33 $\text{min}^{-1}$ 50.2 mM 1100 $\mu$ mol of sucrose consumed $\text{min}^{-1} \text{g}^{-1}$	20 $\text{min}^{-1}$ 38.7 mM 472 $\mu$ mol of glucose released $\text{min}^{-1} \text{g}^{-1}$	10 $\text{min}^{-1}$ 387 mM 1620 $\mu$ mol of glucose incorporated into the $\alpha$ -glucan $\text{min}^{-1} \text{g}^{-1}$
	$K_{cat}$	77 $\text{min}^{-1}$	33 $\text{min}^{-1}$	113 $\text{min}^{-1}$

At sucrose concentrations above 20mM, the theoretical turnover of *Np*AS for sucrose consumption was twice higher than that determined from initial concentrations below 20 mM. But this is accompanied by a 25-fold increase of the  $K_m$  value. Thus, ***Np*AS catalytic efficiency ( $k_{cat}/K_m$ ) is lower at higher sucrose concentration**. In addition, a drop of initial velocity was also observed for sucrose concentration exceeding 300mM, which was attributed to the increase of reaction medium viscosity affecting substrate diffusion, or an inhibition caused by an excess of substrate (Potocki De Montalk et al., 2000).

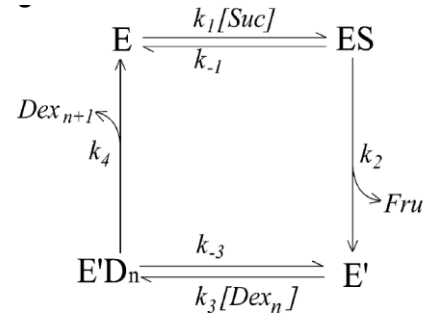
When comparing hydrolysis and polymerization at low sucrose concentrations, *Np*AS appears to favor the hydrolysis reaction, as its maximal velocity and turnover number were shown to be twice higher, at equivalent  $K_m$  (Table 7). In contrary, **at high sucrose concentrations, the polymerization reaction appeared largely favored** (Potocki De Montalk et al., 2000).

To investigate the **kinetic mechanism of the  $\alpha$ -(1→2) branching sucrase**, GBD-CD2 (192 kDa), steady-state kinetic analysis were conducted with sucrose and 70 kDa dextran (Brison et al., 2009). It showed that the enzyme displays a **ping-pong bi-bi mechanism of transglucosylation** and its kinetic parameters were determined (Figure 39).

Notably, the transglucosylation velocity is much higher than the hydrolysis one. Moreover, the  $k_{cat}$  obtained for the transglucosylation reaction is **one of the highest catalytic constants ( $970 \text{ s}^{-1}$ ) reported for glucansucrases, far higher than that of *Np*AS**. The Michaelis constant value for sucrose (42 mM) is close to that of *Np*AS (50 mM).

However, the reported  $K_{mT}$  values indicate a much higher affinity of the enzyme for 70 kDa dextran (0.174 mM). The presence of dextran was suggested to accelerate the deglycosylation of the glucosyl–enzyme intermediate.

Kinetic parameter	Rate constant	
$V_{maxT}$ ( $\mu\text{mol min}^{-1}\text{mg}^{-1}$ of purified enzyme)	$k_2k_4/(k_2 + k_4)$	$303 \pm 5$
$K_{mSucT}$ (mM)	$k_4(k_{-1} + k_2)/k_1(k_2 + k_4)$	$42 \pm 2$
$K_{mDexT}$ (mM of anhydroglucosyl unit equivalents)	$k_2(k_{-3} + k_4)/k_3(k_2 + k_4)$	$75 \pm 3$
$K_{mDexT}$ (mM)	$k_2(k_{-3} + k_4)/k_3(k_2 + k_4)$	$0.174 \pm 0.008$



**Figure 39. Kinetic parameters computed for the  $\alpha$ -(1 $\rightarrow$ 2) transglucosylation of  $\alpha$ -(1 $\rightarrow$ 6) dextrans catalyzed by the  $\alpha$ -(1 $\rightarrow$ 2) branching sucrase (Brison et al., 2009). A King–Altman diagram is also given.**

The  $\alpha$ -(1 $\rightarrow$ 2) branching sucrase activity well-fit to a Ping Pong Bi Bi mechanism for sucrose concentrations ranging from 10 to 300 mM. At higher sucrose concentrations, initial velocities decreased, probably due to the high viscosity of the reaction medium that limits mass transfer (Brison et al., 2009), as also reported for *NpAS*. Activity measurements were tested in the presence of  $\text{Ca}^{2+}$  chelating agent EDTA and showed that the  $\alpha$ -(1 $\rightarrow$ 2) branching sucrase is **not a calcium-dependent enzyme**.

The truncated form of the,  $\Delta\text{N}_{123}$ -GBD-CD2 kinetic parameters were close to that determined for the entire  $\alpha$ -(1 $\rightarrow$ 2) branching sucrase GBD-CD2 (Brison et al., 2012) (Table 8).

**Table 8. Comparison of the apparent kinetic parameters determined for sucrose hydrolysis and  $\alpha$ -(1 $\rightarrow$ 2) dextran branching activity of GBD-CD2 and  $\Delta\text{N}_{123}$ -GBD-CD2 (Brison et al., 2012)**

Apparent kinetic parameters	GBD-CD2 <sup>a</sup>	$\Delta\text{N}_{123}$ -GBD-CD2
<b>In the presence of sucrose</b>		
$V_{maxH}$ ( $\mu\text{mol}\cdot\text{min}^{-1}\cdot\text{mg}^{-1}$ of purified enzyme)	$34.6 \pm 0.5$	$36.3 \pm 0.6$
$K_{mSucH}$ (mM)	$10.8 \pm 0.8$	$7.5 \pm 1.0$
$k_{cat,SucH}$ ( $\text{s}^{-1}$ )	109	76
<b>In the presence of sucrose and dextran 70 kDa</b>		
$V_{maxT}$ ( $\mu\text{mol}\cdot\text{min}^{-1}\cdot\text{mg}^{-1}$ of purified enzyme)	$303 \pm 5$	$462 \pm 45$
$K_{m,SucT}$ (mM)	$42 \pm 2$	$206 \pm 34$
$K_{m,DexT}$ (mM of anhydroglucosyl units)	$75 \pm 3$	$125 \pm 21$
$K_{m,DexT}$ (mM)	$0.174 \pm 0.008$	$0.30 \pm 0.05$
$k_{catT}$ ( $\text{s}^{-1}$ )	970	947
$k_{catT}/K_{m,SucT}$ ( $\text{s}^{-1}\cdot\text{mM}^{-1}$ )	23	4.6
$k_{catT}/K_{m,DexT}$ ( $\text{s}^{-1}\cdot\text{mM}^{-1}$ )	13	7.6

In this thesis, all the experiments were conducted with this truncated form. Thus, in the following part of this thesis,  $\alpha$ -(1 $\rightarrow$ 2) branching sucrase appellation will always refer to the truncated  $\Delta\text{N}_{123}$ -GBD-CD2 form.

### III.3.2. Engineering of glucansucrases

Directed molecular evolution was first performed to increase catalytic efficiency of *NpAS* (van der Veen et al., 2004, 2006). A library of 100,000 variants was constructed, using error-prone PCR and gene shuffling, and screened for activity on sucrose. Two double mutants, V389L-N503I and R20C-F598S, were exhibited a 4-fold increase in their catalytic efficiency. A second library of 30,000 clones was also constructed to improve *NpAS* thermostability. The best variants, R20C-A451T, showed a 10-fold increase in thermostability at 50°C compared to the wild-type. When operated at a 600 mM sucrose concentration, the variant R20C-A451T was reported to catalyze the synthesis of twice longer amylose chains than those obtained with the wild-type enzyme (Emond et al., 2008b, 2008c).

In order to improve ***NpAS*** polymer synthesis, a rational engineering strategy allowed to isolate an amylosucrase variant with reduced side reactions (Albenne et al., 2004). The single **mutation R226A at subsite +2 enhanced transglucosylation activity** and led to a three-fold increase of insoluble amylose-like polymer.

Various linkage specificities are found in **GH70** enzymes and the identification of the structural determinants responsible for linkage specificity has been a challenge since more than 10 years. Many **engineering studies** were conducted to improve the structure-function relationships knowledge and propose new enzymatic tools **dedicated to tailor-made glucooligosaccharides or polymers**. An overview of these works is proposed in Table 9.

**Table 9. GH70 glucansucrases engineering studies for polymerization efficiency and/or specificity changes** (reported in literature at present).

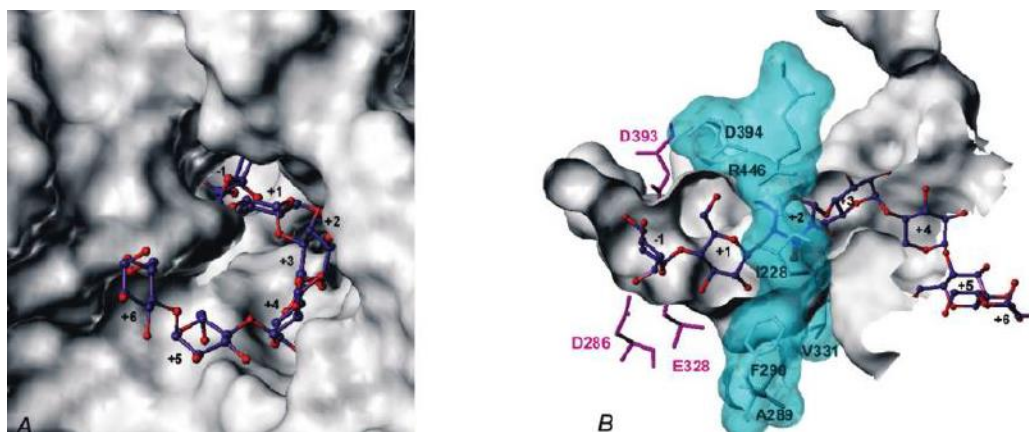
GH70 glucansucrase	Evolution strategy	Interesting mutants	Mutation effects	References
GTF-I	Site directed mutagenesis	Q937 and D569 mutants	Linkages specificity changed	Monchois et al., 2000a, 2000b
DSR-B 742	Random mutagenesis	DSRB742CK	Increase in activity and $\alpha$ -(1→3) branching.	Kang et al., 2003
DSR-S and DSR-T5 B-512F	Chimeric enzymes construction	DSRT5 (TST1)	Polymerization mainly with $\alpha$ -(1→4) linkages	Funane et al., 2004
GTF-A	Site directed mutagenesis	N1134Q, N1134G, N1134H	Loss in transglucosylation activity	Kralj et al., 2004
		N1134A, N1134D, N1134S	2-fold increase in transglucosylation activity	
		P1026V-I1029V-N1134S-N1135E-S1136V	High increase of $\alpha$ -(1→6) linkages	Kralj et al., 2005
DSR-E B-1299	Truncation	$\Delta$ N <sub>123</sub> -GBD-CD2	$\square$ -(1→2) branching enzyme	Fabre et al., 2005
DSR-S B-512F	Site directed mutagenesis	T350K, S455K	Dextrans with enhanced amounts of $\alpha$ -(1→6) linkages	Funane et al., 2005
		T350K-S455K	New unusual $\alpha$ -(2→6) linkages	
AlternansucraseB-1355	Truncation	ASR C-APY-del	Higher activity; linkage specificity retained	Joucla et al., 2006; Monsan et al., 2009
GTF-A and GTF-O	Chimeric enzymes construction	GTF-O-A-dN, GTF-O-dN-RS	Reduced hydrolysis; linkages specificity unchanged	Kralj et al., 2008



GTF-R	Random mutagenesis	R624G-V630I-D717A	Dextrans with increased $\alpha$ -(1 $\rightarrow$ 3) linkages	Hellmuth et al., 2008
	Site directed mutagenesis	S268D, S268R	Loss of polymerization	
GTF-180	Site directed mutagenesis	V1027P-S1137N-A1139S	New linkages specificity: $\alpha$ -(1 $\rightarrow$ 4)	van Leeuwen et al., 2008, 2009
DSR-B CB4 and DEX2 dextranase	Fusion proteins	DXSR	Improved synthesis of isomalto-oligosaccharides	Kim et al., 2009; Ryu et al., 2010
DSR-B CB4	Site directed mutagenesis and directed mutagenesis	V532P-E643N-V644S	Glucan with increased $\alpha$ -(1 $\rightarrow$ 3) linkages; new linkages specificity: $\alpha$ -(1 $\rightarrow$ 4)	Kang et al., 2011
		V532P-V535I-S642N-E643N-V644S	new linkages specificity: $\alpha$ -(1 $\rightarrow$ 4) improved	
DSR-S vardel $\Delta$ 4N (B512F) truncated mutant	Combinatorial engineering	F353T, S512C, F353W, H463R-T464D-S512C, H463R-T464V-S512C, D460A-H463S-T464L, D460M-H463Y-T464M-S512C	Dextrans with increased $\alpha$ -(1 $\rightarrow$ 3) linkage contents and gel-like properties in solution.	Irague et al., 2011, 2012
DSR-E 563	Truncation	DsrE563 $\Delta$ CD2 $\Delta$ GBD, DsrE563 $\Delta$ CD2 $\Delta$ VR	Increase in $\alpha$ -(1 $\rightarrow$ 3) branching.; importance of GBD	Kang et al., 2013
DSR-I	Site directed mutagenesis	T654K, T654R, T654N, T654Q, T654I	Glucan with increased $\alpha$ -(1 $\rightarrow$ 3) linkages	Côté and Skory, 2014

The natural promiscuity of glucanases towards various acceptors make them useful to generate new glucoderivatives (Daudé et al., 2013a). Nevertheless, effective glucosylation of **unnatural acceptors** often requires the generation of engineered enzymes with remodeled active-sites. **Structurally-guided engineering** is usually preferred in this case and has been shown to be really efficient and reduce screening efforts due to the smaller size of the library. This was well exemplified by Champion et al (2009, 2010).

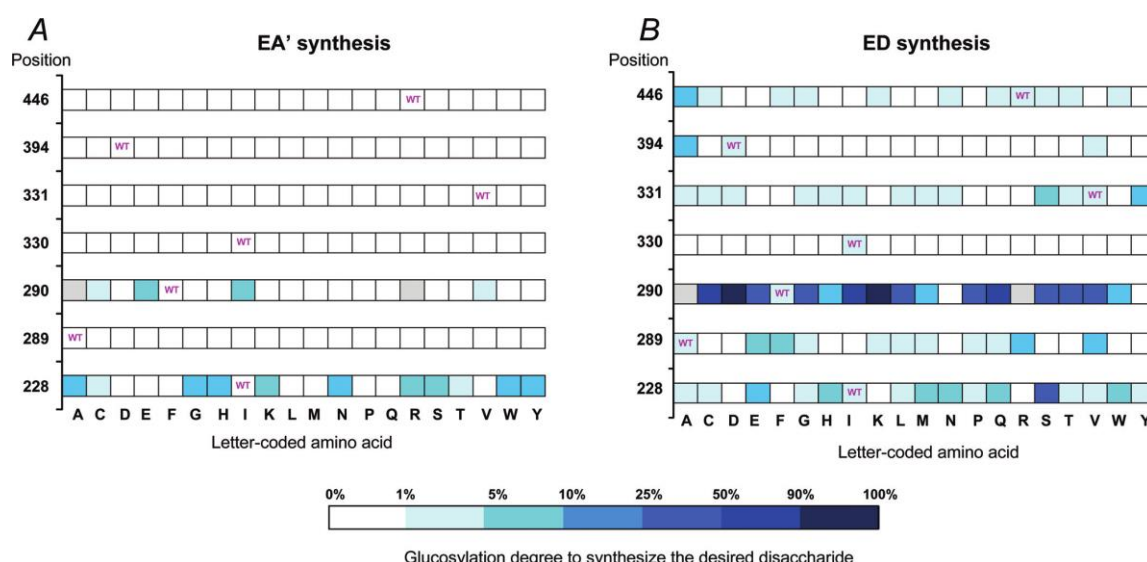
**NpAS** was **semi-rationally engineered** to generate mutants able to synthesize a carbohydrate part of the *Shigella flexneri* antigen O, difficult to obtain chemically and required for vaccine development (Champion et al., 2009, 2010).



**Figure 40. View of Amylosucrase active site in complex with maltoheptaose (PDB: 1MW0).** (A) Amylosucrase is viewed as a Connolly surface, as calculated by the MOLCAD module implemented in Syby17.3 (Champion et al., 2009). (B) Cross view of the active site. Residues forming the catalytic triad (D286, E328 and D393), located at the bottom of the pocket, are colored in magenta. The seven positions identified for site directed mutagenesis (I228, A289, F290, I330, V331, D394, and R446) are shown in cyan color.

The objective was to render the enzyme able to catalyze the regiospecific glucosylation of methyl  $\alpha$ -L-rhamnopyranoside (A') and allyl 2-acetamido-2-deoxy- $\alpha$ -D-glucopyranoside (allyl N-acetyl-D-glucosamine, D'), **two non-natural and protected osidic acceptors** entering in the preparation of third generation vaccines against shigellosis (Champion et al., 2009).

According to molecular modeling studies and docking analyzes, 7 positions out of the 18 residues that constitute the first shell of **subsite +1** were judged to be critical for acceptor binding (Figure 40). They were selected for **site-directed mutagenesis** and systematically replaced by the 19 other possible residues, to create a **small size library: I228, A289, F290, I330, V331, D394, and R446** (Figure 40). In contrast to wt-*NpAS*, which does not glucosylate methyl  $\alpha$ -L-rhamnopyranoside (A') to produce  $\alpha$ -DGlc-(1 $\rightarrow$ 3)- $\alpha$ -L-Rhap-O-Me (EA'), 15 mutants displayed a totally novel specificity toward this acceptor (Figure 41 A). **Remodeling of subsite +1 also led to marked improvement for the glucosylation** of N-acetyl-D-glucosamine (D) and the production  $\alpha$ -D-Glc-(1 $\rightarrow$ 4)-D-GlcNAc (ED) (Figure 41 B). From these results, F290 residue appears highly critical to improve ED production. (Champion et al., 2009).



**Figure 41. Screening of the library of AS monomutants for their ability to synthesize the desired disaccharide:** (A) EA' for methyl R-L-rhamnopyranoside. (B) ED for  $\alpha$ -D-Glcp-(1 $\rightarrow$ 4)-D-GlcpNAc (Champion et al., 2009).

Thus, to further improve NpAS F290 mutants, a new library was constructed focusing onto the adjacent amino acids A289 and F290 (Guérin et al., 2012). Indeed, simultaneous mutations onto vicinal amino acid residues are not frequently observed in nature. By **recombining pairwise mutations** at these positions up to 8-fold and 400-fold **improvements of the catalytic efficiency** towards sucrose and allyl 2- N-acetyl-2-deoxy- $\alpha$ -D-glucopyranoside substrate were respectively obtained for the best mutants (A289P-F290L, A289P-F290C and A289P-F290I) compared to wild-type parental enzyme. Molecular dynamics simulations revealed an alteration of the flexibility of some key structural regions that control active site topology (Guérin et al., 2012).

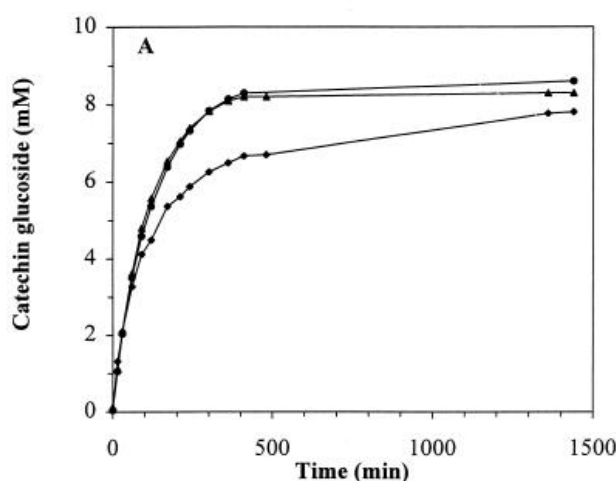
**The sole report about GH70 glucansucrases engineering for the glucosylation of a non osidic exogenous acceptor** was published in 2007 by Nam et al. The dextransucrase DSRB742 from *Leuconostoc mesenteroides* was first submitted to random mutagenesis via ultrasoft X-ray irradiation. Then, one of the generated variants, DSRN, was further rationally engineered and tested for salicin glucosylation. Mutants F196S and K395T successfully exhibited about a 2-fold increase in salicin conversion rates (~40%) compared to DSRB742 (~20%) and a 4-fold increase compared to DSRN (~10%).

### III.3.3. Glucosylation of flavonoids with glucansucrases

Glucansucrases have been successfully used as biocatalyst to glucosylate a wide range of different acceptors (Daudé et al., 2014), including some flavonoids. However, all the glucosylation attempts were **performed using wild-type GSs only**.

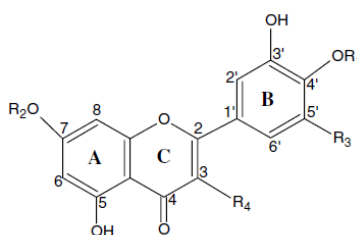
Historically, glucosylation of flavonoids with glucansucrases was **first attempted in 1995** by Nakahara et al. The **glucansucrase (GH70) from *Streptococcus sobrinus* 6715** was assayed for **(+)-catechin** (4 mM) glucosylation with sucrose (60 mM) as glucosyl unit donor. A production of 4'-O- $\alpha$ -D-glucopyranosyl-(+)-catechin (12.8 mg) was achieved, corresponding to a 13.7% conversion rate.

Few years later, **four recombinant glucansucrases (GH70) from the genus *Streptococcus*** were compared for their capacity to glucosylate the **(+)-catechin** (Meulenbeld et al., 1999). The recombinant glucansucrases from *Streptococcus sobrinus* SL-1, and a recombinant GTF-B and GTF-D from *Streptococcus mutans* GS-5, were expressed in *E. coli*. They all glucosylated the (+)-catechin. GTF-D exhibited the highest efficiency, with a (+)-catechin conversion rate about 80% yielded 7.9 mM of (+)-catechin glucosides from 10 mM catechin and 100 mM sucrose (Figure 42). **Two catechin monoglucosides and one diglucoside** were produced by GTF-D: (+)-catechin-4'-O- $\alpha$ -D-glucopyranoside, (+)-catechin-4',7-O- $\alpha$ -di-D-glucopyranoside and catechin-7-O- $\alpha$ -D-glucopyranoside. To further improve catechin glucosylation, **fructose inhibition was minimized by the addition of yeasts for fructose removal**. This resulted in a huge decrease in the reaction duration (Figure 42) together with an increase of catechin glucoside production up to 9.5 mM (95% conversion). Reaction containing water-miscible solvent at 15% (v/v) were also conducted with catechin (10 mM) sucrose (60 mM) and GTF-D (150 mU.mL<sup>-1</sup>), but the amount of catechin glucosides was lowered and only reached 7.5 mM (Meulenbeld, 2001). **Notably, Meulenbeld suggested that vicinal hydroxyl groups on the B ring of flavonoids were required to observe glucosylation with glucansucrase.**



**Figure 42.** Catechin glucoside formation by *S. mutans* GS-5 GTF-D in the absence (♦) or presence of *P. pastoris* (●) or *S. cerevisiae* T2-3D (▲) (Meulenbeld et al., 1999). The reaction solutions (5 mL) contained 60 mM sucrose, 10 mM catechin, and 37 mU of GTF-D per mL.

A few years later, Bertrand et al. (2006) reported the use of two glucansucrases for the glucosylation of **six different flavonoids**: luteolin, myricetin, **quercetin**, **diosmetin**,  $\beta$ -D-Glucopyranosyl-7-diosmetin and diosmin.



Flavonoid	Luteolin <sup>a</sup>		Quercetin <sup>b</sup>		Myricetin <sup>c</sup>	Diosmetin <sup>d</sup>	$\beta$ -D-Glucopyranosyl-7-diosmetin <sup>e</sup>	Diosmin <sup>f</sup>
Enzyme	DS*	ASR*	ASR*	ASR*	DS*	DS*	DS*	DS*
Conversion (%)	44%	8%	4%	4%	49%	0	0	0
Molar ratio of flavone monoglucoside	P1: 17%	P6: 2.5%	P9: 2.5%	P11: 27%	—	—	—	—
	P2: 27%	P7: 1.5%	P10: 1%	P12: 22%	—	—	—	—
Molar ratio of flavone di-glucoside	—	P4: 2%	P8: 0.5%	—	—	—	—	—
Molar ratio of flavone tri-glucoside	—	P5: 1.5%	—	—	—	—	—	—
	—	P3: 0.5%	—	—	—	—	—	—

DS\*: Dextranucrase, ASR\*: alternansucrase.

<sup>a</sup> R<sub>1</sub> = H, R<sub>2</sub> = H, R<sub>3</sub> = H, R<sub>4</sub> = H.

<sup>b</sup> R<sub>1</sub> = H, R<sub>2</sub> = H, R<sub>3</sub> = H, R<sub>4</sub> = OH.

<sup>c</sup> R<sub>1</sub> = H, R<sub>2</sub> = H, R<sub>3</sub> = OH, R<sub>4</sub> = OH.

<sup>d</sup> R<sub>1</sub> = CH<sub>3</sub>, R<sub>2</sub> = H, R<sub>3</sub> = H, R<sub>4</sub> = H.

<sup>e</sup> R<sub>1</sub> = CH<sub>3</sub>, R<sub>2</sub> = glucose, R<sub>3</sub> = H, R<sub>4</sub> = H.

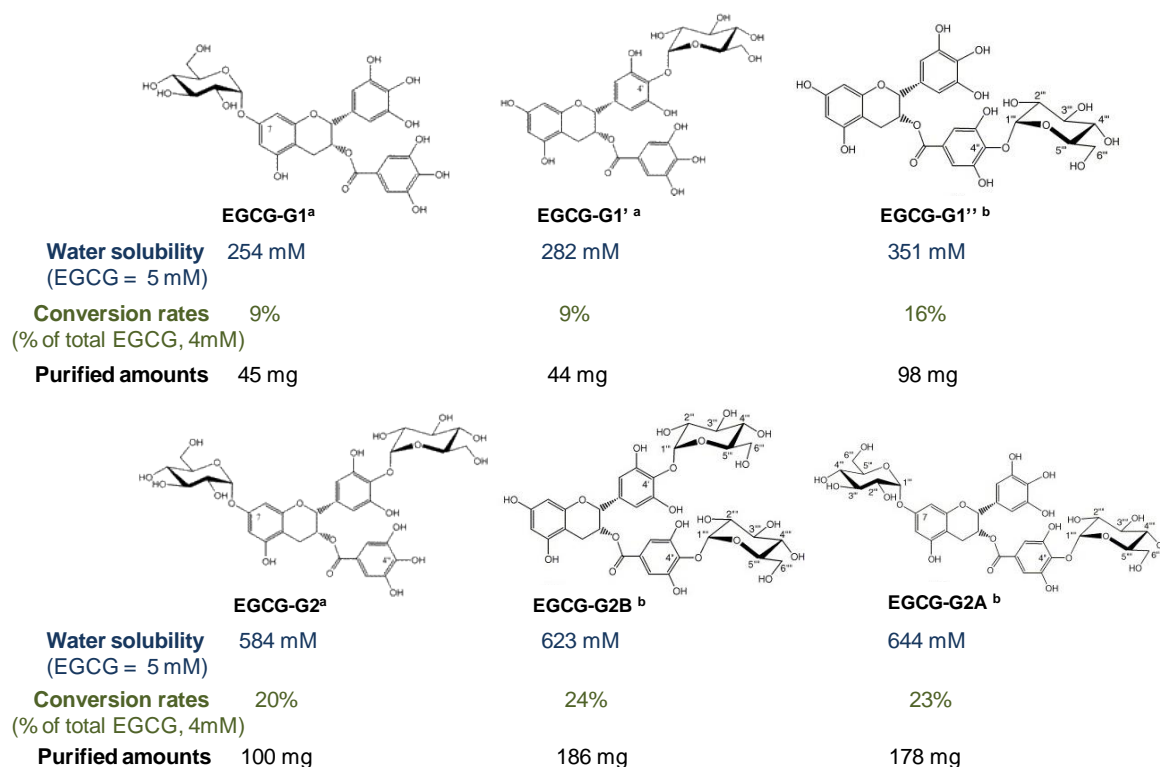
<sup>f</sup> R<sub>1</sub> = CH<sub>3</sub>, R<sub>2</sub> = rutinose, R<sub>3</sub> = H, R<sub>4</sub> = H.

**Table 10.** Transglucosylation of diverse flavonoids using either dextranucrase B-512F or alternansucrase B-23192 and molar ratio of the various flavonoid glucosides (Bertrand et al., 2006).

To improve flavonoid solubility during the reactions catalyzed by the **dextranucrase** from *Leuconostoc mesenteroides* NRRL B-512F (DS) or the **alternansucrase** from *L. mesenteroides* NRRL B-23192 (ASR), the acceptor reactions were performed in the presence of 30% (v/v) water miscible organic solvent. In these conditions, 4.5 mM luteolin and myricetin were glucosylated by DS to form **luteolin-4'-O- $\alpha$ -D-glucopyranoside** (P1 or P12) and **luteolin-3'-O- $\alpha$ -D-glucopyranoside** (P2 or P11). The regioselectivity of DS glucanucrase was similar for both flavonoids (Table 10). The two new luteolin  $\alpha$ -monoglucosides were found to be **8-times more soluble than the aglycon**. The position of the glucosyl residue onto the aromatic B-ring did not influence the solubility. Notably, all the other tested flavonoids were only poorly or not glucosylated by both ASR and DS glucanucrases, highlighting the **limited promiscuity of the wild-type GH70 glucanucrases for these acceptors**.

**Six epigallocatechin gallate (EGCG) glucosides** were also synthesized with **DSR-E glucanucrase** from *Leuconostoc citreum* NRRL B-1299 (2.4 U.mL<sup>-1</sup>) from EGCG (4mM) and sucrose (80 mM) (Moon et al., 2006a, 2006b). Three monoglucosides (EGCG-G1, EGCG-G1' and EGCG-G1'') and three diglucosides (EGCG-G2, EGCG-G2A and EGCG-G2B) were produced and purified at the milligram scale (Figure 43). The monoglucosides

and the diglucosides exhibited 50-fold and 120-fold higher water solubility than EGCG, respectively (Figure 43).



**Figure 43.** Structures of the EGCG glucosides reported in <sup>a</sup> Moon et al., 2006b and <sup>b</sup> Moon et al., 2006a. Water solubility, conversion rates and purified amounts of each glucosides were also given.

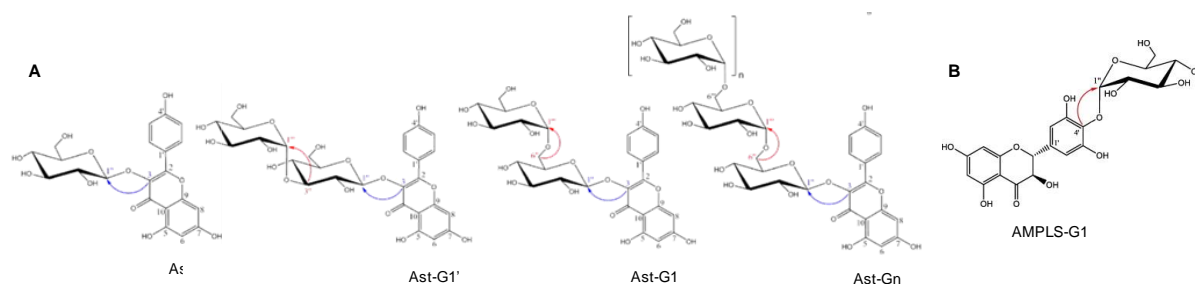
Reaction conditions were further optimized for a maximal production of EGCG-G1 glucoside (Hyun et al., 2007). The reaction was performed with 3.78 U.mL<sup>-1</sup> of enzyme, 25 mM sucrose and 1.5 mM EGCG at pH 5.0 and 25°C. EGCG conversion reached 67% yielding of 1.0 mM EGCG-G1.

**DSR-E glucansucrase** from *L. citreum* B-1299 (2.4 U.mL<sup>-1</sup>) was also used to catalyze quercetin (2 mM) glucosylation in a 400 mL reaction volume (Moon et al., 2007). Two quercetin monoglucosides were produced: quercetin-4'-O- $\alpha$ -D-glucopyranoside (amount, 50 mg; conversion rate, 20%) and **quercetin-3'-O- $\alpha$ -D-glucopyranoside** (amount, 7 mg; conversion rate, 3%). The solubility **Quercetin-4'-O- $\alpha$ -D-glucopyranoside in water was approximately 13 mM in water**, which represents a huge improvement as quercetin aglycon is only barely soluble in water.

The **sole study reporting the use of a glucansucrase from GH family 13** for flavonoid glucosylation was **published in 2011** by Cho et al. The **amylosucrase from *Deinococcus geothermalis* (DgAS)** was used for the glucosylation of the (+)-catechin (25 mM) using sucrose donor (25mM). Conversion of **(+)-catechin** was higher than 90%. Increasing *DgAS* concentration from 50 to 500 U.mL<sup>-1</sup> reduced the reaction time. The same

amount of (+)-catechin glucosides was produced in only 1 h rather compared to 24 h, with no significant changes in the product's profile. Two major products were synthesized: (+)-catechin-3'-O- $\alpha$ -D-glucopyranoside and (+)-catechin-3'-O- $\alpha$ -D-maltoside. The presence of (+)-catechin maltooligosaccharides at the end of the reaction was also confirmed.

In 2012, **astragalín** (kaempferol-3-O- $\beta$ -D-glucopyranoside, Ast) glucosides were synthesized using the **dextranucrase** from *Leuconostoc mesenteroides* B-512FMCM as biocatalyst ( $1.8 \text{ U.mL}^{-1}$ ) with astragalín (10 mM) and sucrose (200 mM) (Kim et al., 2012). Two monoglucosides were produced and their structures were assigned to kaempferol-3-O- $\beta$ -D-glucopyranosyl-(1 $\rightarrow$ 6)-O- $\alpha$ -D-glucopyranoside (Ast-G1) and kaempferol-3-O- $\beta$ -D-glucopyranosyl-(1 $\rightarrow$ 3)-O- $\alpha$ -D-glucopyranoside (Ast-G1') (Figure 44 A). Additionally, very low amounts of minor products corresponding to kaempferol-3-O- $\beta$ -D-isomaltooligosaccharides (Ast-Gn; n = 2–8) were produced (Figure 44 A). 97.6 mg and 12.3 mg of Ast-G1 and Ast-G1' were synthesized, corresponding to 21,8 % and 2.7% production yield, respectively. The solubility of these glucosides in water was not investigated.



**Figure 44. Structures of the: (A) astragalín glucosides** reported in Kim et al., 2012<sup>2</sup>; **(B) ampelopsin glucoside AMPLS-G1** reported in Woo et al., 2012.

The synthesis and characterization of **ampelopsin** (AMPLS) glucosides was reported the same year (Woo et al., 2012). The **dextranucrase DSR-E** from *L. citreum* B-1299 ( $1 \text{ U.mL}^{-1}$ ) was used to glucosylate the favanonol ampelopsin (70 mM) from sucrose (150 mM). About 87% of the AMPLS were converted to ampelopsin glucosides. AMPLS-G1 was characterized as an **ampelopsin-4'-O- $\alpha$ -D-glucopyranoside** (Figure 44 B). The D-glucopyranosyl residue was initially transferred to the B ring of AMPLS (AMPLS-G1) and additional glucosyl units were transferred via an  **$\alpha$ -(1 $\rightarrow$ 6) linkage** to this glucose unit to produce **AMPLS-glucoside with different degrees of polymerization** (AMPLS-Gn, n = 2–5). Then, an **experimental response surface methodology** was carried out to **optimize three reaction parameters** (dextranucrase unit, sucrose and ampelopsin concentration) to improve AMPLS-G1 production. The best production yielded 66 mM of AMPLS-G1 and was achieved using  $1 \text{ U.mL}^{-1}$  dextranucrase, 150 mM sucrose, and 70 mM AMPLS. Finally,



solubility of AMPLS-G1 and AMPLS-G2 in water were analyzed and reached 60 mM and 508 mM, respectively. The AMPLS solubility in water is 0.7 mM.

## IV. Thesis objectives

We have seen that flavonoid glycosides are very attractive compounds for pharmaceutical formulations. However, the access to natural glycosides is difficult due to their usually low level of production in plants. Our work was centered on the implementation of novel enzymatic routes for flavonoid glucosides diversification using engineered  $\alpha$ -transglycosylases. In particular, our approach aimed at taking advantage of the advancements of protein engineering in order to deliver enzymes adapted to the synthesis of new flavonoid  $\alpha$ -glucosides, hopefully revealing new properties in terms of solubility in aqueous medium, resistance to oxidation, bioavailability and bioactivity. Transglycosylases, and particularly the promiscuous glucansucrases, were shown to be promising biocatalysts to glucosylate this kind of molecules. They produce new non-natural  $\alpha$ -glucosides with original glucosylation patterns and improved water solubility. This has already led to patent registrations (Auriol et al., 2007, 2011, 2013). But much more work is still required to achieve efficient glucosylation of different types of flavonoids. Indeed, no reports of glucansucrases engineering for flavonoid glucosylation were found in the literature, despite the demonstrated success of such strategies for changing the specificity of these enzymes and improving the glucosylation of many acceptors.

Thus, the aim of this thesis was to use and generate new engineered glucansucrases able to efficiently glucosylate a selected panel of flavonoids. Two glucansucrases appeared very promising for this work: the *NpAs* and the  $\alpha$ -(1 $\rightarrow$ 2) branching sucrase. These enzymes use sucrose as glucosyl donor and not expensive nucleotide activated sugars. In addition, they were already both well-studied and their 3D structures were available, enabling the development of structurally-guided engineering strategies to improve and/or modify their specificity toward flavonoids glucosylation.

To achieve this goal, several challenges had to be taken up:

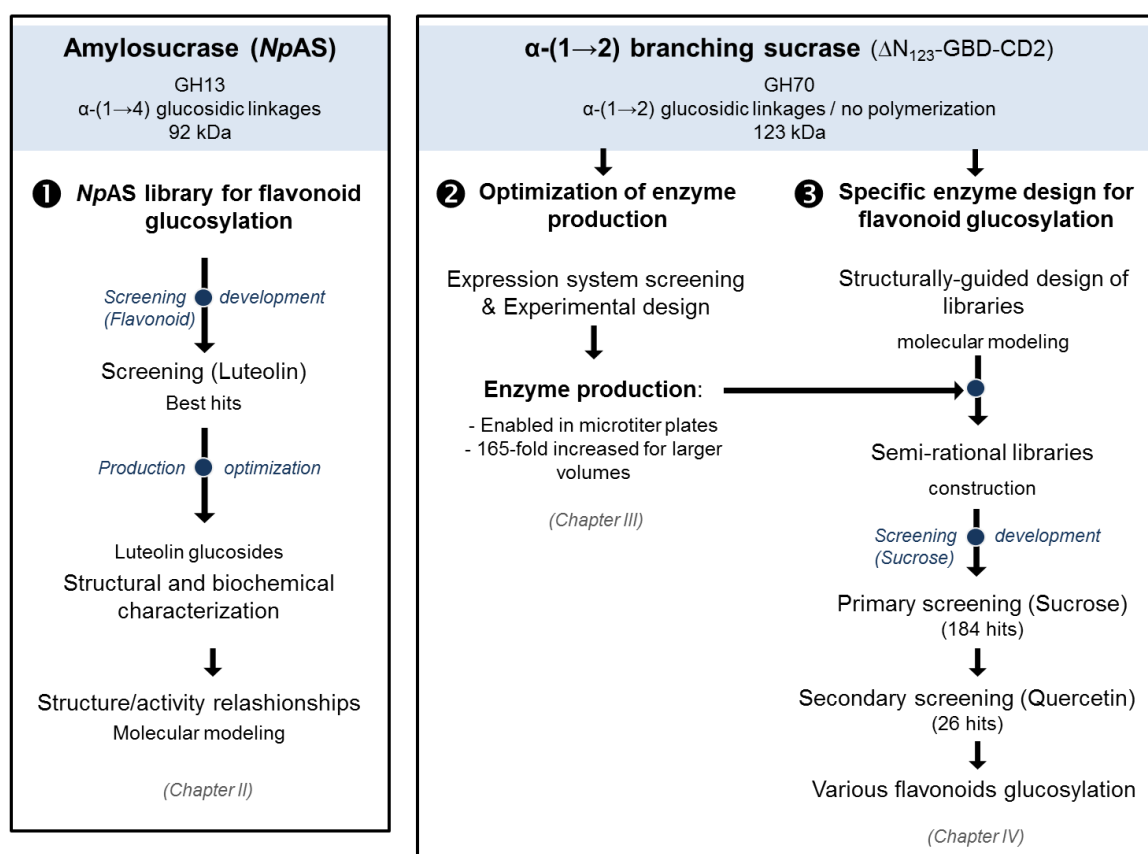
- Ensure an easy and screening-compatible production of these enzymes in micro scale volumes as well as in larger culture volumes.
- Develop efficient screening procedures enabling the accurate and efficient isolation of the improved mutants
- Improve the production of the new flavonoid glucosides to allow their purification in sufficient amounts for their structural and physicochemical characterization.



- Define and conduct a structurally-guided strategy using molecular modeling allowing the reduction of library size while improving the chance of finding improved variants.
- Achieved screening on various types of flavonoids.
- Generate mutants able to efficiently glucosylate many flavonoids, including some which had never been successfully glucosylated with this class of enzymes.
- Increase the knowledge on the structure/activity relationships of glucansucrases, in particular on their ability to recognize flavonoids as acceptor and facilitate future engineering work.

The workflow of the general strategy and organization of the work conducted during this thesis is given below:

### Engineering sucrose-active enzymes for flavonoid glucodiversification



## V. References

### A

- Achnine, L., Huhman, D.V., Farag, M.A., Sumner, L.W., Blount, J.W., and Dixon, R.A. (2005). Genomics-based selection and functional characterization of triterpene glycosyltransferases from the model legume *Medicago truncatula*. *Plant J.* 41, 875–887.
- Agati, G., and Tattini, M. (2010). Multiple functional roles of flavonoids in photoprotection. *New Phytol.* 186, 786–793.
- Agati, G., Azzarello, E., Pollastri, S., and Tattini, M. (2012). Flavonoids as antioxidants in plants: Location and functional significance. *Plant Sci.* 196, 67–76.
- Albenne, C., Skov, L.K., Mirza, O., Gajhede, M., Potocki-Véronèse, G., Monsan, P., and Remaud-Simeon, M. (2002a). Maltooligosaccharide disproportionation reaction: an intrinsic property of amylosucrase from *Neisseria polysaccharea*. *FEBS Lett.* 527, 67–70.
- Albenne, C., Potocki de Montalk, G., Monsan, P., Skov, L.K., Mirza, O., Gajhede, M., and Remaud-Simeon, M. (2002b). Site-directed mutagenesis of key amino acids in the active site of amylosucrase from *Neisseria polysaccharea*. *Biol. Bratisl.* 57, 119–128.
- Albenne, C., Skov, L.K., Mirza, O., Gajhede, M., Feller, G., D'Amico, S., André, G., Potocki-Véronèse, G., Veen, B.A. van der, Monsan, P., et al. (2004). Molecular Basis of the Amylose-like Polymer Formation Catalyzed by *Neisseria polysaccharea* Amylosucrase. *J. Biol. Chem.* 279, 726–734.
- Albenne, C., Skov, L.K., Tran, V., Gajhede, M., Monsan, P., Remaud-Siméon, M., and André-Leroux, G. (2007). Towards the molecular understanding of glycogen elongation by amylosucrase. *Proteins Struct. Funct. Bioinforma.* 66, 118–126.
- Alcaráz, L.E., Blanco, S.E., Puig, O.N., Tomàs, F., and Ferreti, F.H. (2000). Antibacterial Activity of Flavonoids Against Methicillin-resistant *Staphylococcus aureus* strains. *J. Theor. Biol.* 205, 231–240.
- Amellal, M., Bronner, C., Briancon, F., Haag, M., Anton, R., and Landry, Y. (2007). Inhibition of Mast Cell Histamine Release by Flavonoids and Biflavonoids. *Planta Med.* 51, 16–20.
- Amic, D., Davidovic-Amic, D., Beslo, D., Rastija, V., Lucic, B., and Trinajstić, N. (2007). SAR and QSAR of the Antioxidant Activity of Flavonoids. *Curr. Med. Chem.* 14, 827–845.
- 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.
- André, I., Potocki-Véronèse, G., Barbe, S., Moulis, C., and Remaud-Siméon, M. (2014). CAZyme discovery and design for sweet dreams. *Curr. Opin. Chem. Biol.* 19, 17–24.
- Anouar, E.H., Gierschner, J., Duroux, J.-L., and Trouillas, P. (2012). UV/Visible spectra of natural polyphenols: A time-dependent density functional theory study. *Food Chem.* 131, 79–89.
- Ara, K.Z.G., Khan, S., Kulkarni, T.S., Pozzo, T., and Karlsson, E.N. (2013). Glycoside Hydrolases for Extraction and Modification of Polyphenolic Antioxidants. In *Advances in Enzyme Biotechnology*, P. Shukla, and B.I. Pletschke, eds. (Springer India), pp. 9–21.
- Aramsangtienchai, P., Chavasiri, W., Ito, K., and Pongsawasdi, P. (2011). Synthesis of epicatechin glucosides by a  $\beta$ -cyclodextrin glycosyltransferase. *J. Mol. Catal. B Enzym.* 73, 27–34.
- Arita, M., and Suwa, K. (2008). Search extension transforms Wiki into a relational system: A case for flavonoid metabolite database. *BioData Min.* 1, 7.
- Arthur, C.M., Cummings, R.D., and Stowell, S.R. (2014). Using glycan microarrays to understand immunity. *Curr. Opin. Chem. Biol.* 18, 55–61.
- Auriol, D., Lefevre, F., Nalin, R., and Robe, P. (2007). Water soluble and activable phenolics derivatives with dermocosmetic and therapeutic applications and process for preparing said derivatives. U.S. Patent Application 12/304,212.
- Auriol, D., Lefevre, F., Nalin, R., and Robe, P. (2011). Water soluble and activable phenolics derivatives with dermocosmetic and therapeutic applications and process for preparing said derivatives. U.S. Patent Application 13/052,633.
- Auriol, D., Nalin, R., Robe, P., and Lefevre, F. (2013). Water soluble and activable phenolics derivatives with dermocosmetic and therapeutic applications and process for preparing said derivatives. U.S. Patent Application. 14/084,793.

- Badel, S., Bernardi, T., and Michaud, P. (2011). New perspectives for Lactobacilli exopolysaccharides. *Biotechnol. Adv.* 29, 54–66.
- Beil, W., Birkholz, C., and Sewing, K.F. (1995). Effects of flavonoids on parietal cell acid secretion, gastric mucosal prostaglandin production and *Helicobacter pylori* growth. *Arzneimittelforschung*. 45, 697–700.
- Bertrand, A., Morel, S., Lefoulon, F., Rolland, Y., Monsan, P., and Remaud-Simeon, M. (2006a). *Leuconostoc mesenteroides* glucansucrase synthesis of flavonoid glucosides by acceptor reactions in aqueous-organic solvents. *Carbohydr. Res.* 341, 855–863.
- Bertrand, A., Morel, S., Lefoulon, F., Rolland, Y., Monsan, P., and Remaud-Simeon, M. (2006b). *Leuconostoc mesenteroides* glucansucrase synthesis of flavonoid glucosides by acceptor reactions in aqueous-organic solvents. *Carbohydr. Res.* 341, 855–863.
- Bhagwat, S., Haytowitz, D.B., Wasswa-Kintu, S.I., and Holden, J.M. (2013). USDA Develops a Database for Flavonoids to Assess Dietary Intakes. *Procedia Food Sci.* 2, 81–86.
- Biesaga, M., and Pyrżyńska, K. (2013). Stability of bioactive polyphenols from honey during different extraction methods. *Food Chem.* 136, 46–54.
- Blanchard, S., and Thorson, J.S. (2006). Enzymatic tools for engineering natural product glycosylation. *Curr. Opin. Chem. Biol.* 10, 263–271.
- Boege, F., Straub, T., Kehr, A., Boesenberg, C., Christiansen, K., Andersen, A., Jakob, F., and Köhrle, J. (1996). Selected novel flavones inhibit the DNA binding or the DNA religation step of eukaryotic topoisomerase I. *J. Biol. Chem.* 271, 2262–2270.
- Bohm (1999). Introduction to Flavonoids. Chemistry and biochemistry of organic natural products, volume 2, CRC Press.
- Bohmont, C., Aaronson, L.M., Mann, K., and Pardini, R.S. (1987). Inhibition of mitochondrial NADH oxidase, succinoxidase, and ATPase by naturally occurring flavonoids. *J. Nat. Prod.* 50, 427–433.
- Bonina, F., Lanza, M., Montenegro, L., Puglisi, C., Tomaino, A., Trombetta, D., Castelli, F., and Saija, A. (1996). Flavonoids as potential protective agents against photo-oxidative skin damage. *Int. J. Pharm.* 145, 87–94.
- Bourne, Y., and Henrissat, B. (2001). Glycoside hydrolases and glycosyltransferases: families and functional modules. *Curr. Opin. Struct. Biol.* 11, 593–600.
- Bowles, D., Isayenkova, J., Lim, E.-K., and Poppenberger, B. (2005). Glycosyltransferases: managers of small molecules. *Curr. Opin. Plant Biol.* 8, 254–263.
- Bowles, D., Lim, E.-K., Poppenberger, B., and Vaistij, F.E. (2006). Glycosyltransferases of Lipophilic Small Molecules. *Annu. Rev. Plant Biol.* 57, 567–597.
- Bozonnet, S., Dols-Laffargue, M., Fabre, E., Pizzut, S., Remaud-Simeon, M., Monsan, P., and Willemot, R.-M. (2002). Molecular characterization of DSR-E, an  $\alpha$ -1,2 linkage-synthesizing dextranucrase with two catalytic domains. *J. Bacteriol.* 184, 5753–5761.
- Brazier-Hicks, M., Offen, W.A., Gershater, M.C., Revett, T.J., Lim, E.-K., Bowles, D.J., Davies, G.J., and Edwards, R. (2007). Characterization and engineering of the bifunctional N- and O-glucosyltransferase involved in xenobiotic metabolism in plants. *Proc. Natl. Acad. Sci.* 104, 20238–20243.
- Bresson, R.D., Mariotte, A.M., Boumendjel, A., and Perrier, E. (1999). Cosmetic, dermatological, pharmaceutical, dietetic or food compositions, e.g. for improving skin condition, comprise flavonoid esters.
- Brison, Y. (2010). Contribution à la compréhension du mécanisme de formation de dextrans ou gluco-oligosaccharides ramifiés en  $\alpha$ -1,2 par l'enzyme GBD-CD2: études cinétique et structurale (INSA de Toulouse).
- Brison, Y., Fabre, E., Moulis, C., Portais, J.-C., Monsan, P., and Remaud-Siméon, M. (2009). Synthesis of dextrans with controlled amounts of  $\alpha$ -1,2 linkages using the transglucosidase GBD-CD2. *Appl. Microbiol. Biotechnol.* 86, 545–554.
- Brison, Y., Pijning, T., Malbert, Y., Fabre, É., Mourey, L., Morel, S., Potocki-Véronèse, G., Monsan, P., Tranier, S., Remaud-Siméon, M., et al. (2012). Functional and structural characterization of  $\alpha$ -(1→2) branching sucrase derived from DSR-E glucansucrase. *J. Biol. Chem.* 287, 7915–7924.
- Brison, Y., Laguerre, S., Lefoulon, F., Morel, S., Monties, N., Potocki-Véronèse, G., Monsan, P., and Remaud-Simeon, M. (2013). Branching pattern of gluco-oligosaccharides and 1.5 kDa dextran grafted by the  $\alpha$ -1,2 branching sucrase GBD-CD2. *Carbohydr. Polym.* 94, 567–576.
- Brown, A.K., Papaemmanouil, A., Bhowruth, V., Bhatt, A., Dover, L.G., and Besra, G.S. (2007). Flavonoid inhibitors as novel antimycobacterial agents targeting Rv0636, a putative dehydratase enzyme involved in *Mycobacterium tuberculosis* fatty acid synthase II. *Microbiology* 153, 3314–3322.

Brown, J.E., Khodr, H., Hider, R.C., and Rice-Evans, C.A. (1998). Structural dependence of flavonoid interactions with Cu<sup>2+</sup> ions: implications for their antioxidant properties. *Biochem. J.* 330 ( Pt 3), 1173–1178.

Brusselmans, K., Vrolix, R., Verhoeven, G., and Swinnen, J.V. (2005). Induction of cancer cell apoptosis by flavonoids is associated with their ability to inhibit fatty acid synthase activity. *J. Biol. Chem.* 280, 5636–5645.

Burbulis, I.E., and Winkel-Shirley, B. (1999). Interactions among enzymes of the Arabidopsis flavonoid biosynthetic pathway. *Proc. Natl. Acad. Sci.* 96, 12929–12934.

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.

## C

Calias, P., Galanopoulos, T., Maxwell, M., Khayat, A., Graves, D., Antoniadis, H.N., and d' Alarcao, M. (1996). Synthesis of inositol 2-phosphate-quercetin conjugates. *Carbohydr. Res.* 292, 83–90.

Campbell, J.A., Davies, G.J., Bulone, V., and Henrissat, B. (1997). A classification of nucleotide-diphospho-sugar glycosyltransferases based on amino acid sequence similarities. *Biochem. J.* 326 ( Pt 3), 929–939.

Cantarel, B.L., Coutinho, P.M., Rancurel, C., Bernard, T., Lombard, V., and Henrissat, B. (2009a). The Carbohydrate-Active EnZymes database (CAZy): an expert resource for Glycogenomics. *Nucleic Acids Res.* 37, D233–238.

Cantarel, B.L., Coutinho, P.M., Rancurel, C., Bernard, T., Lombard, V., and Henrissat, B. (2009b). The Carbohydrate-Active EnZymes database (CAZy): an expert resource for Glycogenomics. *Nucleic Acids Res.* 37, D233–238.

Capasso, A., Pinto, A., Sorrentino, R., and Capasso, F. (1991a). Inhibitory effects of quercetin and other flavonoids on electrically-induced contractions of guinea pig isolated ileum. *J. Ethnopharmacol.* 34, 279–281.

Capasso, A., Pinto, A., Mascolo, N., Autore, G., and Capasso, F. (1991b). Reduction of agonist-induced contractions of guinea-pig isolated ileum by flavonoids. *Phytother. Res.* 5, 85–87.

Cartwright, A.M., Lim, E.-K., Kleanthous, C., and Bowles, D.J. (2008). A Kinetic Analysis of Regiospecific Glucosylation by Two Glycosyltransferases of *Arabidopsis thaliana*. *J. Biol. Chem.* 283, 15724–15731.

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.

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.

Chang, T.-S. (2009). An Updated Review of Tyrosinase Inhibitors. *Int. J. Mol. Sci.* 10, 2440–2475.

Chang, M.C.Y., and Keasling, J.D. (2006). Production of isoprenoid pharmaceuticals by engineered microbes. *Nat. Chem. Biol.* 2, 674–681.

Chang, M.C.Y., Eachus, R.A., Trieu, W., Ro, D.-K., and Keasling, J.D. (2007). Engineering *Escherichia coli* for production of functionalized terpenoids using plant P450s. *Nat. Chem. Biol.* 3, 274–277.

Chang, W.S., Lee, Y.J., Lu, F.J., and Chiang, H.C. (1993). Inhibitory effects of flavonoids on xanthine oxidase. *Anticancer Res.* 13, 2165–2170.

Chebil, L., Humeau, C., Falcimaigne, A., Engasser, J.-M., and Ghoul, M. (2006). Enzymatic acylation of flavonoids. *Process Biochem.* 41, 2237–2251.

Chebil, L., Humeau, C., Anthoni, J., Dehez, F., Engasser, J.M., and Ghoul, M. (2007). Solubility of flavonoids in organic solvents. *J. Chem. Eng. Data* 52, 1552–1556.

Chen, D., Daniel, K.G., Chen, M.S., Kuhn, D.J., Landis-Piowar, K.R., and Dou, Q.P. (2005). Dietary flavonoids as proteasome inhibitors and apoptosis inducers in human leukemia cells. *Biochem. Pharmacol.* 69, 1421–1432.

Chen, S., Xing, X.-H., Huang, J.-J., and Xu, M.-S. (2011). Enzyme-assisted extraction of flavonoids from Ginkgo biloba leaves: Improvement effect of flavonol transglycosylation catalyzed by *Penicillium decumbens* cellulase. *Enzyme Microb. Technol.* 48, 100–105.

Cheynier, V., Sarni-Manchado, P., and Quideau, S. (2012). Recent advances in polyphenol research (John Wiley & Sons).

- 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., et al. (2011). Biosynthesis of (+)-catechin glycosides using recombinant amylosucrase from *Deinococcus geothermalis* DSM 11300. *Enzyme Microb. Technol.* 49, 246–253.
- Cho, J.S., Yoo, S.S., Cheong, T.K., Kim, M.J., Kim, Y., and Park, K.H. (2000). Transglycosylation of neohesperidin dihydrochalcone by *Bacillus stearothermophilus* maltogenic amylase. *J. Agric. Food Chem.* 48, 152–154.
- Choi, C.-H., Kim, J.-W., Park, C.-S., Park, K.-H., Kyung, M.-O., Park, J.-T., Lee, C.-K., Kim, Y.-W., Lee, J.-H., Park, S.-H., et al. (2010a). Enzymatic biosynthesis of a puerarin–cycloamylose inclusion complex by 4- $\alpha$ -glucanotransferase and maltogenic amylase. *Biocatal. Biotransformation* 28, 209–214.
- Choi, C.-H., Kim, S.-H., Jang, J.-H., Park, J.-T., Shim, J.-H., Kim, Y.-W., and Park, K.-H. (2010b). Enzymatic synthesis of glycosylated puerarin using maltogenic amylase from *Bacillus stearothermophilus* expressed in *Bacillus subtilis*. *J. Sci. Food Agric.* 90, 1179–1184.
- Choi, S.H., Ryu, M., Yoon, Y.J., Kim, D.-M., and Lee, E.Y. (2012). Glycosylation of various flavonoids by recombinant oleandomycin glycosyltransferase from *Streptomyces antibioticus* in batch and repeated batch modes. *Biotechnol. Lett.* 34, 499–505.
- Clere, N., Faure, S., Martinez, M.C., and Andriantsitohaina, R. (2011). Anticancer properties of flavonoids: roles in various stages of carcinogenesis. *Cardiovasc. Hematol. Agents Med. Chem.* 9, 62–77.
- Cody, V., Middleton, E., and Harborne, J.B. (1986). Plant flavonoids in biology and medicine: biochemical, pharmacological, and structure-activity relationships: proceedings of a symposium held in Buffalo, New York, July 22–26, 1985. *Prog. Clin. Biol. Res. USA*.
- Conseil, G., Baubichon-Cortay, H., Dayan, G., Jault, J.-M., Barron, D., and Pietro, A.D. (1998). Flavonoids: A class of modulators with bifunctional interactions at vicinal ATP- and steroid-binding sites on mouse P-glycoprotein. *Proc. Natl. Acad. Sci.* 95, 9831–9836.
- Cook, N.C., and Samman, S. (1996). Flavonoids-Chemistry, metabolism, cardioprotective effects, and dietary sources. *J. Nutr. Biochem.* 7, 66–76.
- Corcoran, M.P., McKay, D.L., and Blumberg, J.B. (2012). Flavonoid basics: chemistry, sources, mechanisms of action, and safety. *J. Nutr. Gerontol. Geriatr.* 31, 176–189.
- Côté, G.L., and Skory, C.D. (2014). Effects of mutations at threonine-654 on the insoluble glucan synthesized by *Leuconostoc mesenteroides* NRRL B-1118 glucansucrase. *Appl. Microbiol. Biotechnol.*
- Coutinho, P.M., Deleury, E., Davies, G.J., and Henrissat, B. (2003). An evolving hierarchical family classification for glycosyltransferases. *J. Mol. Biol.* 328, 307–317.
- Crespy, V., Morand, C., Besson, C., Cotellet, N., Vézin, H., Demigné, C., and Rémésy, C. (2003). The splanchnic metabolism of flavonoids highly differed according to the nature of the compound. *Am. J. Physiol. - Gastrointest. Liver Physiol.* 284, G980–G988.
- Croft, K.D. (1998). The Chemistry and Biological Effects of Flavonoids and Phenolic Acids. *Ann. N. Y. Acad. Sci.* 854, 435–442.
- Cushnie, T.P.T., and Lamb, A.J. (2005). Antimicrobial activity of flavonoids. *Int. J. Antimicrob. Agents* 26, 343–356.

## D

- Dajas, F., Juan Andres, A.-C., Florencia, A., Carolina, E., and Felicia, R.-M. (2013). Neuroprotective Actions of Flavones and Flavonols: Mechanisms and Relationship to Flavonoid Structural Features. *Cent. Nerv. Syst. Agents Med. Chem. - Cent. Nerv. Syst. Agents* 13, 30–35.
- Daudé, D., Champion, E., Morel, S., Guieysse, D., Remaud-Siméon, M., and André, I. (2013a). Probing Substrate Promiscuity of Amylosucrase from *Neisseria polysaccharea*. *ChemCatChem* 5, 2288–2295.
- Daudé, D., Topham, C.M., Remaud-Siméon, M., and André, I. (2013b). Probing impact of active site residue mutations on stability and activity of *Neisseria polysaccharea* amylosucrase. *Protein Sci.* 22, 1754–1765.
- Daudé, D., André, I., Monsan, P., and Remaud-Siméon, M. (2014). Chapter 28. Successes in engineering glucansucrases to enhance glycodiversification. In *Carbohydrate Chemistry*, A. Pilar Rauter, T. Lindhorst, and Y. Queneau, eds. (Cambridge: Royal Society of Chemistry), pp. 624–645.
- Davies, G., and Henrissat, B. (1995). Structures and mechanisms of glycosyl hydrolases. *Structure* 3, 853–859.
- Davies, G.J., Wilson, K.S., and Henrissat, B. (1997). Nomenclature for sugar-binding subsites in glycosyl hydrolases. *Biochem. J.* 557–559.

Desmet, T., Soetaert, W., Bojarová, P., Křen, V., Dijkhuizen, L., Eastwick-Field, V., and Schiller, A. (2012). Enzymatic Glycosylation of Small Molecules: Challenging Substrates Require Tailored Catalysts. *Chem. – Eur. J.* 18, 10786–10801.

Devulapalle, K.S., Goodman, S.D., Gao, Q., Hemsley, A., and Mooser, G. (2008). Knowledge-based model of a glucosyltransferase from the oral bacterial group of mutans streptococci. *Protein Sci.* 6, 2489–2493.

Di Carlo, G., Autore, G., Izzo, A.A., Maiolino, P., Mascolo, N., Viola, P., Diurno, M.V., and Capasso, F. (1993). Inhibition of intestinal motility and secretion by flavonoids in mice and rats: structure-activity relationships. *J. Pharm. Pharmacol.* 45, 1054–1059.

Di Carlo, G., Mascolo, N., Izzo, A.A., and Capasso, F. (1999). Flavonoids: Old and new aspects of a class of natural therapeutic drugs. *Life Sci.* 65, 337–353.

Dijkhuizen, L., Dobruchowska, J.M., Kamerling, J.P., Kralj, S., and Leemhuis, R.J. (2010). Glucooligosaccharides comprising (alpha 1-> 4) and (alpha 1-> 6) glycosidic bonds, use thereof, and methods for providing them.

Dixon, R.A., and Steele, C.L. (1999). Flavonoids and isoflavonoids – a gold mine for metabolic engineering. *Trends Plant Sci.* 4, 394–400.

Dobruchowska, J.M., Gerwig, G.J., Kralj, S., Grijpstra, P., Leemhuis, H., Dijkhuizen, L., and Kamerling, J.P. (2012). Structural characterization of linear isomalto-/malto-oligomer products synthesized by the novel GTFB 4, 6- $\alpha$ -glucanotransferase enzyme from *Lactobacillus reuteri* 121. *Glycobiology* 22, 517–528.

Duarte, S., Gregoire, S., Singh, A.P., Vorsa, N., Schaich, K., Bowen, W.H., and Koo, H. (2006). Inhibitory effects of cranberry polyphenols on formation and acidogenicity of *Streptococcus mutans* biofilms. *FEMS Microbiol. Lett.* 257, 50–56.

## E

Einerhand, A., Lopez, M., Monsan, P.F.E., Naeye, T., Potter, S.M., and Remaud-Siméon, M. (2010). Compositions and methods for making alpha-(1, 2)-branched alpha-(1, 6) oligodextrans. U.S. Patent Application 12/775,656.

Emim, J.A.D.S., Oliveira, A.B., and Lapa, A.J. (1994). Pharmacological Evaluation of the Anti-inflammatory Activity of a Citrus Bioflavonoid, Hesperidin, and the Isoflavonoids, Diquartin and Clausenquinone, in Rats and Mice. *J. Pharm. Pharmacol.* 46, 118–122.

Emond, S., Mondeil, S., Jaziri, K., André, I., Monsan, P., Remaud-Siméon, M., and Potocki-Véronèse, G. (2008a). Cloning, purification and characterization of a thermostable amylosucrase from *Deinococcus geothermalis*. *FEMS Microbiol. Lett.* 285, 25–32.

Emond, S., André, I., Jaziri, K., Potocki-Véronèse, G., Mondon, P., Bouayadi, K., Kharrat, H., Monsan, P., and Remaud-Simeon, M. (2008b). Combinatorial engineering to enhance thermostability of amylosucrase. *Protein Sci. Publ. Protein Soc.* 17, 967–976.

Emond, S., Mondon, P., Pizzut-Serin, S., Douchy, L., Crozet, F., Bouayadi, K., Kharrat, H., Potocki-Véronèse, G., Monsan, P., and Remaud-Simeon, M. (2008c). A novel random mutagenesis approach using human mutagenic DNA polymerases to generate enzyme variant libraries. *Protein Eng. Des. Sel.* 21, 267–274.

Engelbertz, J., Lechtenberg, M., Studt, L., Hensel, A., and Verspohl, E.J. (2012). Bioassay-guided fractionation of a thymol-deprived hydrophilic thyme extract and its antispasmodic effect. *J. Ethnopharmacol.* 141, 848–853.

Erleijman, A.G., Verstraeten, S.V., Fraga, C.G., and Oteiza, P.I. (2004). The Interaction of Flavonoids with Membranes: Potential Determinant of Flavonoid Antioxidant Effects. *Free Radic. Res.* 38, 1311–1320.

## F

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

Fang, Z., and Bhandari, B. (2010). Encapsulation of polyphenols – a review. *Trends Food Sci. Technol.* 21, 510–523.

Fang, J.-Y., Hwang, T.-L., Huang, Y.-L., and Fang, C.-L. (2006). Enhancement of the transdermal delivery of catechins by liposomes incorporating anionic surfactants and ethanol. *Int. J. Pharm.* 310, 131–138.

Faria, P.A., Fernandes, I., Mateus, N., and Calhau, C. (2013). Bioavailability of Anthocyanins. In *Natural Products*, K.G. Ramawat, and J.-M. Mérillon, eds. (Springer Berlin Heidelberg), pp. 2465–2487.

- Fernández, S.P., Wasowski, C., Loscalzo, L.M., Granger, R.E., Johnston, G.A.R., Paladini, A.C., and Marder, M. (2006). Central nervous system depressant action of flavonoid glycosides. *Eur. J. Pharmacol.* 539, 168–176.
- Ferrell, J.E., Jr, Chang Sing, P.D., Loew, G., King, R., Mansour, J.M., and Mansour, T.E. (1979). Structure/activity studies of flavonoids as inhibitors of cyclic AMP phosphodiesterase and relationship to quantum chemical indices. *Mol. Pharmacol.* 16, 556–568.
- Ferrer, J.-L., Austin, M.B., Stewart, C., Jr, and Noel, J.P. (2008). Structure and function of enzymes involved in the biosynthesis of phenylpropanoids. *Plant Physiol. Biochem.*, 46, 356–370.
- Ferretti, J.J., Gilpin, M., and Russell, R. (1987). Nucleotide sequence of a glucosyltransferase gene from *Streptococcus sobrinus* MFe28. *J. Bacteriol.* 169, 4271–4278.
- Ferriola, P.C., Cody, V., and Middleton Jr, E. (1989). Protein kinase C inhibition by plant flavonoids: Kinetic mechanisms and structure-activity relationships. *Biochem. Pharmacol.* 38, 1617–1624.
- Ford, C.M., Boss, P.K., and Hoj, P.B. (1998). Cloning and characterization of *Vitis vinifera* UDP-glucose:flavonoid 3-O-glucosyltransferase, a homologue of the enzyme encoded by the maize Bronze-1 locus that may primarily serve to glucosylate anthocyanidins in vivo. *J. Biol. Chem.* 273, 9224–9233.
- Forkmann, G., and Heller, W. (1999). 1.26 - Biosynthesis of Flavonoids. In *Comprehensive Natural Products Chemistry*, S.D. Barton, K. Nakanishi, and O. Meth-Cohn, eds. (Oxford: Pergamon), pp. 713–748.
- Formica, J.V., and Regelson, W. (1995). Review of the biology of quercetin and related bioflavonoids. *Food Chem. Toxicol.* 33, 1061–1080.
- Funane, K., Ishii, T., Terasawa, K., Yamamoto, T., and Kobayashi, M. (2004). Construction of chimeric glucansucrases for analyzing substrate-binding regions that affect the structure of glucan products. *Biosci. Biotechnol. Biochem.* 68, 1912–1920.
- Funane, K., Ishii, T., Ono, H., and Kobayashi, M. (2005). Changes in linkage pattern of glucan products induced by substitution of Lys residues in the dextranase. *FEBS Lett.* 579, 4739–4745.
- Funayama, M., Nishino, T., Hirota, A., Murao, S., Takenisi, S., and Nakano, H. (1993). Enzymatic Synthesis of (+)Catechin- $\alpha$ -glucoside and Its Effect on Tyrosinase Activity. *Biosci. Biotechnol. Biochem.* 57, 1666–1669.

## G

- Gachon, C.M.M., Langlois-Meurinne, M., and Saindrenan, P. (2005). Plant secondary metabolism glucosyltransferases: the emerging functional analysis. *Trends Plant Sci.* 10, 542–549.
- Gamet-Payraastre, L., Manenti, S., Gratacap, M.-P., Tulliez, J., Chap, H., and Payraastre, B. (1999). Flavonoids and the inhibition of PKC and PI 3-kinase. *Gen. Pharmacol. Vasc. Syst.* 32, 279–286.
- Gao, C., Mayon, P., MacManus, D.A., and Vulfson, E.N. (2000). Novel enzymatic approach to the synthesis of flavonoid glycosides and their esters. *Biotechnol. Bioeng.* 71, 235–243.
- García-Mediavilla, V., Crespo, I., Collado, P.S., Esteller, A., Sánchez-Campos, S., Tuñón, M.J., and González-Gallego, J. (2007). The anti-inflammatory flavones quercetin and kaempferol cause inhibition of inducible nitric oxide synthase, cyclooxygenase-2 and reactive C-protein, and down-regulation of the nuclear factor kappaB pathway in Chang Liver cells. *Eur. J. Pharmacol.* 557, 221–229.
- Ghayur, M.N., Khan, H., and Gilani, A.H. (2007). Antispasmodic, bronchodilator and vasodilator activities of (+)-catechin, a naturally occurring flavonoid. *Arch. Pharm. Res.* 30, 970–975.
- Giffard, P., and Jacques, N. (1994). Definition of a fundamental repeating unit in streptococcal glucosyltransferase glucan-binding regions and related sequences. *J. Dent. Res.* 73, 1133–1141.
- Gilbert, H.J. (2010). The Biochemistry and Structural Biology of Plant Cell Wall Deconstruction. *Plant Physiol.* 153, 444–455.
- Girish, C., Raj, V., Arya, J., and Balakrishnan, S. (2012). Evidence for the involvement of the monoaminergic system, but not the opioid system in the antidepressant-like activity of ellagic acid in mice. *Eur. J. Pharmacol.* 682, 118–125.
- Go, Y.-H., Kim, T.-K., Lee, K.-W., and Lee, Y.-H. (2007). Functional characteristics of cyclodextrin glucanotransferase from alkalophilic *Bacillus* sp. BL-31 highly specific for intermolecular transglycosylation of bioflavonoids. *J. Microbiol. Biotechnol.* 17, 1550–1553.
- Gong, Z., Yamazaki, M., Sugiyama, M., Tanaka, Y., and Saito, K. (1997). Cloning and molecular analysis of structural genes involved in anthocyanin biosynthesis and expressed in a forma-specific manner in *Perilla frutescens*. *Plant Mol. Biol.* 35, 915–927.
- Groenwall, A., and Ingelman, B. (1948). Manufacture of infusion and injection fluids. US Pat. 2, 2–437.

Guardia, T., Rotelli, A.E., Juarez, A.O., and Pelzer, L.E. (2001). Anti-inflammatory properties of plant flavonoids. Effects of rutin, quercetin and hesperidin on adjuvant arthritis in rat. *Il Farm.* 56, 683–687.

Guerin, F. (2012). Caractérisation structurale et fonctionnelle d'amylosaccharases. Thesis, LISBP, INSA, France.

Guérin, F., Barbe, S., Pizzut-Serin, S., Potocki-Véronèse, G., Guieysse, D., Guillet, V., Monsan, P., Mourey, L., Remaud-Siméon, M., and André, I. (2012). Structural investigation of the thermostability and product specificity of amylosucrase from the bacterium *Deinococcus geothermali*s. *J. Biol. Chem.* 287, 6642–6654.

Guerrero, L., Castillo, J., Quiñones, M., Garcia-Vallvé, S., Arola, L., Pujadas, G., and Muguerza, B. (2012). Inhibition of angiotensin-converting enzyme activity by flavonoids: structure-activity relationship studies. *PLoS One* 7, e49493.

Gurung, R.B., Kim, E.-H., Oh, T.-J., and Sohng, J.K. (2013). Enzymatic synthesis of apigenin glucosides by glucosyltransferase (YjiC) from *Bacillus licheniformis* DSM 13. *Mol. Cells* 36, 355–361.

## H

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.

Halliwell, B. (2007). Flavonoids: a re-run of the carotenoids story? *Novartis Found. Symp.* 282, 93–101; discussion 101–104, 212–218.

Halliwell, B., and Gutteridge, J.M.C. (1985). *Free radicals in biology and medicine* (Clarendon Press).

Hamada, H. (2007). The High Functionality of Glycosylation Using Plant Suspension Cells. TCIMAIL No. 151..

Hancock, S.M., Vaughan, M.D., and Withers, S.G. (2006). Engineering of glycosidases and glycosyltransferases. *Curr. Opin. Chem. Biol.* 10, 509–519.

Harborne, J.B. (1973). *Phytochemical methods: a guide to modern techniques of plant analysis* (Chapman & Hall, London)

Harborne, J.B. (1989). *Methods in plant biochemistry*. Volume 1. Plant phenolics. xii + 552 pp.

Harborne, J.B., and Mabry, T.J. (1982). *The Flavonoids: advances in research* (Chapman & Hall, London).

Harborne, J.B., and Williams, C.A. (2000). Advances in flavonoid research since 1992. *Phytochemistry* 55, 481–504.

Harborne, J.B. (Jeffrey B., Mabry, T.J. (Tom J.), and Mabry, H. (1975). *The Flavonoids* (Chapman & Hall, London).

Havsteen, B. (1983). Flavonoids, a class of natural products of high pharmacological potency. *Biochem. Pharmacol.* 32, 1141–1148.

Havsteen, B.H. (novembre). The biochemistry and medical significance of the flavonoids. *Pharmacol. Ther.* 96, 67–202.

He, X.-Z., Wang, X., and Dixon, R.A. (2006). Mutational Analysis of the Medicago Glycosyltransferase UGT71G1 Reveals Residues That Control Regioselectivity for (Iso)flavonoid Glycosylation. *J. Biol. Chem.* 281, 34441–34447.

He, X.-Z., Li, W.-S., Blount, J.W., and Dixon, R.A. (2008). Regioselective synthesis of plant (iso)flavone glycosides in *Escherichia coli*. *Appl. Microbiol. Biotechnol.* 80, 253–260.

Hehre, E.J. (1949). Synthesis of a polysaccharide of the starch-glycogen class from sucrose by a cell free, bacterial enzyme system. *J. Biol. Chem.* 177, 267–279.

Heijnen, C.G.M., Haenen, G.R.M.M., van Acker, F.A.A., van der Vijgh, W.J.F., and Bast, A. (2001). Flavonoids as peroxynitrite scavengers: the role of the hydroxyl groups. *Toxicol. In Vitro* 15, 3–6.

Heim, K.E., Tagliaferro, A.R., and Bobilya, D.J. (2002). Flavonoid antioxidants: chemistry, metabolism and structure-activity relationships. *J. Nutr. Biochem.* 13, 572–584.

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

Hertog, M.G.L., Hollman, P.C.H., and van de Putte, B. (1993). Content of potentially anticarcinogenic flavonoids of tea infusions, wines, and fruit juices. *J. Agric. Food Chem.* 41, 1242–1246.

Hijiya, H., and Miyake, T. (1990). Alpha-Glycosyl hesperidin, and its preparation and uses. U.S. Patent Application 5,652,124.



- Hijiya, H., Miyake, T., Suzuki, Y., Suzuki, K., and Yoneyama, M. (1992). Preparation and uses of alpha-glycosyl rutin. U.S. Patent Application. 5,145,781.
- Hirano, R., Sasamoto, W., Matsumoto, A., Itakura, H., Igarashi, O., and Kondo, K. (2001). Antioxidant ability of various flavonoids against DPPH radicals and LDL oxidation. *J. Nutr. Sci. Vitaminol. (Tokyo)* 47, 357–362.
- Hiromoto, T., Honjo, E., Tamada, T., Noda, N., Kazuma, K., Suzuki, M., and Kuroki, R. (2013). Crystal structure of UDP-glucose:anthocyanidin 3- O -glucosyltransferase from *Clitoria ternatea*. *J. Synchrotron Radiat.* 20, 894–898.
- Hodek, P., Trefil, P., and Stiborová, M. (2002). Flavonoids-potent and versatile biologically active compounds interacting with cytochromes P450. *Chem. Biol. Interact.* 139, 1–21.
- Hollman, P.C.H., and Arts, I.C.W. (2000). Flavonols, flavones and flavanols – nature, occurrence and dietary burden. *J. Sci. Food Agric.* 80, 1081–1093.
- Holm, L., Kääriäinen, S., Rosenström, P., and Schenkel, A. (2008). Searching protein structure databases with DaliLite v. 3. *Bioinformatics* 24, 2780–2781.
- Hrazdina, G. (1992). Compartmentation in Aromatic Metabolism. In *Phenolic Metabolism in Plants*, H.A. Stafford, and R.K. Ibrahim, eds. (Springer US), 1–23.
- Hrazdina, G., and Wagner, G.J. (1985). Metabolic pathways as enzyme complexes: Evidence for the synthesis of phenylpropanoids and flavonoids on membrane associated enzyme complexes. *Arch. Biochem. Biophys.* 237, 88–100.
- Hutzler, P., Fischbach, R., Heller, W., Jungblut, T.P., Reuber, S., Schmitz, R., Veit, M., Weissenböck, G., and Schnitzler, J.-P. (1998). Tissue localization of phenolic compounds in plants by confocal laser scanning microscopy. *J. Exp. Bot.* 49, 953–965.
- Hyun, E.-K., Park, H.-Y., Kim, H.-J., Lee, J.-K., Kim, D., and Oh, D.-K. (2007). Production of epigallocatechin gallate 7-O-alpha-D-glucopyranoside (EGCG-G1) using the glucosyltransferase from *Leuconostoc mesenteroides*. *Biotechnol. Prog.* 23, 1082–1086.
- Hyung Ko, J., Gyu Kim, B., and Joong-Hoon, A. (2006). Glycosylation of flavonoids with a glucosyltransferase from *Bacillus cereus*. *FEMS Microbiol. Lett.* 258, 263–268.

## I

- Ioannou, I., and Ghoul, M. (2012). Biological Activities and Effects of Food Processing on Flavonoids as Phenolic Antioxidants. In *Advances in Applied Biotechnology*, pp. 101–124.
- Irague, R., Massou, S., Moulis, C., Saurel, O., Milon, A., Monsan, P., Remaud-Siméon, M., Portais, J.-C., and Potocki-Véronèse, G. (2011). NMR-based structural glycomics for high-throughput screening of carbohydrate-active enzyme specificity. *Anal. Chem.* 83, 1202–1206.
- Irague, R., Rolland-Sabaté, A., Tarquis, L., Doublier, J.L., Moulis, C., Monsan, P., Remaud-Siméon, M., Potocki-Véronèse, G., and Buléon, A. (2012). Structure and Property Engineering of  $\alpha$ -D-Glucans Synthesized by Dextranucrase Mutants. *Biomacromolecules* 13, 187–195.
- Ito, K., Ito, S., Shimamura, T., Kawarasaki, Y., Abe, K., Misaka, T., Kobayashi, T., and Iwata, S. (2010). Crystallization and preliminary X-ray analysis of a glucansucrase from the dental caries pathogen *Streptococcus mutans*. *Acta Crystallograph. Sect. F Struct. Biol. Cryst. Commun.* 66, 1086–1088.
- 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.

## J

- Jackman, R.L., Yada, R.Y., Tung, M.A., and Speers, R.A. (1987). Anthocyanins as Food Colorants —a Review. *J. Food Biochem.* 11, 201–247.
- Jayaraj, R., Deb, U., Bhaskar, A.S.B., Prasad, G.B.K.S., and Rao, P.V.L. (2007). Hepatoprotective efficacy of certain flavonoids against microcystin induced toxicity in mice. *Environ. Toxicol.* 22, 472–479.
- Jedinák, A., Maliar, T., Grancai, D., and Nagy, M. (2006). Inhibition activities of natural products on serine proteases. *Phytother. Res. PTR* 20, 214–217.
- 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 (Mosc.)* 43, 3104–3110.

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.

Ji, H.-F., and Zhang, H.-Y. (2006). Theoretical evaluation of flavonoids as multipotent agents to combat Alzheimer's disease. *J. Mol. Struct. THEOCHEM* 767, 3–9.

Jiang, J., Yuan, S., Ding, J., Zhu, S., Xu, H., Chen, T., Cong, X., Xu, W., Ye, H., and Dai, Y. (2008). Conversion of puerarin into its 7-O-glycoside derivatives by *Microbacterium oxydans* (CGMCC 1788) to improve its water solubility and pharmacokinetic properties. *Appl. Microbiol. Biotechnol.* 81, 647–657.

Jones, P., and Vogt, T. (2001). Glycosyltransferases in secondary plant metabolism: tranquilizers and stimulant controllers. *Planta* 213, 164–174.

Jørgensen, K., Rasmussen, A.V., Morant, M., Nielsen, A.H., Bjarnholt, N., Zagrobelny, M., Bak, S., and Møller, B.L. (2005). Metabolon formation and metabolic channeling in the biosynthesis of plant natural products. *Curr. Opin. Plant Biol.* 8, 280–291.

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

Jung, N.R., Joe, E.J., Kim, B.-G., Ahn, B.C., Park, J.C., Chong, Y., and Ahn, J.-H. (2010a). Change of *Bacillus cereus* flavonoid O-triglucosyltransferase into flavonoid O-monoglucosyltransferase by error-prone polymerase chain reaction. *J. Microbiol. Biotechnol.* 20, 1393–1396.

Jung, U.J., Kim, H.J., Lee, J.S., Lee, M.K., Kim, H.O., Park, E.J., Kim, H.K., Jeong, T.S., and Choi, M.S. (2003). Naringin supplementation lowers plasma lipids and enhances erythrocyte antioxidant enzyme activities in hypercholesterolemic subjects. *Clin. Nutr.* 22, 561–568.

Jung, W.Y., Park, S.J., Park, D.H., Kim, J.M., Kim, D.H., and Ryu, J.H. (2010b). Quercetin impairs learning and memory in normal mice via suppression of hippocampal phosphorylated cyclic AMP response element-binding protein expression. *Toxicol. Lett.* 197, 97–105.

## K

Kang, H.K., Kimura, A., and Kim, D. (2011). Bioengineering of *Leuconostoc mesenteroides* Glucansucrases That Gives Selected Bond Formation for Glucan Synthesis and/or Acceptor-Product Synthesis. *J. Agric. Food Chem.* 59, 4148–4155.

Kang, H.-K., Seo, E.-S., Robyt, J.F., and Kim, D. (2003). Directed evolution of a dextranucrase for increased constitutive activity and the synthesis of a highly branched dextran. *J. Mol. Catal. B Enzym.* 26, 167–176.

Kang, H.-K., Ecklund, D., Liu, M., and Datta, S.K. (2009). Apigenin, a non-mutagenic dietary flavonoid, suppresses lupus by inhibiting autoantigen presentation for expansion of autoreactive Th1 and Th17 cells. *Arthritis Res. Ther.* 11, R59.

Kang, H.-K., Ko, E.-A., Kim, J.-H., and Kim, D. (2013). Molecular cloning and characterization of active truncated dextranucrase from *Leuconostoc mesenteroides* B-1299CB4. *Bioprocess Biosyst. Eng.* 36, 857–865.

Katalinić, M., Rusak, G., Domaćinović Barović, J., Sinko, G., Jelić, D., Antolović, R., and Kovarik, Z. (2010). Structural aspects of flavonoids as inhibitors of human butyrylcholinesterase. *Eur. J. Med. Chem.* 45, 186–192.

Katsoura, M.H., Polydera, A.C., Katapodis, P., Kolis, F.N., and Stamatis, H. (2007). Effect of different reaction parameters on the lipase-catalyzed selective acylation of polyhydroxylated natural compounds in ionic liquids. *Process Biochem.* 42, 1326–1334.

Kaviarasan, K., and Pugalendi, K.V. (2009). Influence of flavonoid-rich fraction from *Spermocoe hispida* seed on PPAR- $\alpha$  gene expression, antioxidant redox status, protein metabolism and marker enzymes in high-fat-diet fed STZ diabetic rats. *J. Basic Clin. Physiol. Pharmacol.* 20, 141–158.

Keegstra, K., and Raikhel, N. (2001). Plant glycosyltransferases. *Curr. Opin. Plant Biol.* 4, 219–224.

Kim, B.-G., Kim, H.J., and Ahn, J.-H. (2012a). Production of Bioactive Flavonol Rhamnosides by Expression of Plant Genes in *Escherichia coli*. *J. Agric. Food Chem.* 60, 11143–11148.

Kim, B.H., Cho, S.M., Reddy, A.M., Kim, Y.S., Min, K.R., and Kim, Y. (2005). Down-regulatory effect of quercitrin gallate on nuclear factor-kappa B-dependent inducible nitric oxide synthase expression in lipopolysaccharide-stimulated macrophages RAW 264.7. *Biochem. Pharmacol.* 69, 1577–1583.

Kim, G.-E., Kang, H.-K., Seo, E.-S., Jung, S.-H., Park, J.-S., Kim, D.-H., Kim, D.-W., Ahn, S.-A., Sunwoo, C., and Kim, D. (2012b). Glucosylation of the flavonoid, astragalin by *Leuconostoc mesenteroides* B-512FMCM dextranucrase acceptor reactions and characterization of the products. *Enzyme Microb. Technol.* 50, 50–56.

- Kim, H.J., Kim, B.G., Kim, J.A., Park, Y., Lee, Y.J., Lim, Y., and Ahn, J.-H. (2007). Glycosylation of flavonoids with *E. coli* expressing glycosyltransferase from *Xanthomonas campestris*. *J. Microbiol. Biotechnol.* 17, 539–542.
- Kim, H.J., Kim, B.-G., and Ahn, J.-H. (2013a). Regioselective synthesis of flavonoid bisglycosides using *Escherichia coli* harboring two glycosyltransferases. *Appl. Microbiol. Biotechnol.* 97, 5275–5282.
- Kim, H.S., Kim, B.-G., Sung, S., Kim, M., Mok, H., Chong, Y., and Ahn, J.-H. (2013b). Engineering flavonoid glycosyltransferases for enhanced catalytic efficiency and extended sugar-donor selectivity. *Planta* 238, 683–693.
- Kim, J.H., Kim, B.G., Park, Y., Ko, J.H., Lim, C.E., Lim, J., Lim, Y., and Ahn, J.-H. (2006a). Characterization of flavonoid 7-O-glucosyltransferase from *Arabidopsis thaliana*. *Biosci. Biotechnol. Biochem.* 70, 1471–1477.
- Kim, J.H., Shin, K.H., Ko, J.H., and Ahn, J.-H. (2006b). Glucosylation of flavonols by *Escherichia coli* expressing glycosyltransferase from rice (*Oryza sativa*). *J. Biosci. Bioeng.* 102, 135–137.
- Kim, Y.-M., Seo, M.-Y., Kang, H.-K., Atsuo, K., and Kim, D. (2009). Construction of a fusion enzyme of dextranucrase and dextranase: Application for one-step synthesis of isomalto-oligosaccharides. *Enzyme Microb. Technol.* 44, 159–164.
- Kim, Y.-M., Kang, H.-K., Moon, Y.-H., Nguyen, T.T.H., Day, D.F., and Kim, D. (2014). 10 Production and Bioactivity of Glucooligosaccharides and Glucosides Synthesized using Glucansucrases. *Food Oligosacch. Prod. Anal. Bioactivity* 168.
- Kimura, Y., Ito, H., Ohnishi, R., and Hatano, T. (2010). Inhibitory effects of polyphenols on human cytochrome P450 3A4 and 2C9 activity. *Food Chem. Toxicol. Int. J. Publ. Br. Ind. Biol. Res. Assoc.* 48, 429–435.
- Kinoshita, T., Lepp, Z., Kawai, Y., Terao, J., and Chuman, H. (2006). An integrated database of flavonoids. *BioFactors* 26, 179–188.
- Kitagawa, S. (2006). Inhibitory effects of polyphenols on p-glycoprotein-mediated transport. *Biol. Pharm. Bull.* 29, 1–6.
- Kitagawa, S., Nabekura, T., Takahashi, T., Nakamura, Y., Sakamoto, H., Tano, H., Hirai, M., and Tsukahara, G. (2005). Structure-Activity Relationships of the Inhibitory Effects of Flavonoids on P-Glycoprotein-Mediated Transport in KB-C2 Cells. *Biol. Pharm. Bull.* 28, 2274–2278.
- Kitagawa, S., Tanaka, Y., Tanaka, M., Endo, K., and Yoshii, A. (2009). Enhanced skin delivery of quercetin by microemulsion. *J. Pharm. Pharmacol.* 61, 855–860.
- Knekt, P., Jarvinen, R., Reunanen, A., and Maatela, J. (1996). Flavonoid intake and coronary mortality in Finland: a cohort study. *BMJ* 312, 478–481.
- Koes, R.E., Quattrocchio, F., and Mol, J.N.M. (1994). The flavonoid biosynthetic pathway in plants: Function and evolution. *BioEssays* 16, 123–132.
- Kometani, T., Terada, Y., Nishimure, T., Takii, H., and Okada, S. (1994). Purification and Characterization of Cyclodextrin Glucanotransferase from an Alkalophilic *Bacillus* Species and Transglycosylation at Alkaline pHs. *Biosci. Biotechnol. Biochem.* 58, 517–520.
- Kometani, T., Nishimura, T., Nakae, T., Takii, H., and Okada, S. (1996). Synthesis of Neohesperidin Glycosides and Naringin Glycosides by Cyclodextrin Glucanotransferase from an Alkalophilic *Bacillus* Species. *Biosci. Biotechnol. Biochem.* 60, 645–649.
- Koo, H., Rosalen, P.L., Cury, J.A., Park, Y.K., and Bowen, W.H. (2002). Effects of Compounds Found in Propolis on *Streptococcus mutans* Growth and on Glucosyltransferase Activity. *Antimicrob. Agents Chemother.* 46, 1302–1309.
- Koshland, D.E. (1953). Stereochemistry and the Mechanism of Enzymatic Reactions. *Biol. Rev.* 28, 416–436.
- Kovinich, N., Saleem, A., Arnason, J.T., and Miki, B. (2010). Functional characterization of a UDP-glucose:flavonoid 3-O-glucosyltransferase from the seed coat of black soybean (*Glycine max* (L.) Merr.). *Phytochemistry* 71, 1253–1263.
- Kralj, S., van Geel-Schutten, G., Van Der Maarel, M., and Dijkhuizen, L. (2004). Biochemical and molecular characterization of *Lactobacillus reuteri* 121 reuteransucrase. *Microbiology* 150, 2099–2112.
- Kralj, S., van Geel-Schutten, G., Faber, E.J., van der Maarel, M.J.E.C., and Dijkhuizen, L. (2005). Rational Transformation of *Lactobacillus reuteri* 121 Reuteransucrase into a Dextranucrase. *Biochemistry (Mosc.)* 44, 9206–9216.
- Kralj, S., van Leeuwen, S.S., Valk, V., Eeuwema, W., Kamerling, J.P., and Dijkhuizen, L. (2008). Hybrid reuteransucrase enzymes reveal regions important for glucosidic linkage specificity and the transglucosylation/hydrolysis ratio: Hybrid reuteransucrases and linkage specificity. *FEBS J.* 275, 6002–6010.
- Kralj, S., Grijpstra, P., van Leeuwen, S.S., Leemhuis, H., Dobruchowska, J.M., van der Kaaij, R.M., Malik, A., Oetari, A., Kamerling, J.P., and Dijkhuizen, L. (2011). 4, 6- $\alpha$ -Glucanotransferase, a novel enzyme that structurally

and functionally provides an evolutionary link between glycoside hydrolase enzyme families 13 and 70. *Appl. Environ. Microbiol.* 77, 8154–8163.

Kren, V., and Martínková, L. (2001). Glycosides in medicine: “The role of glycosidic residue in biological activity.” *Curr. Med. Chem.* 8, 1303–1328.

Křen, V., and Thiem, J. (1997). Glycosylation employing bio-systems: from enzymes to whole cells. *Chem. Soc. Rev.* 26, 463–473.

## L

Lairson, L.L., Henrissat, B., Davies, G.J., and Withers, S.G. (2008). Glycosyltransferases: Structures, Functions, and Mechanisms. *Annu. Rev. Biochem.* 77, 521–555.

Latham, G., Fang, X., Conrad, R., Kemppainen, J., Setterquist, R., and Pasloske, B. (2013). Modified surfaces as solid supports for nucleic acid purification. U.S. Patent Application 13/850,275.

Le Marchand, L. (2002). Cancer preventive effects of flavonoids - A review. *Biomed. Pharmacother.* 56, 296–301.

Lea, M.A., Xiao, Q., Sadhukhan, A.K., Cottle, S., Wang, Z.-Y., and Yang, C.S. (1993). Inhibitory effects of tea extracts and (–)-epigallocatechin gallate on DNA synthesis and proliferation of hepatoma and erythroleukemia cells. *Cancer Lett.* 68, 231–236.

Lee, H.S., and Thorson, J.S. (2011). Development of a universal glycosyltransferase assay amenable to high-throughput formats. *Anal. Biochem.* 418, 85–88.

Lee, E.Y., Choi, S.H., and Kim, H.S. (2010). Nano-immobilization of oleandomycin glycosyltransferase for preparation of glycoflavonoids. *Proc. ICCE-18 Int. Community Compos. Eng. Anchorage Supplement 2*, 256.

Lee, S.J., Son, K.H., Chang, H.W., Do, J.C., Jung, K.Y., Kang, S.S., and Kim, H.P. (1993). Antiinflammatory activity of naturally occurring flavone and flavonol glycosides. *Arch. Pharm. Res.* 16, 25–28.

Lee, S.J., Kim, J.C., Kim, M.J., Kitaoka, M., Park, C.S., Lee, S.Y., Ra, M.J., Moon, T.W., Robyt, J.F., and Park, K.H. (1999). Transglycosylation of naringin by *Bacillus stearothermophilus* Maltogenic amylase to give glycosylated naringin. *J. Agric. Food Chem.* 47, 3669–3674.

Lee, S.S., Hong, S.Y., Errey, J.C., Izumi, A., Davies, G.J., and Davis, B.G. (2011). Mechanistic evidence for a front-side, S<sub>N</sub>i-type reaction in a retaining glycosyltransferase. *Nat. Chem. Biol.* 7, 631–638.

Le-Hilz, Y.Y., Boerboom, A.-M.J.F., Westphal, A.H., van Berkel, W.J.H., Aarts, J.M.M.J.G., and Rietjens, I.M.C.M. (2006). Pro-Oxidant Activity of Flavonoids Induces EpRE-Mediated Gene Expression. *Chem. Res. Toxicol.* 19, 1499–1505.

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

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

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

Lepiniec, L., Debeaujon, I., Routaboul, J.-M., Baudry, A., Pourcel, L., Nesi, N., and Caboche, M. (2006). Genetics and Biochemistry of Seed Flavonoids. *Annu. Rev. Plant Biol.* 57, 405–430.

Lewandowska, U., Szewczyk, K., Hrabec, E., Janecka, A., and Gorlach, S. (2013). Overview of Metabolism and Bioavailability Enhancement of Polyphenols. *J. Agric. Food Chem.* 61, 12183–12199.

Li, B.H., and Tian, W.X. (2004). Inhibitory effects of flavonoids on animal fatty acid synthase. *J. Biochem. (Tokyo)* 135, 85–91.

Li, B.Q., Fu, T., Dongyan, Y., Mikovits, J.A., Ruscetti, F.W., and Wang, J.M. (2000). Flavonoid Baicalin Inhibits HIV-1 Infection at the Level of Viral Entry. *Biochem. Biophys. Res. Commun.* 276, 534–538.

Li, D., Park, S.-H., Shim, J.-H., Lee, H.-S., Tang, S.-Y., Park, C.-S., and Park, K.-H. (2004a). In vitro enzymatic modification of puerarin to puerarin glycosides by maltogenic amylase. *Carbohydr. Res.* 339, 2789–2797.

Li, D., Park, J.-H., Park, J.-T., Park, C.S., and Park, K.-H. (2004b). Biotechnological production of highly soluble daidzein glycosides using *Thermotoga maritima* maltosyltransferase. *J. Agric. Food Chem.* 52, 2561–2567.

- Li, L., Modolo, L.V., Escamilla-Trevino, L.L., Achnine, L., Dixon, R.A., and Wang, X. (2007). Crystal Structure of *Medicago truncatula* UGT85H2 – Insights into the Structural Basis of a Multifunctional (Iso)flavonoid Glycosyltransferase. *J. Mol. Biol.* 370, 951–963.
- Li, X., Li, D., Park, S.-H., Gao, C., Park, K.-H., and Gu, L. (2011). Identification and antioxidative properties of transglycosylated puerarins synthesised by an archaeal maltogenic amylase. *Food Chem.* 124, 603–608.
- Li, Y.Q., Zhou, F.C., Gao, F., Bian, J.S., and Shan, F. (2009). Comparative evaluation of quercetin, isoquercetin and rutin as inhibitors of alpha-glucosidase. *J. Agric. Food Chem.* 57, 11463–11468.
- Lim, E.-K. (2005). Plant Glycosyltransferases: Their Potential as Novel Biocatalysts. *Chem. Eur. J.* 11, 5486–5494.
- Lim, E.-K., Ashford, D.A., Hou, B., Jackson, R.G., and Bowles, D.J. (2004). Arabidopsis glycosyltransferases as biocatalysts in fermentation for regioselective synthesis of diverse quercetin glucosides. *Biotechnol. Bioeng.* 87, 623–631.
- Limem, I., Guedon, E., Hehn, A., Bourgaud, F., Chekir Ghedira, L., Engasser, J.-M., and Ghoul, M. (2008). Production of phenylpropanoid compounds by recombinant microorganisms expressing plant-specific biosynthesis genes. *Process Biochem.* 43, 463–479.
- Lindahl, M., and Tagesson, C. (1997). Flavonoids as phospholipase A2 inhibitors: importance of their structure for selective inhibition of group II phospholipase A2. *Inflammation* 21, 347–356.
- Liu, A.-L., Wang, H.-D., Lee, S.M., Wang, Y.-T., and Du, G.-H. (2008). Structure–activity relationship of flavonoids as influenza virus neuraminidase inhibitors and their in vitro anti-viral activities. *Bioorg. Med. Chem.* 16, 7141–7147.
- Liu, S., Li, D., Huang, B., Chen, Y., Lu, X., and Wang, Y. (2013). Inhibition of pancreatic lipase,  $\alpha$ -glucosidase,  $\alpha$ -amylase, and hypolipidemic effects of the total flavonoids from *Nelumbo nucifera* leaves. *J. Ethnopharmacol.* 149, 263–269.
- Lo Piparo, E., Scheib, H., Frei, N., Williamson, G., Grigorov, M., and Chou, C.J. (2008). Flavonoids for controlling starch digestion: structural requirements for inhibiting human alpha-amylase. *J. Med. Chem.* 51, 3555–3561.
- Loizzo, M.R., Said, A., Tundis, R., Rashed, K., Statti, G.A., Hufner, A., and Menichini, F. (2007). Inhibition of angiotensin converting enzyme (ACE) by flavonoids isolated from *Ailanthus excelsa* (Roxb) (Simaroubaceae). *Phytother. Res. PTR* 21, 32–36.
- Lombard, V., Golaconda Ramulu, H., Drula, E., Coutinho, P.M., and Henrissat, B. (2013). The carbohydrate-active enzymes database (CAZy) in 2013. *Nucleic Acids Res.* 42, D490–D495.
- Lombard, V., Golaconda Ramulu, H., Drula, E., Coutinho, P.M., and Henrissat, B. (2014). The carbohydrate-active enzymes database (CAZy) in 2013. *Nucleic Acids Res.* 42, D490–D495.
- Lorenc-Kukuła, K., Korobczak, A., Aksamit-Stachurska, A., Kostyń, K., Lukaszewicz, M., and Szopa, J. (2004). Glucosyltransferase: the gene arrangement and enzyme function. *Cell. Mol. Biol. Lett.* 9, 935–946.
- Lu, M.-F., Xiao, Z.-T., and Zhang, H.-Y. (2013). Where do health benefits of flavonoids come from? Insights from flavonoid targets and their evolutionary history. *Biochem. Biophys. Res. Commun.* 434, 701–704.

## M

- MacGregor, E.A., Jespersen, H.M., and Svensson, B. (1996). A circularly permuted alpha-amylase-type alpha/beta-barrel structure in glucan-synthesizing glucosyltransferases. *FEBS Lett.* 378, 263–266.
- Mackenzie, L.F., Wang, Q., Warren, R.A.J., and Withers, S.G. (1998). Glycosynthases: Mutant Glycosidases for Oligosaccharide Synthesis. *J. Am. Chem. Soc.* 120, 5583–5584.
- Mackenzie, P.I., Owens, I.S., Burchell, B., Bock, K.W., Bairoch, A., Bélanger, A., Fournel-Gigleux, S., Green, M., Hum, D.W., Iyanagi, T., et al. (1997). The UDP glycosyltransferase gene superfamily: recommended nomenclature update based on evolutionary divergence. *Pharmacogenetics* 7, 255–269.
- Madeswaran, A., Umamaheswari, M., Asokkumar, K., Sivashanmugam, T., Subhadra Devi, V., and Jagannath, P. (2012). In silico docking studies of phosphodiesterase inhibitory activity of commercially available flavonoids. *Orient. Pharm. Exp. Med.* 12, 301–306.
- Maher, P., Akaishi, T., and Abe, K. (2006). Flavonoid fisetin promotes ERK-dependent long-term potentiation and enhances memory. *Proc. Natl. Acad. Sci.* 103, 16568–16573.
- Makino, T., Shimizu, R., Kanemaru, M., Suzuki, Y., Moriwaki, M., and Mizukami, H. (2009). Enzymatically Modified Isoquercitrin,  $\alpha$ -Oligoglucosyl Quercetin 3-O-Glucoside, Is Absorbed More Easily than Other Quercetin Glycosides or Aglycone after Oral Administration in Rats. *Biol. Pharm. Bull.* 32, 2034–2040.

- Maliar, T., Jedinák, A., Kadrabová, J., and Sturdík, E. (2004). Structural aspects of flavonoids as trypsin inhibitors. *Eur. J. Med. Chem.* 39, 241–248.
- Manach, C., Regerat, F., Texier, O., Agullo, G., Demigne, C., and Remesy, C. (1996). Bioavailability, metabolism and physiological impact of 4-oxo-flavonoids. *Nutr. Res.* 16, 517–544.
- Mantas, A., Deretey, E., Ferretti, F.H., Estrada, M.R., and Csizmadia, I.G. (2000). Structural analysis of flavonoids with anti-HIV activity. *J. Mol. Struct. Theochem* 504, 171–179.
- Marais, J.P.J., Deavours, B., Dixon, R.A., and Ferreira, D. (2006). The Stereochemistry of Flavonoids. In *The Science of Flavonoids*, E. Grotewold, ed. (Springer New York), pp. 1–46.
- Marinova, K., Kleinschmidt, K., Weissenböck, G., and Klein, M. (2007). Flavonoid Biosynthesis in Barley Primary Leaves Requires the Presence of the Vacuole and Controls the Activity of Vacuolar Flavonoid Transport. *Plant Physiol.* 144, 432–444.
- Markham, K.R. (1982). *Techniques of flavonoid identification* (Academic Press, London).
- Marti-Mestres, G., Mestres, J.P., Bres, J., Martin, S., Ramos, J., and Vian, L. (2007). The “in vitro” percutaneous penetration of three antioxidant compounds. *Int. J. Pharm.* 331, 139–144.
- De Martino, L., Mencherini, T., Mancini, E., Aquino, R.P., De Almeida, L.F.R., and De Feo, V. (2012). In Vitro Phytotoxicity and Antioxidant Activity of Selected Flavonoids. *Int. J. Mol. Sci.* 13, 5406–5419.
- Masada, S., Kawase, Y., Nagatoshi, M., Oguchi, Y., Terasaka, K., and Mizukami, H. (2007). An efficient chemoenzymatic production of small molecule glucosides with in situ UDP-glucose recycling. *FEBS Lett.* 581, 2562–2566.
- Matsubara, K., Ishihara, K., Mizushina, Y., Mori, M., and Nakajima, N. (2004). Anti-Angiogenic Activity of Quercetin and its Derivatives. *Lett. Drug Des. Discov.* 1, 329–333.
- Matsuda, H., Morikawa, T., Toguchida, I., and Yoshikawa, M. (2002). Structural requirements of flavonoids and related compounds for aldose reductase inhibitory activity. *Chem. Pharm. Bull. (Tokyo)* 50, 788–795.
- Matsuda, H., Morikawa, T., Ando, S., Toguchida, I., and Yoshikawa, M. (2003). Structural requirements of flavonoids for nitric oxide production inhibitory activity and mechanism of action. *Bioorg. Med. Chem.* 11, 1995–2000.
- Matsukawa, N., Matsumoto, M., Shinoki, A., Hagio, M., Inoue, R., and Hara, H. (2009). Nondigestible Saccharides Suppress the Bacterial Degradation of Quercetin Aglycone in the Large Intestine and Enhance the Bioavailability of Quercetin Glucoside in Rats. *J. Agric. Food Chem.* 57, 9462–9468.
- Mattos-Graner, R., Smith, D., King, W., and Mayer, M. (2000). Water-insoluble glucan synthesis by mutans streptococcal strains correlates with caries incidence in 12-to 30-month-old children. *J. Dent. Res.* 79, 1371–1377.
- Mercader, A.G., Duchowicz, P.R., Fernández, F.M., Castro, E.A., Bennardi, D.O., Autino, J.C., and Romanelli, G.P. (2008). QSAR prediction of inhibition of aldose reductase for flavonoids. *Bioorg. Med. Chem.* 16, 7470–7476.
- Meulenbeld, G.H. (2001). *Enzymatic glucosylation: sucrose glucosyltransferases and glucosidases in O- and S-glucoside synthesis*. Thesis, Wageningen University, The Netherlands.
- Meulenbeld, G.H., Zuilhof, H., van Veldhuizen, A., van den Heuvel, R.H., and Hartmans, S. (1999). Enhanced (+)-catechin transglucosylating activity of *Streptococcus mutans* GS-5 glucosyltransferase-D due to fructose removal. *Appl. Environ. Microbiol.* 65, 4141–4147.
- Miller, A. (1996). Antioxidant Flavonoids: Structure, Function and Clinical Usage. *Altern. Med. Rev.* 1, 103–111.
- Minic, Z. (2008). Physiological roles of plant glycoside hydrolases. *Planta* 227, 723–740.
- Minic, Z., and Jouanin, L. (2006). Plant glycoside hydrolases involved in cell wall polysaccharide degradation. *Plant Physiol. Biochem.* 44, 435–449.
- Moco, S., Tseng, L.-H., Spraul, M., Chen, Z., and Vervoort, J. (2006). Building-Up a Comprehensive Database of Flavonoids Based on Nuclear Magnetic Resonance Data. *Chromatographia* 64, 503–508.
- Modolo, L.V., Li, L., Pan, H., Blount, J.W., Dixon, R.A., and Wang, X. (2009a). Crystal Structures of Glycosyltransferase UGT78G1 Reveal the Molecular Basis for Glycosylation and Deglycosylation of (Iso)flavonoids. *J. Mol. Biol.* 392, 1292–1302.
- Modolo, L.V., Escamilla-Treviño, L.L., Dixon, R.A., and Wang, X. (2009b). Single amino acid mutations of Medicago glycosyltransferase UGT85H2 enhance activity and impart reversibility. *FEBS Lett.* 583, 2131–2135.
- Monchois, V. (1999). Glucansucrases: mechanism of action and structure–function relationships. *FEMS Microbiol. Rev.* 23, 131–151.

- Monchois, V., Willemot, R.M., Remaud-Simeon, M., Croux, C., and Monsan, P. (1996). Cloning and sequencing of a gene coding for a novel dextransucrase from *Leuconostoc mesenteroides* NRRL B-1299 synthesizing only alpha (1-6) and alpha (1-3) linkages. *Gene* 182, 23–32.
- Monchois, V., Remaud-Simeon, M., Russell, R.R., Monsan, P., and Willemot, R.M. (1997). Characterization of *Leuconostoc mesenteroides* NRRL B-512F dextransucrase (DSRS) and identification of amino-acid residues playing a key role in enzyme activity. *Appl. Microbiol. Biotechnol.* 48, 465–472.
- Monchois, V., Vignon, M., Escalier, P.C., Svensson, B., and Russell, R.R. (2000a). Involvement of Gln937 of *Streptococcus downei* GTF-I glucansucrase in transition-state stabilization. *Eur. J. Biochem. FEBS* 267, 4127–4136.
- Monchois, V., Vignon, M., and Russell, R.R.B. (2000b). Mutagenesis of Asp-569 of Glucosyltransferase I Glucansucrase Modulates Glucan and Oligosaccharide Synthesis. *Appl. Environ. Microbiol.* 66, 1923–1927.
- Monsan, P., Potocki de Montalk, G., Sarçabal, P., Remaud-Siméon, M., and Willemot, R. (2000). Glucansucrases: efficient tools for the synthesis of oligosaccharides of nutritional interest. *Prog. Biotechnol.* 17, 115–122.
- Monsan, P., Remaud-Simeon, M., and Joucla, G. (2009). Fully active alternansucrases partially deleted in its carboxy-terminal and amino-terminal domains and mutants thereof. U.S. Patent Application. 7,524,645.
- Monsan, P., Remaud-Siméon, M., and André, I. (2010a). Transglucosidases as efficient tools for oligosaccharide and glucoconjugate synthesis. *Curr. Opin. Microbiol.* 13, 293–300.
- Monsan, P., Remaud-Siméon, M., and André, I. (2010b). Transglucosidases as efficient tools for oligosaccharide and glucoconjugate synthesis. *Curr. Opin. Microbiol.* 13, 293–300.
- Moon, Y.-H., Lee, J.-H., Ahn, J.-S., Nam, S.-H., Oh, D.-K., Park, D.-H., Chung, H.-J., Kang, S., Day, D.F., and Kim, D. (2006a). Synthesis, structure analyses, and characterization of novel epigallocatechin gallate (EGCG) glycosides using the glucansucrase from *Leuconostoc mesenteroides* B-1299CB. *J. Agric. Food Chem.* 54, 1230–1237.
- Moon, Y.-H., Kim, G., Lee, J.-H., Jin, X.-J., Kim, D.-W., and Kim, D. (2006b). Enzymatic synthesis and characterization of novel epigallocatechin gallate glucosides. *J. Mol. Catal. B Enzym.* 40, 1–7.
- Moon, Y.-H., Lee, J.-H., Jhon, D.-Y., Jun, W.-J., Kang, S.-S., Sim, J., Choi, H., Moon, J.-H., and Kim, D. (2007). Synthesis and characterization of novel quercetin-[alpha]-d-glucopyranosides using glucansucrase from *Leuconostoc mesenteroides*. *Enzyme Microb. Technol.* 40, 1124–1129.
- Mooser, G. (1992). Glycosidases and Glycosyltransferases. *The enzymes*, 20, 187–233.
- Mooser, G., Hefta, S., Paxton, R., Shively, J., and Lee, T. (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.
- Moriwaki, M., Emura, K., and Tanaka, H. (2010). Method for manufacturing  $\alpha$ -glycosylisoquercitrin, intermediate product and by-product thereof. U.S. Patent Application. 7,691,425.
- Mostrag-Szlichtyng, A., and Worth, A. (2010). Review of QSAR models and software tools for predicting biokinetic properties. Luxembourg: European Commission.
- Moulis, C., Joucla, G., Harrison, D., Fabre, E., Potocki-Veronese, G., Monsan, P., and Remaud-Simeon, M. (2006). Understanding the polymerization mechanism of glycoside-hydrolase family 70 glucansucrases. *J. Biol. Chem.* 281, 31254–31267.
- Mulvihill, E.E., and Huff, M.W. (2010). Antiatherogenic properties of flavonoids: Implications for cardiovascular health. *Can. J. Cardiol.* 26, Supplement A, 17A–21A.
- Murota, K., Matsuda, N., Kashino, Y., Fujikura, Y., Nakamura, T., Kato, Y., Shimizu, R., Okuyama, S., Tanaka, H., Koda, T., et al. (2010).  $\alpha$ -Oligoglucosylation of a sugar moiety enhances the bioavailability of quercetin glucosides in humans. *Arch. Biochem. Biophys.* 501, 91–97.

## N

- Nakahara, K., Kontani, M., Ono, H., Kodama, T., Tanaka, T., Ooshima, T., and Hamada, S. (1995). Glucosyltransferase from *Streptococcus sobrinus* Catalyzes Glucosylation of Catechin. *Appl. Environ. Microbiol.* 61, 2768–2770.
- Nakamura, A., Haga, K., and Yamane, K. (1993). Three histidine residues in the active center of cyclodextrin glucanotransferase from alkalophilic *Bacillus* sp. 1011: effects of the replacement on pH dependence and transition-state stabilization. *Biochemistry (Mosc.)* 32, 6624–6631.

Nakatani, H. (2001). Analysis of Glycosidase-Catalyzed Transglycosylation Reaction Using Probabilistic Model. *Arch. Biochem. Biophys.* 385, 387–391.

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

Naoumkina, M., and Dixon, R.A. Subcellular localization of flavonoid natural products: A signaling function? *Plant Signal. Behav.* 3, 573–575.

Németh, K., Plumb, G.W., Berrin, J.-G., Juge, N., Jacob, R., Naim, H.Y., Williamson, G., Swallow, D.M., and Kroon, P.A. (2003). Deglycosylation by small intestinal epithelial cell  $\beta$ -glucosidases is a critical step in the absorption and metabolism of dietary flavonoid glycosides in humans. *Eur. J. Nutr.* 42, 29–42.

Nerland, D.E. (2007). The Antioxidant/electrophile Response Element Motif. *Drug Metab. Rev.* 39, 235–248.

Nielsen, I.L.F., Chee, W.S.S., Poulsen, L., Offord-Cavin, E., Rasmussen, S.E., Frederiksen, H., Enslen, M., Barron, D., Horcajada, M.-N., and Williamson, G. (2006). Bioavailability is improved by enzymatic modification of the citrus flavonoid hesperidin in humans: a randomized, double-blind, crossover trial. *J. Nutr.* 136, 404–408.

Nijveldt, R.J., Nood, E. van, Hoorn, D.E. van, Boelens, P.G., Norren, K. van, and Leeuwen, P.A. van (2001). Flavonoids: a review of probable mechanisms of action and potential applications. *Am. J. Clin. Nutr.* 74, 418–425.

Nishimura, T., Kometani, T., Takii, H., Terada, Y., and Okada, S. (1994). Acceptor specificity in the glucosylation reaction of *Bacillus subtilis* X-23  $\alpha$ -amylase towards various phenolic compounds and the structure of kojic acid glucoside. *J. Ferment. Bioeng.* 78, 37–41.

Noguchi, A., Sasaki, N., Nakao, M., Fukami, H., Takahashi, S., Nishino, T., and Nakayama, T. (2008). cDNA cloning of glycosyltransferases from Chinese wolfberry (*Lycium barbarum* L.) fruits and enzymatic synthesis of a catechin glucoside using a recombinant enzyme (UGT73A10). *J. Mol. Catal. B Enzym.* 55, 84–92.

## O

Ochiai, M., Fukami, H., Nakao, M., and Noguchi, A. (2010). Method for glycosylation of flavonoid compounds. U.S. Patent Application 12/523,065.

Offen, W., Martinez-Fleites, C., Yang, M., Kiat-Lim, E., Davis, B.G., Tarling, C.A., Ford, C.M., Bowles, D.J., and Davies, G.J. (2006). Structure of a flavonoid glucosyltransferase reveals the basis for plant natural product modification. *EMBO J.* 25, 1396–1405.

Oh, H., Kim, D.-H., Cho, J.-H., and Kim, Y.-C. (2004). Hepatoprotective and free radical scavenging activities of phenolic petrosins and flavonoids isolated from *Equisetum arvense*. *J. Ethnopharmacol.* 95, 421–424.

Okuda, J., Miwa, I., Inagaki, K., Horie, T., and Nakayama, M. (1984). Inhibition of aldose reductase by 3',4'-dihydroxyflavones. *Chem. Pharm. Bull. (Tokyo)* 32, 767–772.

Olszanecki, R., Gebska, A., Kozlovski, V.I., and Gryglewski, R.J. (2002). Flavonoids and nitric oxide synthase. *J. Physiol. Pharmacol. Off. J. Pol. Physiol. Soc.* 53, 571–584.

Ono, K., Nakane, H., Fukushima, M., Chermann, J.-C., and Barré-Sinoussi, F. (1990). Differential inhibitory effects of various flavonoids on the activities of reverse transcriptase and cellular DNA and RNA polymerases. *Eur. J. Biochem.* 190, 469–476.

Ono, Y., Tomimori, N., Tateishi, N., Moriwaki, M., Emura, K., and Okuyama, S. (2009). Quercetin Glycoside Composition and Method of Preparing the Same. U.S. Patent Application 11/794,218.

Osmani, S.A., Bak, S., Imbert, A., Olsen, C.E., and Møller, B.L. (2008). Catalytic Key Amino Acids and UDP-Sugar Donor Specificity of a Plant Glucuronosyltransferase, UGT94B1: Molecular Modeling Substantiated by Site-Specific Mutagenesis and Biochemical Analyses. *Plant Physiol.* 148, 1295–1308.

Osmani, S.A., Bak, S., and Møller, B.L. (2009). Substrate specificity of plant UDP-dependent glycosyltransferases predicted from crystal structures and homology modeling. *Phytochemistry* 70, 325–347.

## P

Paliwal, S., Sundaram, J., and Mitragotri, S. (2005). Induction of cancer-specific cytotoxicity towards human prostate and skin cells using quercetin and ultrasound. *Br. J. Cancer* 92, 499–502.

Pandey, R.P., Malla, S., Simkhada, D., Kim, B.-G., and Sohng, J.K. (2013). Production of 3-O-xylosyl quercetin in *Escherichia coli*. *Appl. Microbiol. Biotechnol.* 97, 1889–1901.

Parichatkanond, W., Pinthong, D., and Mangmool, S. (2012). Blockade of the renin-angiotensin system with delphinidin, cyanin, and quercetin. *Planta Med.* 78, 1626–1632.



- Park, K.-H. (2008). *Carbohydrate-Active Enzymes: Structure, Function and Applications* (Woodhead Publishing Ltd, Cambridge).
- Park, S.-H., Park, H.-Y., Sohng, J.K., Lee, H.C., Liou, K., Yoon, Y.J., and Kim, B.-G. (2009). Expanding substrate specificity of GT-B fold glycosyltransferase via domain swapping and high-throughput screening. *Biotechnol. Bioeng.* 102, 988–994.
- Peng, B., Zi, J., and Yan, W. (2006). Measurement and correlation of solubilities of luteolin in organic solvents at different temperatures. *J. Chem. Eng. Data* 51, 2038–2040.
- Perez-Cenci, M., and Salerno, G.L. (2014). Functional characterization of *Synechococcus* amylosucrase and fructokinase encoding genes discovers two novel actors on the stage of cyanobacterial sucrose metabolism. *Plant Sci.* 224, 95–102.
- Perez-Vizcaino, F., Duarte, J., and Santos-Buelga, C. (2012). The flavonoid paradox: conjugation and deconjugation as key steps for the biological activity of flavonoids. *J. Sci. Food Agric.* 92, 1822–1825.
- Phosrithong, N., Samee, W., and Ungwitayatorn, J. (2012). 3D-QSAR studies of natural flavonoid compounds as reverse transcriptase inhibitors. *Med. Chem. Res.* 21, 559–567.
- Phuwamongkolwivat, P., Suzuki, T., Hira, T., and Hara, H. Fructooligosaccharide augments benefits of quercetin-3-O- $\beta$ -glucoside on insulin sensitivity and plasma total cholesterol with promotion of flavonoid absorption in sucrose-fed rats. *Eur. J. Nutr.* 1–12.
- Piazza, C., Privitera, M.G., Melilli, B., Incognito, T., Marano, M.R., Leggio, G.M., Roxas, M.A., and Drago, F. (2007). Influence of inulin on plasma isoflavone concentrations in healthy postmenopausal women. *Am. J. Clin. Nutr.* 86, 775–780.
- Pietta, P.-G. (2000). Flavonoids as Antioxidants. *J. Nat. Prod.* 63, 1035–1042.
- Pijning, T., Vujicic-Zagar, A., Kralj, S., Dijkhuizen, L., and Dijkstra, B.W. (2012). Structure of the-1, 6/-1, 4-specific glucansucrase GTFA from *Lactobacillus reuteri* 121. *Acta Crystallograph. Sect. F Struct. Biol. Cryst. Commun.* 68, 1448–1454.
- Pijning, T., Vujičić-Zagar, A., Kralj, S., Dijkhuizen, L., and Dijkstra, B.W. (2014). Flexibility of truncated and full-length glucansucrase GTF180 enzymes from *Lactobacillus reuteri* 180. *FEBS J.*, 281(9), 2159–2171.
- 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.
- Plaza, M., Pozzo, T., Liu, J., Gulshan Ara, K.Z., Turner, C., and Nordberg Karlsson, E. (2014). Substituent Effects on in Vitro Antioxidizing Properties, Stability, and Solubility in Flavonoids. *J. Agric. Food Chem.* 62, 3321–3333.
- Plochmann, K., Korte, G., Koutsilieri, E., Richling, E., Riederer, P., Rethwilm, A., Schreier, P., and Scheller, C. (2007). Structure-activity relationships of flavonoid-induced cytotoxicity on human leukemia cells. *Arch. Biochem. Biophys.* 460, 1–9.
- 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.
- Potocki De Montalk, G., 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.
- 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.
- 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.
- Pozzo, T. (2012). *Thermostable glycosidases and glycosynthases as biocatalysts in green chemistry applications*. Doctoral dissertation. Lund University.
- Prasath, G.S., and Subramanian, S.P. (2013). Fisetin, a tetra hydroxy flavone recuperates antioxidant status and protects hepatocellular ultrastructure from hyperglycemia mediated oxidative stress in streptozotocin induced experimental diabetes in rats. *Food Chem. Toxicol.* 59, 249–255.
- Procházková, D., Boušová, I., and Wilhelmová, N. (2011). Antioxidant and prooxidant properties of flavonoids. *Fitoterapia* 82, 513–523.
- Pulley, G.N. (1936). Solubility of Naringin in Water. *Ind. Eng. Chem. Anal. Ed.* 8, 360–360.

## Q

Quideau, S., Deffieux, D., Douat-Casassus, C., and Pouységu, L. (2011). Plant Polyphenols: Chemical Properties, Biological Activities, and Synthesis. *Angew. Chem. Int. Ed.* 50, 586–621.

## R

Raab, T., Barron, D., Vera, F.A., Crespy, V., Oliveira, M., and Williamson, G. (2010). Catechin Glucosides: Occurrence, Synthesis, and Stability. *J. Agric. Food Chem.* 58, 2138–2149.

Raj, N., Sripal, R., Chaluvadi, M., and Krishna, R. (2001). Bioflavonoids classification, pharmacological, biochemical effects and therapeutic potential. *Indian J. Pharmacol.* 33, 2.

Ramos, S. (2007). Effects of dietary flavonoids on apoptotic pathways related to cancer chemoprevention. *J. Nutr. Biochem.* 18, 427–442.

Rang, H., Dale, M., Ritter, J., and Flower, R. (2007). *Rand and Dale's Pharmacology* 6<sup>th</sup> edition (Churchill Livingstone, Edinburgh)

Ravishankar, D., Rajora, A.K., Greco, F., and Osborn, H.M.I. (2013). Flavonoids as prospective compounds for anti-cancer therapy. *Int. J. Biochem. Cell Biol.* 45, 2821–2831.

Redrejo-Rodriguez, M., Tejeda-Cano, A., del Carmen Pinto, M., and Macías, P. (2004). Lipoygenase inhibition by flavonoids: semiempirical study of the structure–activity relation. *J. Mol. Struct. THEOCHEM* 674, 121–124.

Remaud-Simeon, M., Albaret, F., Canard, B., Varlet, I., Colonna, P., Willemot, R.M., and Monsan, P. (1995). Studies on a recombinant amylosucrase. In *Progress in Biotechnology*, (Elsevier), pp. 313–320.

Remaud-Simeon, M., Willemot, R.-M., Sarçabal, P., Potocki de Montalk, G., and Monsan, P. (2000). Glucansucrases: molecular engineering and oligosaccharide synthesis. *J. Mol. Catal. B Enzym.* 10, 117–128.

Ren, G., Hou, J., Fang, Q., Sun, H., Liu, X., Zhang, L., and Wang, P.G. (2012). Synthesis of flavonol 3-O-glycoside by UGT78D1. *Glycoconj. J.* 29, 425–432.

Ribeiro, D., Freitas, M., Tomé, S.M., Silva, A.M.S., Porto, G., Cabrita, E.J., Marques, M.M.B., and Fernandes, E. (2014). Inhibition of LOX by flavonoids: a structure–activity relationship study. *Eur. J. Med. Chem.* 72, 137–145.

Rice-Evans, C. (2001). Flavonoid antioxidants. *Curr. Med. Chem.* 8, 797–807.

Rice-Evans, C.A., Miller, N.J., and Paganga, G. (1996). Structure-antioxidant activity relationships of flavonoids and phenolic acids. *Free Radic. Biol. Med.* 20, 933–956.

Del Rio, D., Rodriguez-Mateos, A., Spencer, J.P.E., Tognolini, M., Borges, G., and Crozier, A. (2013). Dietary (Poly)phenolics in Human Health: Structures, Bioavailability, and Evidence of Protective Effects Against Chronic Diseases. *Antioxid. Redox Signal.* 18, 1818–1892.

Roberts, S.C. (2007). Production and engineering of terpenoids in plant cell culture. *Nat. Chem. Biol.* 3, 387–395.

Ross, J., Li, Y., Lim, E.-K., and Bowles, D.J. (2001). Higher plant glycosyltransferases. *Genome Biol.* 2, reviews3004.1–reviews3004.6.

Routaboul, J.-M., Kerhoas, L., Debeaujon, I., Pourcel, L., Caboche, M., Einhorn, J., and Lepiniec, L. (2006). Flavonoid diversity and biosynthesis in seed of *Arabidopsis thaliana*. *Planta* 224, 96–107.

Rubbo, H., Radi, R., Trujillo, M., Telleri, R., Kalyanaraman, B., Barnes, S., Kirk, M., and Freeman, B.A. (1994). Nitric oxide regulation of superoxide and peroxynitrite-dependent lipid peroxidation. Formation of novel nitrogen-containing oxidized lipid derivatives. *J. Biol. Chem.* 269, 26066–26075.

Russo, M., Spagnuolo, C., Tedesco, I., Bilotto, S., and Russo, G.L. (2012). The flavonoid quercetin in disease prevention and therapy: Facts and fancies. *Biochem. Pharmacol.* 83, 6–15.

Rusznýák, S., and Szent-Györgyi, A. (1936). Vitamin P: Flavonols as Vitamins. *Nature* 138, 27–27.

Rye, C.S., and Withers, S.G. (2000). Glycosidase mechanisms. *Curr. Opin. Chem. Biol.* 4, 573–580.

Ryu, H.-J., Jin, X., Lee, J.-H., Woo, H.-J., Kim, Y.-M., Kim, G.J., Seo, E.-S., Kang, H.-K., Kim, J., Cho, D.-L., et al. (2010). Optimal expression and characterization of a fusion enzyme having dextransucrase and dextranase activities. *Enzyme Microb. Technol.* 47, 212–215.

## S

- Sadik, C.D., Sies, H., and Schewe, T. (2003). Inhibition of 15-lipoxygenases by flavonoids: structure–activity relations and mode of action. *Biochem. Pharmacol.* **65**, 773–781.
- Saidman, E., Yurquina, A., Rudyk, R., Molina, M.A.A., and Ferretti, F.H. (2002). A theoretical and experimental study on the solubility, dissolution rate, structure and dipolar moment of flavone in ethanol. *J. Mol. Struct. THEOCHEM* **585**, 1–13.
- Saito, K., Yonekura-Sakakibara, K., Nakabayashi, R., Higashi, Y., Yamazaki, M., Tohge, T., and Fernie, A.R. (2013). The flavonoid biosynthetic pathway in Arabidopsis: Structural and genetic diversity. *Plant Physiol. Biochem.* **72**, 21–34.
- Samuelsson, G., Bohlin, L. (2009). *Drugs of natural origin: a treatise of pharmacognosy*, 6th revised edition (Stockholm, Sweden: Apotekarsocieteten).
- Sarbini, S.R., Kolida, S., Naeye, T., Einerhand, A., Brison, Y., Remaud-Simeon, M., Monsan, P., Gibson, G.R., and Rastall, R.A. (2011). In vitro fermentation of linear and {alpha}-1,2 branched dextrans by the human faecal microbiota. *Appl. Environ. Microbiol.*, **77** (15), 5307–5315.
- Sarbini, S.R., Kolida, S., Naeye, T., Einerhand, A.W., Gibson, G.R., and Rastall, R.A. (2013). The prebiotic effect of  $\alpha$ -1, 2 branched, low molecular weight dextran in the batch and continuous faecal fermentation system. *J. Funct. Foods* **5**, 1938–1946.
- Sarçabal, P., Remaud-Simeon, M., Willemot, R.-M., 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.
- Saslowsky, D.E., Warek, U., and Winkel, B.S.J. (2005). Nuclear Localization of Flavonoid Enzymes in Arabidopsis. *J. Biol. Chem.* **280**, 23735–23740.
- Sato, T., Nakagawa, H., Kurosu, J., Yoshida, K., Tsugane, T., Shimura, S., Kirimura, K., Kino, K., and Usami, S. (2000).  $\alpha$ -Anomer-selective glucosylation of (+)-catechin by the crude enzyme, showing glucosyl transfer activity, of *Xanthomonas campestris* WU-9701. *J. Biosci. Bioeng.* **90**, 625–630.
- Scheepens, A., Tan, K., and Paxton, J.W. (2010). Improving the oral bioavailability of beneficial polyphenols through designed synergies. *Genes Nutr.* **5**, 75–87.
- 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.
- Seo, D.-H. (2012). Functional Expression of Amylosucrase, a Glucan-Synthesizing Enzyme, from *Arthrobacter chlorophenolicus* A6. *J. Microbiol. Biotechnol.* **22**, 1253–1257.
- Shah, D.S.H., Joucla, G., Remaud-Simeon, M., and Russell, R.R.B. (2004). Conserved Repeat Motifs and Glucan Binding by Glucansucrases of Oral Streptococci and *Leuconostoc mesenteroides*. *J. Bacteriol.* **186**, 8301–8308.
- Shao, H., He, X., Achnine, L., Blount, J.W., Dixon, R.A., and Wang, X. (2005). Crystal Structures of a Multifunctional Triterpene/Flavonoid Glycosyltransferase from *Medicago truncatula*. *Plant Cell Online* **17**, 3141–3154.
- Shimoda, K., Hamada, H., and Hamada, H. (2011). Synthesis of xylooligosaccharides of daidzein and their anti-oxidant and anti-allergic activities. *Int. J. Mol. Sci.* **12**, 5616–5625.
- Shinoki, A., Lang, W., Thawornkuno, C., Kang, H.-K., Kumagai, Y., Okuyama, M., Mori, H., Kimura, A., Ishizuka, S., and Hara, H. (2013). A novel mechanism for the promotion of quercetin glycoside absorption by megalo  $\alpha$ -1,6-glucosaccharide in the rat small intestine. *Food Chem.* **136**, 293–296.
- Simkhada, D., Kim, E., Lee, H.C., and Sohng, J.K. (2009). Metabolic engineering of *Escherichia coli* for the biological synthesis of 7-O-xylosyl naringenin. *Mol. Cells* **28**, 397–401.
- Simkhada, D., Lee, H.C., and Sohng, J.K. (2010a). Genetic engineering approach for the production of rhamnosyl and allosyl flavonoids from *Escherichia coli*. *Biotechnol. Bioeng.* **107**, 154–162.
- Simkhada, D., Kurumbang, N.P., Lee, H.C., and Sohng, J.K. (2010b). Exploration of glycosylated flavonoids from metabolically engineered *E. coli*. *Biotechnol. Bioprocess Eng.* **15**, 754–760.
- Singh, S., Saini, S., Verma, B., and Mishra, D. (2009). Quantitative Structure Pharmacokinetic Relationship Using Artificial Neural Network: A Review. *Int. J. Pharm. Sci. Drug Res.* **1**, 144–153.
- 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.

- 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. *J. Biol. Chem.* 277, 47741–47747.
- Smith, G.J., Thomsen, S.J., Markham, K.R., Andary, C., and Cardon, D. (2000). The photostabilities of naturally occurring 5-hydroxyflavones, flavonols, their glycosides and their aluminium complexes. *J. Photochem. Photobiol. Chem.* 136, 87–91.
- Snyder, R.D., and Gillies, P.J. (2002). Evaluation of the clastogenic, DNA intercalative, and topoisomerase II-interactive properties of bioflavonoids in Chinese hamster V79 cells. *Environ. Mol. Mutagen.* 40, 266–276.
- Sousa, Â., Sousa, F., and Queiroz, J.A. (2012). Advances in chromatographic supports for pharmaceutical-grade plasmid DNA purification. *J. Sep. Sci.* 35, 3046–3058.
- Stafford, H. (1974). Metabolism of Aromatic-Compounds. *Annu. Rev. Plant Physiol. Plant Mol. Biol.* 25, 459–486.
- Stam, M.R., Danchin, E.G., 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.
- Steffen, Y., Gruber, C., Schewe, T., and Sies, H. (2008). Mono-O-methylated flavanols and other flavonoids as inhibitors of endothelial NADPH oxidase. *Arch. Biochem. Biophys.* 469, 209–219.
- Suzuki, Y., and Suzuki, K. (1991). Enzymatic formation of 4G- $\alpha$ -D-glucopyranosyl-rutin. *Agric. Biol. Chem.* 55, 181–187.
- Svensson, B. (1994). Protein engineering in the  $\alpha$ -amylase family: catalytic mechanism, substrate specificity, and stability. *Plant Mol. Biol.* 25, 141–157.
- Szent-Györgyi, A. (1963). Lost in the twentieth century. *Annu. Rev. Biochem.* 32, 1–14.

## I

- Tan, L., Su, J., Wu, D., Yu, X., Su, Z., He, J., Wu, X., Kong, S., Lai, X., Lin, J., et al. (2013). Kinetics and Mechanism Study of Competitive Inhibition of Jack-Bean Urease by Baicalin. *Sci. World J.* 2013.
- Tanaka, T. (2013). Flavonoids For Allergic Diseases: Present Evidence And Future Perspective. *Curr. Pharm. Des.* 20 (6), 879-85.
- Tanaka, Y., Yonekura, K., Fukuchi-Mizutani, M., Fukui, Y., Fujiwara, H., Ashikari, T., and Kusumi, T. (1996). Molecular and biochemical characterization of three anthocyanin synthetic enzymes from *Gentiana triflora*. *Plant Cell Physiol.* 37, 711–716.
- Tanaka, Y., Sasaki, N., and Ohmiya, A. (2008). Biosynthesis of plant pigments: anthocyanins, betalains and carotenoids. *Plant J.* 54, 733–749.
- Tapas, A.R., Sakarkar, D.M., and Kakde, R.B. (2008). Flavonoids as Nutraceuticals: A Review. *Trop. J. Pharm. Res.* 7, 1089–1099.
- Tasdemir, D., Lack, G., Brun, R., Rüedi, P., Scapozza, L., and Perozzo, R. (2006). Inhibition of *Plasmodium falciparum* fatty acid biosynthesis: evaluation of FabG, FabZ, and FabI as drug targets for flavonoids. *J. Med. Chem.* 49, 3345–3353.
- Terwisscha van Scheltinga, A.C., Armand, S., Kalk, K.H., Isogai, A., Henrissat, B., and Dijkstra, B.W. (1995). Stereochemistry of chitin hydrolysis by a plant chitinase/lysozyme and X-ray structure of a complex with allosamidin: evidence for substrate assisted catalysis. *Biochemistry (Mosc.)* 34, 15619–15623.
- Thorson, J., Williams, G., and Gantt, R. (2009). Engineered Glycosyltransferases with Expanded Substrate Specificity. U.S. Patent Application 8,093,028.
- Thuan, N.H., Park, J.W., and Sohng, J.K. (2013a). Toward the production of flavone-7-O- $\beta$ -D-glucopyranosides using *Arabidopsis* glycosyltransferase in *Escherichia coli*. *Process Biochem.* 48, 1744–1748.
- Thuan, N.H., Pandey, R.P., Thuy, T.T.T., Park, J.W., and Sohng, J.K. (2013b). Improvement of Regio-Specific Production of Myricetin-3-O- $\alpha$ -L-Rhamnoside in Engineered *Escherichia coli*. *Appl. Biochem. Biotechnol.* 171, 1956–1967.
- Tian, W.-X. (2006). Inhibition of Fatty Acid Synthase by Polyphenols. *Curr. Med. Chem.* 13, 967–977.
- Tohge, T., Yonekura-Sakakibara, K., Niida, R., Watanabe-Takahashi, A., and Saito, K. (2007). Phytochemical genomics in *Arabidopsis thaliana*: A case study for functional identification of flavonoid biosynthesis genes. *Pure Appl. Chem.* 79, 811–823.
- Tripoli, E., Guardia, M.L., Giammanco, S., Majo, D.D., and Giammanco, M. (2007). Citrus flavonoids: Molecular structure, biological activity and nutritional properties: A review. *Food Chem.* 104, 466–479.

**U**

Ungar, Y., Osundahunsi, O.F., and Shimoni, E. (2003). Thermal Stability of Genistein and Daidzein and Its Effect on Their Antioxidant Activity. *J. Agric. Food Chem.* 51, 4394–4399.

Uriarte-Pueyo, I., and I. Calvo, M. (2011). Flavonoids as Acetylcholinesterase Inhibitors. *Curr. Med. Chem.* 18, 5289–5302.

**V**

Vasella, A., Davies, G.J., and Böhm, M. (2002). Glycosidase mechanisms. *Curr. Opin. Chem. Biol.* 6, 619–629.

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.

Van der Veen, B.A., Skov, L.K., Potocki-Véronèse, G., Gajhede, M., Monsan, P., and Remaud-Simeon, M. (2006). Increased amylosucrase activity and specificity, and identification of regions important for activity, specificity and stability through molecular evolution. *FEBS J.* 273, 673–681.

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

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

Van Leeuwen, S.S., Kralj, S., Eeuwema, W., Gerwig, G.J., Dijkhuizen, L., and Kamerling, J.P. (2009). Structural Characterization of Bioengineered  $\alpha$ -D-Glucans Produced by Mutant Glucansucrase GTF180 Enzymes of *Lactobacillus reuteri* Strain 180. *Biomacromolecules* 10, 580–588.

Vessal, M., Hemmati, M., and Vasei, M. (2003). Antidiabetic effects of quercetin in streptozocin-induced diabetic rats. *Comp. Biochem. Physiol. Part C Toxicol. Pharmacol.* 135, 357–364.

Vijayaraghavan, R., Sugendran, K., Pant, S.C., Husain, K., and Malhotra, R.C. (1991). Dermal intoxication of mice with bis(2-chloroethyl)sulphide and the protective effect of flavonoids. *Toxicology* 69, 35–42.

Vocadlo, D.J., and Withers, S.G. (2000). Glycosidase-Catalysed Oligosaccharide Synthesis. In *Carbohydrates in Chemistry and Biology*, B. Ernst, G.W. Hart, and P. Sina, eds. (Weinheim, Germany: Wiley-VCH Verlag GmbH), 724–844.

Vogt, T. (2002). Substrate specificity and sequence analysis define a polyphyletic origin of betanidin 5- and 6-O-glucosyltransferase from *Dortheanthus bellidiformis*. *Planta* 214, 492–495.

Vogt, T., and Jones, P. (2000). Glycosyltransferases in plant natural product synthesis: characterization of a supergene family. *Trends Plant Sci.* 5, 380–386.

Vujičić-Žagar, 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.* 107, 21406–21411.

Vuong, T.V., and Wilson, D.B. (2010). Glycoside hydrolases: catalytic base/nucleophile diversity. *Biotechnol. Bioeng.* 107, 195–205.

**W**

Walle, T. (2004). Absorption and metabolism of flavonoids. *Free Radic. Biol. Med.* 36, 829–837.

Walle, T., Walle, U.K., and Halushka, P.V. (2001a). Carbon dioxide is the major metabolite of quercetin in humans. *J. Nutr.* 131, 2648–2652.

Walle, T., Otake, Y., Brubaker, J.A., Walle, U.K., and Halushka, P.V. (2001b). Disposition and metabolism of the flavonoid chrysin in normal volunteers. *Br. J. Clin. Pharmacol.* 51, 143–146.

Wang, X. (2009). Structure, mechanism and engineering of plant natural product glycosyltransferases. *FEBS Lett.* 583, 3303–3309.

Wang, Y., Chen, S., and Yu, O. (2011). Metabolic engineering of flavonoids in plants and microorganisms. *Appl. Microbiol. Biotechnol.* 91, 949–956.

- Webb, M.R., and Ebeler, S.E. (2004). Comparative analysis of topoisomerase IB inhibition and DNA intercalation by flavonoids and similar compounds: structural determinates of activity. *Biochem. J.* 384, 527.
- Weng, C.-J., Chen, M.-J., Yeh, C.-T., and Yen, G.-C. (2011). Hepatoprotection of quercetin against oxidative stress by induction of metallothionein expression through activating MAPK and PI3K pathways and enhancing Nrf2 DNA-binding activity. *New Biotechnol.* 28, 767–777.
- Werner, S.R., and Morgan, J.A. (2009). Expression of a *Dianthus* flavonoid glucosyltransferase in *Saccharomyces cerevisiae* for whole-cell biocatalysis. *J. Biotechnol.* 142, 233–241.
- Werner, S.R., and Morgan, J.A. (2010). Controlling selectivity and enhancing yield of flavonoid glycosides in recombinant yeast. *Bioprocess Biosyst. Eng.* 33, 863–871.
- Weston, L.A., and Mathesius, U. (2013). Flavonoids: Their Structure, Biosynthesis and Role in the Rhizosphere, Including Allelopathy. *J. Chem. Ecol.* 39, 283–297.
- Williams, C.A., and Grayer, R.J. (2004). Anthocyanins and other flavonoids. *Nat. Prod. Rep.* 21, 539–573.
- Williams, G.J., and Thorson, J.S. (2008). A high-throughput fluorescence-based glycosyltransferase screen and its application in directed evolution. *Nat. Protoc.* 3, 357–362.
- Williams, G.J., Zhang, C., and Thorson, J.S. (2007). Expanding the promiscuity of a natural-product glycosyltransferase by directed evolution. *Nat. Chem. Biol.* 3, 657–662.
- Williamson, G., and Manach, C. (2005). Bioavailability and bioefficacy of polyphenols in humans. II. Review of 93 intervention studies. *Am. J. Clin. Nutr.* 81, 243S–255S.
- Willits, M.G., Giovanni, M., Prata, R.T.N., Kramer, C.M., De Luca, V., Steffens, J.C., and Graser, G. (2004). Bio-fermentation of modified flavonoids: an example of in vivo diversification of secondary metabolites. *Phytochemistry* 65, 31–41.
- Winkel-Shirley, B. (2001). Flavonoid Biosynthesis. A Colorful Model for Genetics, Biochemistry, Cell Biology, and Biotechnology. *Plant Physiol.* 126, 485–493.
- Winkel-Shirley, B. (2002). Biosynthesis of flavonoids and effects of stress. *Curr. Opin. Plant Biol.* 5, 218–223.
- Withers, S.G. (2001a). Mechanisms of glycosyl transferases and hydrolases. *Carbohydr. Polym.* 44, 325–337.
- Withers, S.G. (2001b). Mechanisms of glycosyl transferases and hydrolases. *Carbohydr. Polym.* 44, 325–337.
- Woo, H.-J., Kang, H.-K., Nguyen, T.T.H., Kim, G.-E., Kim, Y.-M., Park, J.-S., Kim, D., Cha, J., Moon, Y.-H., Nam, S.-H., et al. (2012). Synthesis and characterization of ampelopsin glucosides using dextransucrase from *Leuconostoc mesenteroides* B-1299CB4: Glucosylation enhancing physicochemical properties. *Enzyme Microb. Technol.* 51, 311–318.
- Wu, X., Chu, J., Wu, B., Zhang, S., and He, B. (2013). An efficient novel glycosylation of flavonoid by  $\beta$ -fructosidase resistant to hydrophilic organic solvents. *Bioresour. Technol.* 129, 659–662.

## Y

- Yamada, M., Tanabe, F., Arai, N., Mitsuzumi, H., Miwa, Y., Kubota, M., Chaen, H., and Kibata, M. (2006). Bioavailability of Glucosyl Hesperidin in Rats. *Biosci. Biotechnol. Biochem.* 70, 1386–1394.
- Yamashita, Y., Bowen, W., Burne, R., and Kuramitsu, H. (1993). Role of the *Streptococcus mutans* *gtf* genes in caries induction in the specific-pathogen-free rat model. *Infect. Immun.* 61, 3811–3817.
- Yamazaki, M., Yamagishi, E., Gong, Z., Fukuchi-Mizutani, M., Fukui, Y., Tanaka, Y., Kusumi, T., Yamaguchi, M., and Saito, K. (2002). Two flavonoid glucosyltransferases from *Petunia hybrida*: molecular cloning, biochemical properties and developmentally regulated expression. *Plant Mol. Biol.* 48, 401–411.
- Yang, M., Davies, G.J., and Davis, B.G. (2007). A Glycosynthase Catalyst for the Synthesis of Flavonoid Glycosides. *Angew. Chem. Int. Ed.* 46, 3885–3888.
- Yao, X., Zhu, L., Chen, Y., Tian, J., and Wang, Y. (2013). In vivo and in vitro antioxidant activity and  $\alpha$ -glucosidase,  $\alpha$ -amylase inhibitory effects of flavonoids from *Cichorium glandulosum* seeds. *Food Chem.* 139, 59–66.
- Yi, L.-T., Li, C.-F., Zhan, X., Cui, C.-C., Xiao, F., Zhou, L.-P., and Xie, Y. (2010). Involvement of monoaminergic system in the antidepressant-like effect of the flavonoid naringenin in mice. *Prog. Neuropsychopharmacol. Biol. Psychiatry* 34, 1223–1228.
- Yip, V.L.Y., Varrot, A., Davies, G.J., Rajan, S.S., Yang, X., Thompson, J., Anderson, W.F., and Withers, S.G. (2004). An Unusual Mechanism of Glycoside Hydrolysis Involving Redox and Elimination Steps by a Family 4  $\beta$ -Glycosidase from *Thermotoga maritima*. *J. Am. Chem. Soc.* 126, 8354–8355.

Yonekura-Sakakibara, K., Tohge, T., Niida, R., and Saito, K. (2007). Identification of a flavonol 7-O-rhamnosyltransferase gene determining flavonoid pattern in Arabidopsis by transcriptome coexpression analysis and reverse genetics. *J. Biol. Chem.* 282, 14932–14941.

Yonekura-Sakakibara, K., Tohge, T., Matsuda, F., Nakabayashi, R., Takayama, H., Niida, R., Watanabe-Takahashi, A., Inoue, E., and Saito, K. (2008). Comprehensive Flavonol Profiling and Transcriptome Coexpression Analysis Leading to Decoding Gene–Metabolite Correlations in Arabidopsis. *Plant Cell Online* 20, 2160–2176.

Yoon, J.-A., Kim, B.-G., Lee, W.J., Lim, Y., Chong, Y., and Ahn, J.-H. (2012). Production of a novel quercetin glycoside through metabolic engineering of *Escherichia coli*. *Appl. Environ. Microbiol.* 78, 4256–4262.

## Z

Zduńczyk, Z., Juśkiewicz, J., and Estrella, I. (2006). Cecal parameters of rats fed diets containing grapefruit polyphenols and inulin as single supplements or in a combination. *Nutrition* 22, 898–904.

Zhao, J., and Dixon, R.A. (2010). The “ins” and “outs” of flavonoid transport. *Trends Plant Sci.* 15, 72–80.

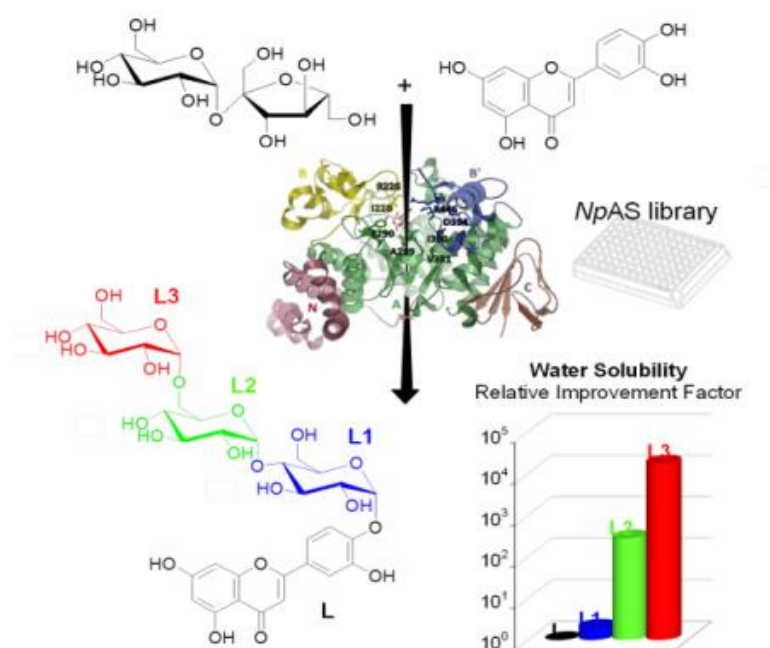
Zhen, L., Zhu, J., Zhao, X., Huang, W., An, Y., Li, S., Du, X., Lin, M., Wang, Q., Xu, Y., et al. (2012). The antidepressant-like effect of fisetin involves the serotonergic and noradrenergic system. *Behav. Brain Res.* 228, 359–366.

Zhou, W., Shan, J., Tan, X., Zou, J., Yin, A., Cai, B., and Di, L. (2013). Effect of chito-oligosaccharide on the oral absorptions of phenolic acids of Flos Lonicerae extract. *Phytomedicine*, 21 (2), 184–194.

Zhu, Q.Y., Huang, Y., and Chen, Z.Y. (2000). Interaction between flavonoids and alpha-tocopherol in human low density lipoprotein. *J. Nutr. Biochem.* 11, 14–21.

# Chapter II

## Extending the structural diversity of $\alpha$ -flavonoid glycosides with engineered glucansucrases



Yannick Malbert, Sandra Pizzut-Serin, Stéphane Massou, Emmanuelle Cambon, Sandrine Laguerre, Pierre Monsan, François Lefoulon, Sandrine Morel, Isabelle André, Magali Remaud-Simeon

ChemcatChem (2014), accepted





---

The first objective of this work was to assess the potential of engineered glucansucrases for flavonoid glucosylation. Within the frame of collaboration with “Les Laboratoires Servier”, we first focused on the amylosucrase from *N. polysaccharea*, which presented some advantages to develop an engineering approach:

- Well-characterized, both structurally and biochemically.
- Previously engineered to glucosylate non-natural glycoside acceptors.
- Efficient and well controlled production of soluble and active enzyme was in both microplates and larger volumes.

Preliminary assays (data not shown) revealed that wild-type *NpAS* was able to glucosylate about 7% of 5 mM luteolin, in the presence of 3% DMSO. Thus, we decided to choose luteolin as a model flavonoid to develop a screening method and evaluate the potential of an engineered amylosucrases library for flavonoid glucosylation and hopefully generate new interesting compounds.

Our objectives included:

- Target the residues potentially involved in the accommodation of flavonoid at the *NpAS* acceptor binding site, and constitute a small library of variants.
- Develop an efficient screening method for luteolin glucosylation (and possibly applicable to flavonoids in general) and screen the *NpAS* library.
- Produce the novel luteolin glucosides in sufficient amounts to allow their structural and biochemical characterization.
- Correlate the activity and regiospecificity of interesting *NpAS* mutants to their molecular structure.

The corresponding results are detailed and discussed in this chapter.

---

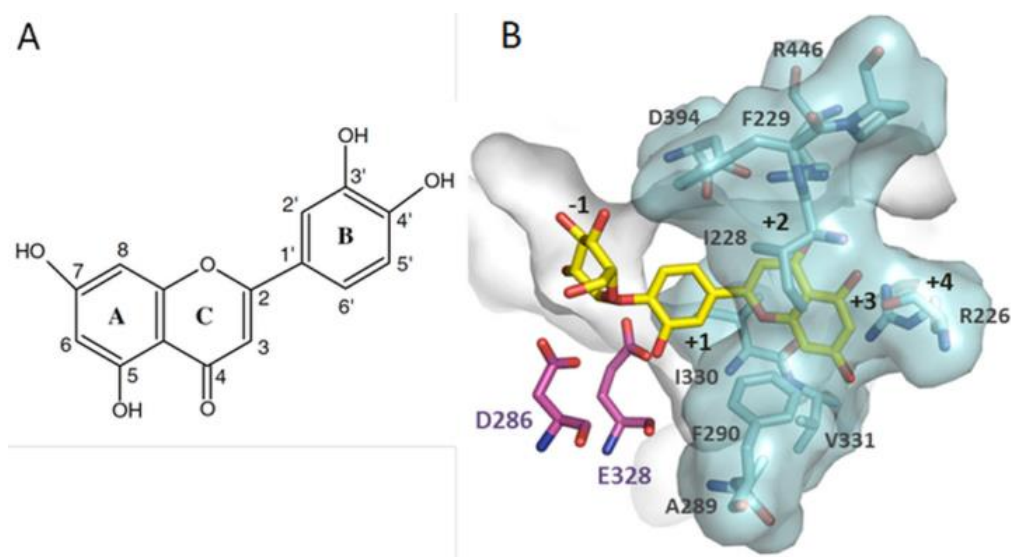
## I. Introduction

Flavonoids are natural polyphenolic compounds that are among the most widely distributed secondary metabolites throughout the plant kingdom. Over 8,000 different structures have been identified in plants and classified into six subclasses (flavonols, flavones, isoflavones, flavanones, anthocyanidins and flavanol) on the basis of the structural variations observed in the rings of their common di-phenylpropane core structure (C6-C3-C6) (Figure 45). These molecules have raised a growing interest in the past ten years and an increasing number of reports describe their physicochemical properties as well as biological activities (Galleano et al., 2010; Quideau et al., 2011).

Among them, Luteolin, (5,7,3',4'-tetrahydroxyflavone) belongs to the flavone subclass and is found in many edible plants. Like for many other flavonoids, numerous investigations indicate that luteolin or luteolin glycosides exhibit anti-oxidant, anti-inflammatory, anti-tumor, antiviral and anti-allergic properties and may stand out as promising candidates for the development of new bioactive molecules for the prevention or treatment of human diseases (Lopez-Lazaro, 2009). However, most of the flavonoids show a poor solubility in water limiting their bioavailability. To circumvent such drawbacks, flavonoid glycosylation was proposed to modify their hydrophilic-lipophilic balance and access to a wider structural diversity. Due to the difficulties and poor regio-specificity of chemical glycosylation, glycodiversification of flavonoids using enzymes has emerged as a promising alternative to generate novel and more soluble structures (Wang et al., 2010). In vivo glycosylation is usually catalyzed by Leloir-type glycosyltransferases (GTs, E.C. 2.4). Recently, the use of native or engineered glycosyltransferases of Leloir-type (GTs, E.C. 2.4) has been rapidly expanding (Coutinho et al., 2003; Palcic, 2011). However, although promising, GTs and their NDP-sugar substrates are not easily available. Alternatively, glycosylation of flavonoids catalyzed by transglycosylases (E.C. 2.4.1) has been attempted. Various glycoderivatives were successfully obtained with cyclodextrin glycosyl-transferases, (Kometani et al., 1996) maltogenic amylases (Cho et al., 2000; Lee et al., 1999; Li et al., 2004) and also glucansucrases from Glycoside-Hydrolase (GH) Families 70 (Bertrand et al., 2006) and 13 (Cho et al., 2011).

Naturally, glucansucrases catalyze the synthesis of  $\alpha$ -glucans with a concomitant release of fructose from sucrose substrate. In the polymer, the  $\alpha$ -D-glucopyranosyl units are linked by  $\alpha$ -(1 $\rightarrow$ 2),  $\alpha$ -(1 $\rightarrow$ 3),  $\alpha$ -(1 $\rightarrow$ 4) or  $\alpha$ -(1 $\rightarrow$ 6) osidic bonds depending on the enzyme specificity (Leemhuis et al., 2013). The reaction also yields limited amounts of by-products such as glucose or sucrose isomers resulting from transfer of the glucosyl units onto either water or fructose, respectively. A remarkable property of these transglycosylases comes from

their natural promiscuity towards various types of acceptor molecules such as saccharides, alcohols, polyols, or flavonoids (Daudé et al., 2013). In particular, various types of flavonoids including catechin (Cho et al., 2011), epigallocatechin gallate (Hyun et al., 2007), quercetin (Bertrand et al., 2006), myricetin (Bertrand et al., 2006), astragalin (Kim et al., 2012), ampelopsin (Woo et al., 2012) and luteolin (Bertrand et al., 2006) were tested as acceptors. Notably, luteolin-4'-O-O- $\alpha$ -D-glucopyranoside and luteolin-3'-O-O- $\alpha$ -D-glucopyranoside were synthesized by acceptor reaction catalyzed by dextran sucrose B-512F, a glucansucrase from GH family 70 (Bertrand et al., 2006) with 44% conversion rates in the presence of 30% organic solvent used to increase the flavone solubility. Alternansucrase, another glucansucrase from GH family 70, was also tested but a very low conversion (8%) of luteolin was reached.



**Figure 45. (A) Chemical structure and numbering of luteolin (3',4',5,7-tetrahydroxy-flavone or 2-(3',4'-dihydroxyphenyl)-5,7-dihydroxy-4H-1-benzopyran-4-one). (B) Side view of *N. polysaccharea* amylosucrase active site pocket showing the 9 amino acid mutagenesis targets encircling the subsite+1. Luteolin-4'-O-O- $\alpha$ -D-glucopyranoside represented in yellow sticks is docked in the active site. Catalytic residues D286 and E328 are shown in magenta sticks. The subsite numbering from -1 to +4 is shown on the figure.**

As for GTs, protein engineering could help to improve and expand the synthetic utility of glucansucrases for flavonoid glycoside synthesis. Recently, amylosucrases with completely novel acceptor specificity were successfully engineered using a structurally-guided approach. The parental enzyme from *N. polysaccharea* (*NpAS*) (EC 2.4.1.4) belongs to GH family 13 and naturally catalyzes from sucrose the *de novo* synthesis of a water-insoluble amylose-like polymer with  $\alpha$ -(1 $\rightarrow$ 4) glycosidic bonds. A set of eight amino acid positions located in flexible loops of the ( $\beta/\alpha$ )<sub>8</sub> catalytic barrel and that encircle acceptor subsite +1 of *NpAS* were replaced to generate a library of 133 mutants from which variants showing novel or highly enhanced specificities toward unnatural acceptors such as protected  $\alpha$ -methyl-L-rhamnoside and  $\alpha$ -allyl-N-acetyl-glucosamine were successfully obtained (Champion et al., 2009, 2012).

However, glucansucrase and more generally  $\alpha$ -transglucosylase engineering has never been attempted to generate catalysts specialized in flavonoid glycosylation. Flavonoids are generally highly hydrophobic whereas sucrose is hydrophilic. In addition, glucansucrases are not very tolerant to organic solvent (Bertrand et al., 2006).

Such incompatibilities clearly hampered the design of miniaturized, accurate and rapid screening protocols. In this paper, these problems have been tackled with the view of diversifying luteolin glycoside structures and improving luteolin solubility. A full procedure was developed and applied to screen a library of amylosucrase mutants. Reaction scale-up with improved variants confirmed their properties and allowed the purification and structural characterization of the luteolin glucosides. Physico-chemical properties such as their water solubility, UV spectra and antioxidant capacity were determined. Finally, docking experiments were conducted to get insight into the structural changes that controlled the catalytic properties of the best mutants.

## II. Results

### II.1. Library and design of the screening assay

The presence of two vicinal groups on catechol ring was previously reported to be important for flavonoid glycosylation with glucansucrases (Bertrand et al., 2006). Thus, we performed docking experiments of luteolin-4'-O- $\alpha$ -D-glucopyranoside and luteolin-3'-O- $\alpha$ -D-glucopyranoside in *NpAS* acceptor binding site. Only the Luteolin-4' derivatives could be easily accommodated due to steric hindrance with catalytic residue E328 (Figure 45 B).

**Table 11. Amounts and relative distribution of luteolin glucoside products synthesized with the best representative mutant of each product profiles shown in Figure 46.**

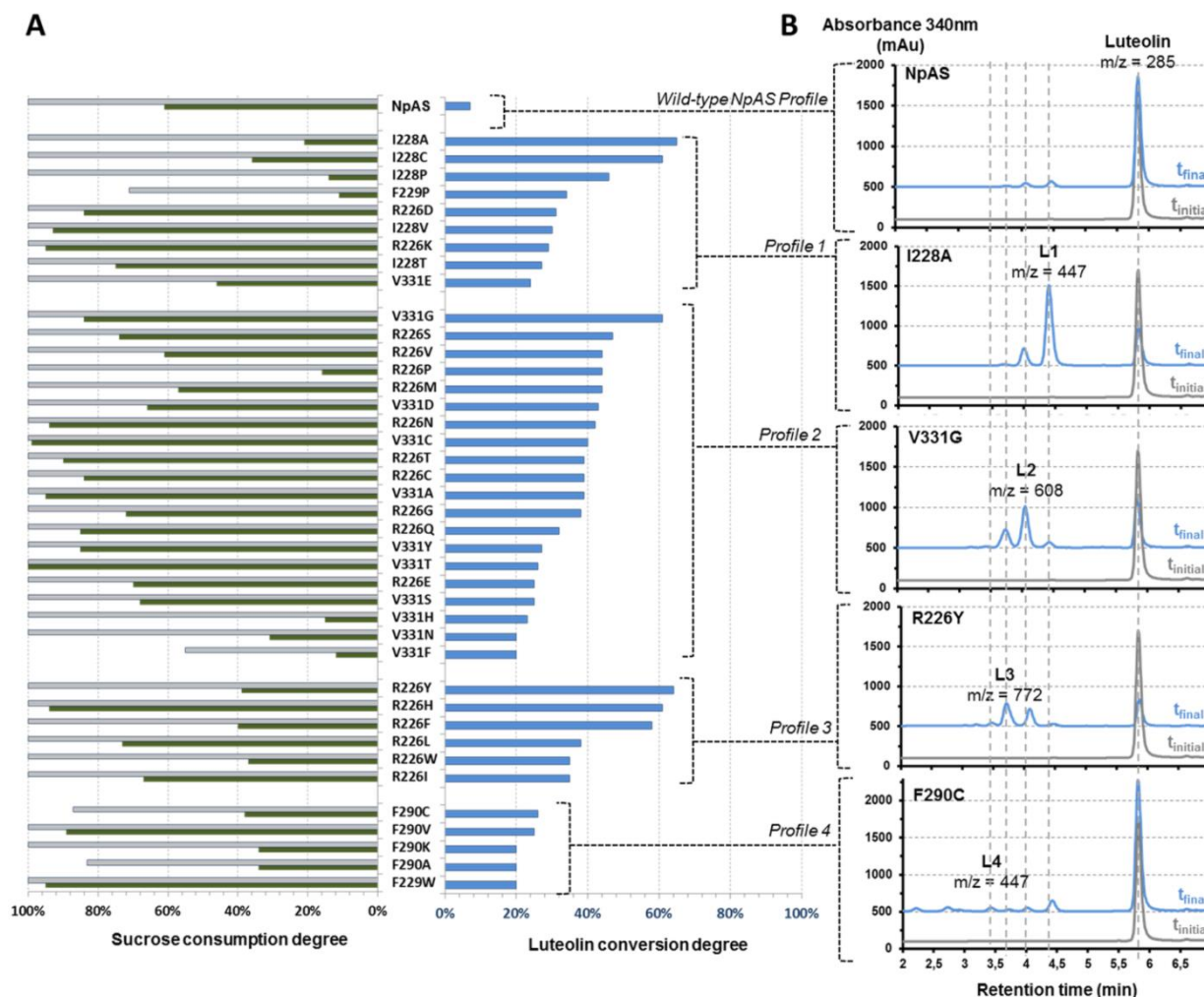
Profile	<i>NpAS</i> mutant	Converted luteolin (mM)	Conversion (%)	Luteolin glucoside (mM)			Luteolin glucoside distribution <sup>[a]</sup>		
				L1	L2	L3	L1	L2	L3
WT	Wild type	0.35	7%	0.21	0.10	0.03	62%	29%	9%
1	I228A	3.25	65%	2.48	0.52	0.04	82%	17%	1%
2	V331G	3.05	61%	0.22	1.83	0.91	7%	62%	31%
3	R226Y	3.20	64%	0.16	1.08	1.68	5%	37%	58%
4	F290C	1.30	26%	0.41	0.32	0.40	36%	29%	35%

<sup>[a]</sup> Relative distributions of the three main luteolin glucosides were determined from HPLC analyzes using the formula :  $(Lx \text{ peak area}) / \sum(L1, L2 \text{ and } L3 \text{ peaks areas})$ , where  $1 \leq x \leq 3$ .

Further analysis of the model revealed that eight amino acids that are located in subsite +1 of the active site and were previously selected for mutation to improve *NpAS* specificity for unnatural acceptors could also help to favor luteolin recognition (Champion et al., 2009) (Figure 45 B). In addition, Arg 226 which occupies subsite +2 was found to induce a steric clash with the flavone. Therefore, the 19 single variants generated at this position were also included in the library (Cambon et al., 2014a), which finally comprised a total of 171 single mutants targeting altogether nine amino acids of the acceptor binding site (Figure 45B). A semi-automated and standardized protocol including cell growth in deep-well, cell lysis, activity assay (sucrose consumption) and acceptor reaction assay was then developed (Figure S1). The acceptor reaction was carried out in microtiter plates in the presence of sucrose (146 mM), luteolin (5mM) and 3% dimethylsulfoxide (DMSO, v/v) to enhance luteolin solubility and preserve at least 80 percent of the initial wild type enzyme activity (Figure S2). These conditions allowed the partial dissolution of 106  $\mu$ M of luteolin (versus 40  $\mu$ M in the absence of DMSO) whereas sucrose was totally dissolved. The relative standard deviation values were of 10% for bacteria growth, less than 20% for activity values and below 5% for luteolin conversion degrees. These controls showed an acceptable variability. In particular, luteolin conversion degree was quite reproducible although the medium was heterogeneous. Consequently, a 15% threshold was fixed as the value above which a variant is considered as significantly improved. *NpAS* library screening.

The library of *NpAS* mutants was screened according to the established protocol. 74% of the mutants were still active on sucrose substrate (Table S1). Compared to the parental enzyme, 50 mutants (29% of the library) were identified to be improved for the conversion of luteolin. 40 variants converted at least 20% of luteolin (1 mM), with conversion degrees ranging from 20% to 66% depending on the mutant (Figure 46 A). Notably, most of them were as active as the wild type enzyme on sucrose substrate alone, with the exception of F229P, F290C, V331F, F290A variants (Figure 46 A). In the presence of luteolin, sucrose consumption was always reduced although to various extents depending on the variant, hence revealing a possible inhibitory effect of the flavone. Despite this, all the variants consumed from 15% to 90% of the sucrose donor in the screening conditions. This represents sucrose concentrations ranging from 21 to 131.4 mM, which are in large excess compared to luteolin conversions, comprised between 1 and 3.3 mM. Of note, the wild type enzyme converted only 0.35 mM of luteolin in similar conditions.

In addition, four different glucosylation profiles could be distinguished (Figure 46 B) and no correlation could be deduced between inhibition of sucrose consumption and luteolin conversion. Luteolin conversion degree and glucoside distribution obtained with the best representatives of each category are reported in Table 11.

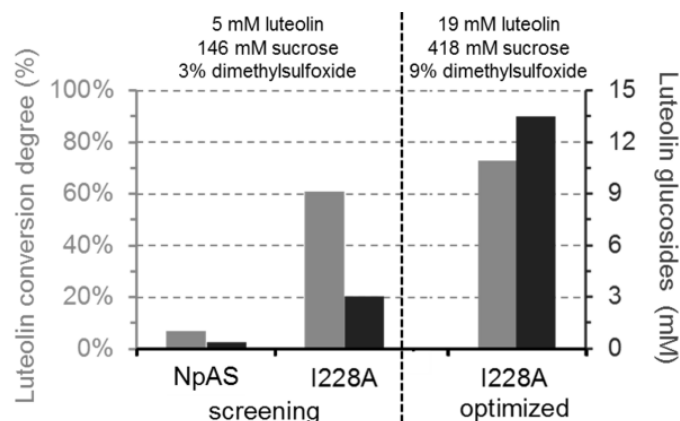


**Figure 46. Results of NpAS single mutant library screening.** (A) Gray bars: sucrose consumption degrees obtained without luteolin; green bars: sucrose consumption degrees observed with luteolin; blue bars: luteolin conversion degrees. Luteolin conversion degree =  $100 \times ([\text{luteolin}]_{\text{ti}} - [\text{luteolin}]_{\text{tf}}) / [\text{luteolin}]_{\text{ti}}$ ; Sucrose consumption degree =  $100 \times ([\text{sucrose}]_{\text{ti}} - [\text{sucrose}]_{\text{tf}}) / [\text{sucrose}]_{\text{ti}}$  where  $t_i$  and  $t_f$  are the initial and final times (20h) of the reaction, respectively. (B) Profiles of luteolin glucosylation products analyzed by HPLC-UV/MS; in gray, initial reaction time; in blue, final reaction time.

Profile 1 regroups variants mutated at positions 228, 226 and 229 that produce mainly compound L1 with a molecular mass of  $448 \text{ g mol}^{-1}$ , corresponding to a mono-glucosylated form of luteolin. (Figure 46 B, profile 1). In profile 2, the formation of di-glucosylated luteolin (L2;  $609 \text{ g mol}^{-1}$ ) is dominant and performed by variants with mutations mainly found at positions 226 and 331 (Figure 46 B, profile 2). A third group of variants mainly produced tri-glucosylated luteolin (L3;  $772 \text{ g mol}^{-1}$ ) (Figure 46 B, profile 3). Finally, a more polydisperse product profile was observed for F290C and F290V variants (Figure 46 B, profile 4). With these variants, the conversion degree of luteolin did not exceed 25% but interestingly, in addition to L1, L2 and L3, another mono-glucosylated form of luteolin was detected indicating that glucosylation occurred at two distinct positions.

## II.2. Production and structural characterization of luteolin glucosides L1, L2 and L3

To improve the production yields of L1 and L2 compounds with I228A mutant, a response surface methodology (RSM) using a Box-Benken experimental design was first performed considering three factors: X1) DMSO concentration (v/v); X2) luteolin amounts; X3) sucrose/luteolin molar ratio (R(S/L)). For all experiments, no significant changes in the product chromatographic profiles were observed. The analysis of variance (ANOVA) showed  $R^2$  values of 92.7% and 83.57% for luteolin conversion degrees and converted luteolin concentration, respectively. From response surface contour plot analyses (Figure S3), the model predicted the conversion of 14.1 mM of luteolin (with an associated conversion degree of 71%) in the following conditions: luteolin (19 mM), DMSO (9%), sucrose (418 mM) and enzyme (0.8 U.mL<sup>-1</sup>). The validation tests resulted in the glucosylation of 13.7 mM luteolin corresponding to 72% conversion, which was in very good agreement with the RSM model predicted values. Overall, a 40-fold increase in glucosylation level associated with a 10-fold increase in glucosylation efficiency were observed compared to that initially obtained with the wild type enzyme (Figure 47). These reaction conditions were used to scale up the production of L1 and L2 to several hundred mg. Regarding L3 production, mutant R226Y was preferred.

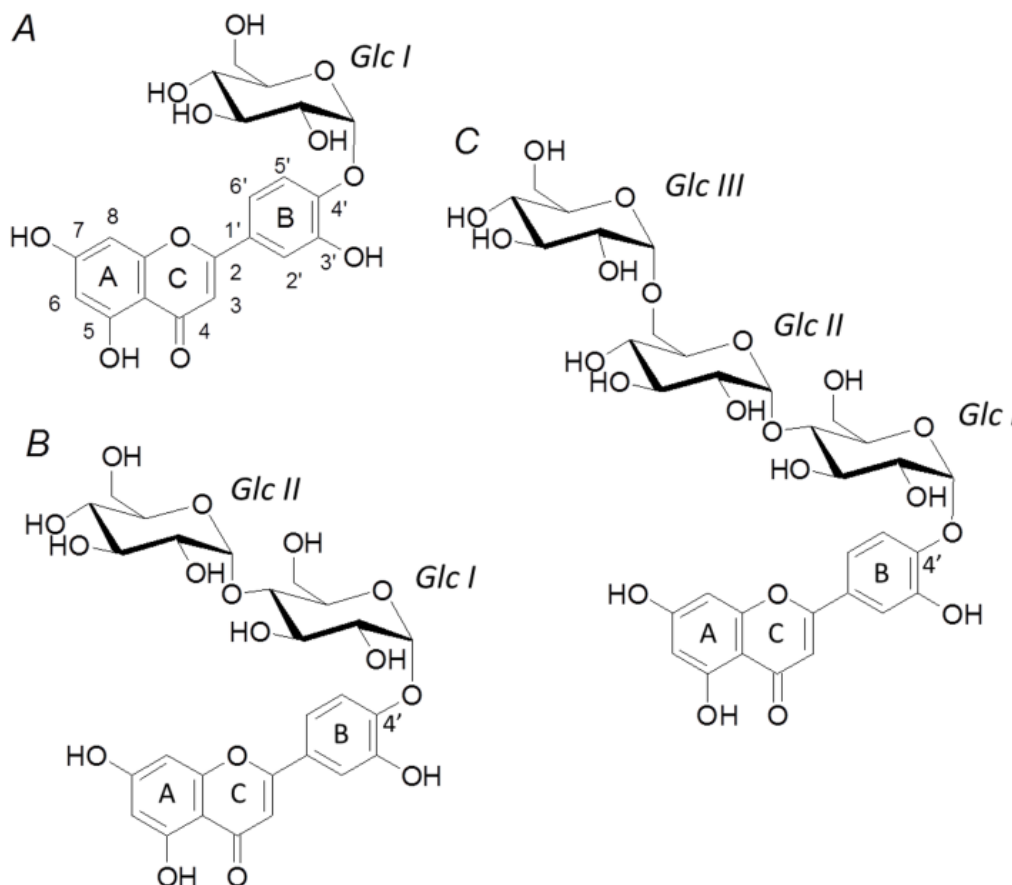


**Figure 47. Comparison between luteolin glucosides productions with wild-type NpAS and I228A mutant in screening and optimized reaction conditions.** Gray bars: luteolin conversion degrees; black bars: amounts of glucosylated luteolin produced.

The structures of the purified compounds were determined by NMR analyses (Table S2 and Figure S4). First of all, the 21 signals arising from L1 in <sup>13</sup>C NMR spectrum were identical to those reported for luteolin-4'-O- $\alpha$ -D-glucopyranoside in Bertrand et al., 2006. Regarding L2, of the twenty seven signals observed in <sup>13</sup>C NMR spectrum, 15 signals were assigned to luteolin moiety ( $\delta$  95.05 – 183.81). Comparison between L1 and L2 spectra also revealed that L2 results from L1 glucosylation and possesses an additional glucosyl unit (*GlcII*) linked



to the glucosyl unit I (*Glc I*) of L1 through an  $\alpha$ -(1 $\rightarrow$ 4) linkage. Indeed, the chemical shift from C4 of *Glc I* moved downfield from 70.4 to 80.90 ppm indicating the presence of an additional glycosidic linkage and the  $\alpha$ -anomeric configuration is attested by the coupling constant value ( $J = 3.6$  Hz) of the *Glc II* anomeric proton at 5.19 ppm. HMBC data of L2 confirmed these results with the carbon-proton coupling identified between H1 of *Glc II* unit and C4 of *Glc I* at 80.90 ppm (Table S2 and Figure S4).



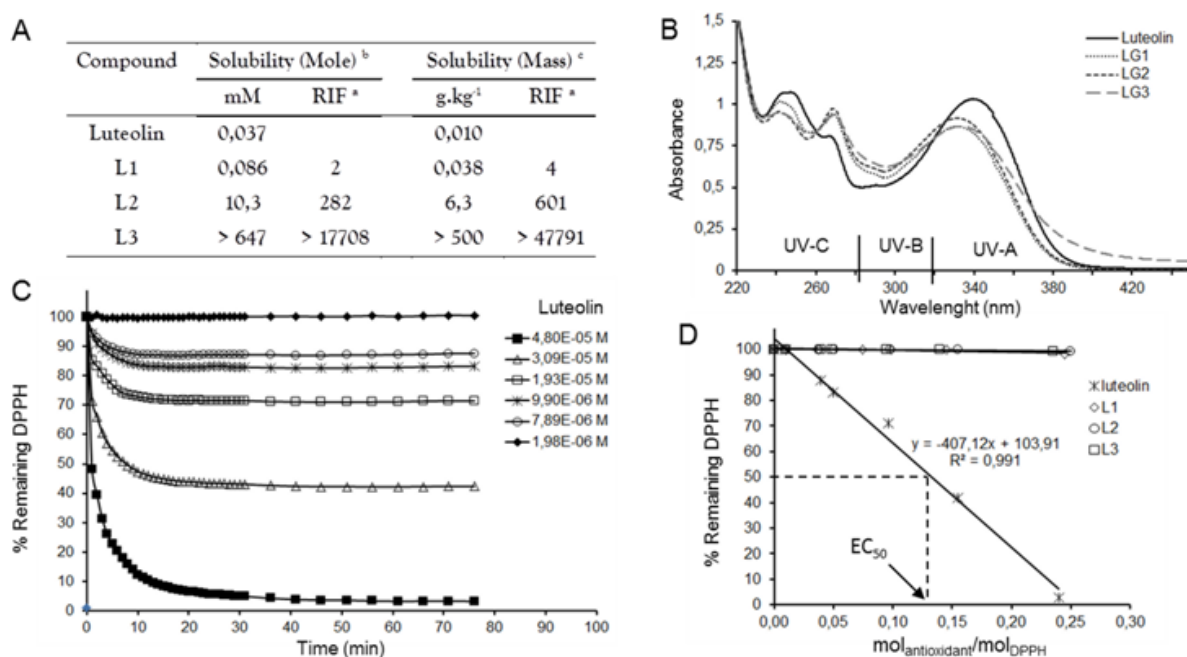
**Scheme 1.** Structures of the luteolin glucosides produced with NpAS mutants I228A and R226Y (A) L1; luteolin-4'-O- $\alpha$ -D-glucopyranoside. (B) L2; luteolin-4'-O- $\alpha$ -D-glucopyranosyl- $\alpha$ (1 $\rightarrow$ 4)-D-glucopyranoside. (C) L3; luteolin-4'-O- $\alpha$ -D-glucopyranosyl- $\alpha$ (1 $\rightarrow$ 4)-D-glucopyranosyl- $\alpha$ (1 $\rightarrow$ 6)-D-glucopyranoside.

Compared with  $^{13}\text{C}$  NMR spectrum of L2, six additional signals were identified in the spectrum of L3. The chemical shift from C6 of *Glc II* was displaced from 62.76 to 68.45 ppm indicating that C6 of *Glc II* unit is engaged into a (1 $\rightarrow$ 6) osidic linkage of  $\alpha$  configuration as shown by the coupling constant value ( $J = 3.6$  Hz). The presence of this  $\alpha$ -(1 $\rightarrow$ 6) linkage was also demonstrated by the carbon-proton coupling between H1 of *Glc III* (signal at 4.80 ppm) and C6 of *Glc II* (Table S2 and Figure S4). As wild-type NpAS and its R226Y mutant only catalyze the formation of  $\alpha$ -(1 $\rightarrow$ 4) glucosidic bonds, the  $\alpha$ -(1 $\rightarrow$ 6) linkage between *Glc II* and *Glc III* of the tri-glucosylated luteolin L3 was totally unexpected. The structures of L1, L2 and L3 correspond to luteolin-4'-O- $\alpha$ -D-glucopyranoside, luteolin-4'-O- $\alpha$ -D-glucopyranosyl- $\alpha$ -

(1 $\rightarrow$ 4)-D-glucopyranoside and luteolin-4'-O- $\alpha$ -D-glucopyranosyl- $\alpha$ -(1 $\rightarrow$ 4)-D-glucopyranosyl- $\alpha$ -(1 $\rightarrow$ 6)-D-glucopyranoside, respectively (Scheme 1).

### II.3. Physico-chemical characterization of the glucosides

As seen in Figure 48, glucosylation modifies the hydrophilic-lipophilic balance of L1, L2 and L3 and clearly promotes their solubility in water. Water solubility of L1 remained quite low with only a 2-fold increase compared to that of luteolin (37  $\mu$ M) whereas a significant rise of solubility is observed for L2 and L3, leading to concentrations up to 10 mM and 647 mM (500 g kg<sup>-1</sup>), respectively. In the case of L3, the solubilization could not exceed 647 mM due to the very high viscosity of the medium at this concentration. Noteworthy, triglucosylation of luteolin enabled to improve by 17,700-fold the water solubility of the flavone (Figure 48 A). Such a level of solubility has never been reported for functionalized forms of flavones. The UV-absorption spectrum of the aglycon luteolin showed two maxima at 340 nm (UV-A) and 250 nm (UV-C range) and a minimum absorption at 284 nm (UV-B zone), which is consistent with previously reported results (Fischer et al., 2011) (Figure 48 B).



**Figure 48. Physico-chemical properties of the luteolin glucosides L1, L2 and L3.** (A) Table of aqueous solubility of luteolin and its glucosides. Values correspond to the mean of three independent measurements; error ranges below 10%. <sup>a</sup> Relative Improvement Factor (RIF) = 100\*([compound]/[luteolin]). <sup>b</sup> solubility is expressed using molar concentrations. <sup>c</sup> solubility is expressed using mass concentrations. (B) UV spectra of luteolin, L1, L2 and L3, measured in acetonitrile. (C) Kinetic curves of scavenged DPPH by several concentrations of luteolin in methanol. (D) Luteolin, L1, L2 and L3 linear fitting of % Remaining DPPH to mol<sub>antioxidant</sub>/mol<sub>DPPH</sub>.

The UV spectra of L1, L2 and L3 were close to that of luteolin (UV-A  $\lambda_{\text{max}} = 330 \text{ nm}$ ; UV-C  $\lambda_{\text{max}} = 250 \text{ nm}$ ; UV-B  $\lambda_{\text{min}} = 294 \text{ nm}$ ) what is in line with previously reported results for the 4'-O- $\beta$ -glucopyranosyl-3',5,7-trihydroxyflavone (Fischer et al., 2011). An additional UV-C maximum was detected at 270 nm for the three luteolin glucosides and slight loss of UV absorption of around 15% and 10% were observed at 330 nm and 250 nm, respectively.

In contrast, an absorption gain of around 15% was detected in the UV-B wavelengths range. No significant differences were noted between the three luteolin glucosides. Finally, the antioxidant activity of luteolin, L1, L2 and L3 were also analyzed using the DPPH test. As seen in Figure 48 C and 48 D, only luteolin showed an antioxidant activity with an EC<sub>50</sub>, expressed as the molar ratio ( $\text{mol}_{\text{antioxidant}}/\text{mol}_{\text{DPPH}}$ ), of 0.13, consistent with previous reports (Tsimogiannis and Oreopoulou, 2004). In case of L1, L2 and L3, no DPPH free radical scavenging was detected showing that glucosylation of the B-ring 4'-hydroxyl group prevents luteolin antioxidant activity.

### III. Discussion

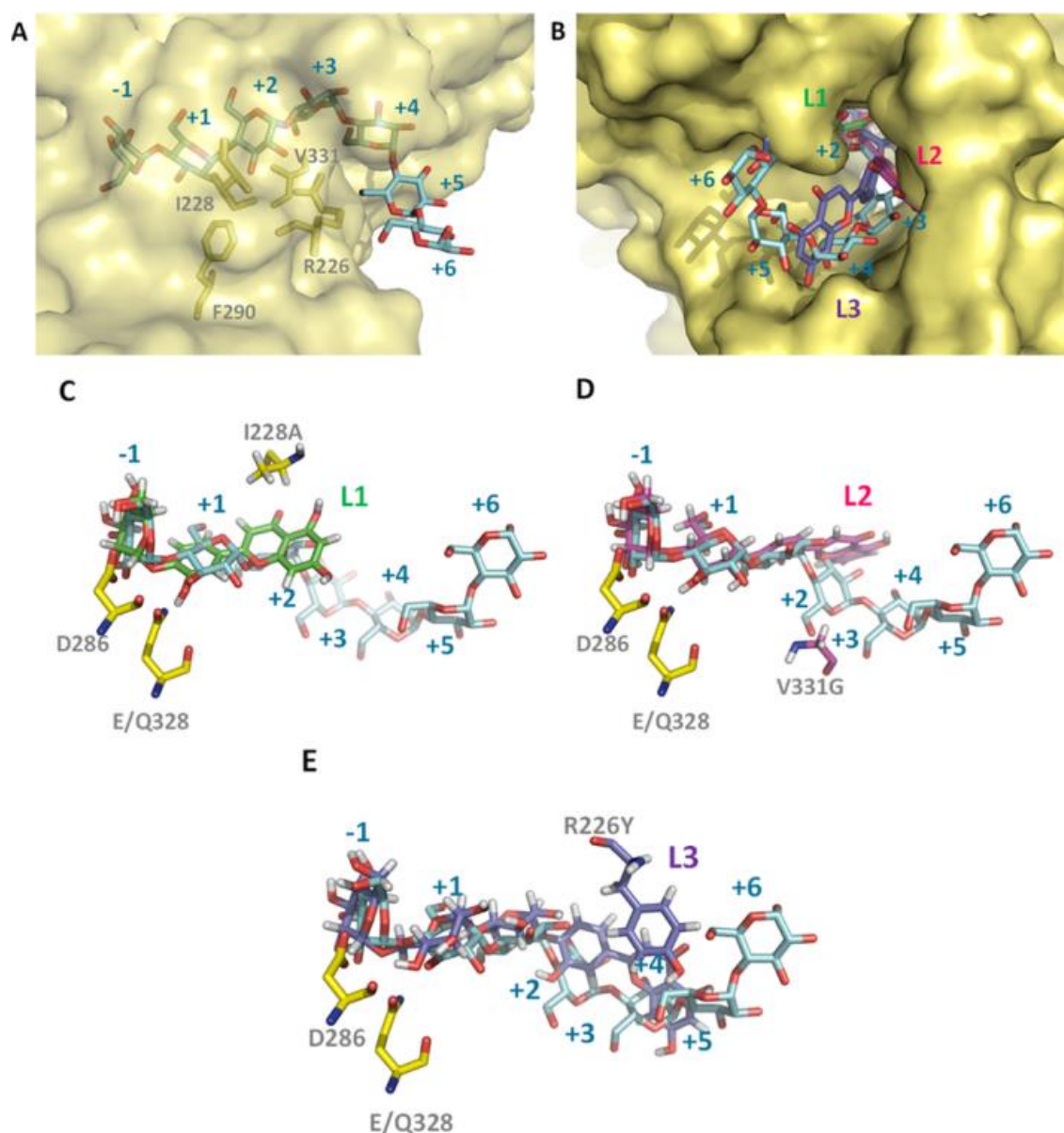
Numerous investigations showed that flavonoids may have beneficial effects on human health. The potency of these products is tightly related to their structures which govern their physicochemical as well as their biological properties. In nature, glycosylated forms of these compounds are produced by Leloir-type glycosyltransferases and most of the glycosyl units are attached to the flavonoids through  $\beta$ -osidic linkages. Moreover, glycosylated materials are usually produced in low amounts and difficulties are encountered to isolate pure and well-defined structures from natural sources. In the present study, the aim was to develop efficient glycosylation routes of flavonoids by taking advantage of enzyme engineering. A single mutant library of amylosucrase from *N. polysaccharea*, a transglucosylase from the GH 13 family (also known as the  $\alpha$ -amylase family) was selected for this purpose and luteolin was used as a model of flavonoid. Of the small library of 171 single mutants targeting nine amino acid positions in this region, 29% of the mutants were found to convert more efficiently luteolin than the parental enzyme did. Notably, three mutants showed luteolin conversion higher than 60%, versus 7% for the parental wild type enzyme. This is to our knowledge the first example of  $\alpha$ -retaining transglucosylase engineering that led to such improvements for flavonoid glycosylation. The present results exemplify the versatility of libraries focused on acceptor binding site to search for enzymes displaying original and non-natural acceptor specificity.

The very poor solubility of luteolin which is only of 37  $\mu\text{M}$  in water clearly represents a difficulty for setting up medium screening assays. We showed that this could be

circumvented by addition of small quantity of DMSO (3%, v/v) and by performing acceptor reaction with an excess of luteolin fixed at 5 mM theoretical concentration. In these conditions, the reaction medium was heterogeneous but such heterogeneity had no severe impact on the reproducibility of the miniaturized assay as revealed by standard deviation values. On the contrary, it allowed a slow phase transfer of luteolin from the solid to the liquid phase during the glucosylation, probably limiting inhibitory activity of luteolin while providing appropriate conditions for increasing production yield. The reduced sucrose consumption observed in the presence of luteolin is an indicator of inhibition. In recent studies, IC<sub>50</sub> values of 18.4  $\mu$ M and 360  $\mu$ M were determined for the inhibition of human salivary  $\alpha$ -amylase (Lo Piparo et al., 2008) and porcine pancreatic  $\alpha$ -amylase (Tadera et al., 2006), respectively. Computational docking suggested that luteolin directly interacts with the catalytic residues of the  $\alpha$ -amylase active site (ie: the nucleophile and the acid/base general catalyst) via the 7-hydroxyl group of A ring and 4'-hydroxyl group of B ring. Notably, the high-resolution structures of human pancreatic  $\alpha$ -amylase in complex with myricetin (a flavonol compound with a structure close to that of luteolin) confirmed that despite its planar structure, myricetin interacts directly with the catalytic residues (Williams et al., 2012). However, in these studies, the flavonoids fit within subsites -3 to +1, of the large binding cleft of salivary  $\alpha$ -amylases as well as in its human pancreatic isoenzyme form. In amylosucrase, there is no subsite -2 or -3; the active site pocket is indeed closed by a salt bridge located just behind subsite -1. Positioning of luteolin may thus be quite different in amylosucrase. Although we may not exclude a possible interaction with subsite -1, luteolin glucosylation clearly demonstrates that the flavone can bind at subsites +1, +2 and +3. Luteolin accommodation in these subsites may block sucrose access due to the pocket funnel topology of the enzyme, resulting in a pseudo-competitive inhibition.

Moreover, library screening further revealed that four positions of the nine targeted positions of mutations are critical for luteolin glucosylation (ie: positions 226, 228, 290 and 331). Position 226 appeared to be clearly strategic to improve luteolin glucosylation, as 18 of the 19 corresponding mutants converted more than 25% luteolin. Mutations of R226 has recently been shown to cause steric hindrance at subsite +2 and +3 of amylosucrase and its mutation by smallest residues enabled *NpAS* polymerase activity to be improved (Albenne et al., 2004). Similarly, replacement of this long side chain may help for luteolin and its glucosylated form binding. In addition, some mutants were more prone to synthesize mono, di or tri-glucosylated forms. Clearly, the different profiles obtained in similar conditions and with equivalent sucrose consumption indicate that the mutations impact the distribution of the various acceptor reaction products. Regarding the three categories, whose representatives are I228A, V331G and R226Y mutants, we carried out docking experiments of L1, L2, and L3

products to investigate further the molecular determinants involved in the product specificity. Energy minimized complexes provided in Figure 49 were used as a basis for detailed structural analysis.



**Figure 49.** View of amylosucrase active site in complex with maltoheptaose (cyan) (PDB: 1MW0) (A). Superimposition of docked L1 (green), L2 (magenta, and L3 (purple) onto maltoheptaose (cyan) (B). Docked L1 into I228A mutant (C), L2 into V331G mutant (D) and L3 into R226Y mutant (E). Catalytic (D286 and E/Q328) and mutated residues are shown in sticks. The subsite numbering from -1 to +6 is shown on the figure.

Comparison of the docking modes adopted by L1, L2 and L3 with the binding mode of maltoheptaose in crystallographic complex with *NpAS* (PDB: 1MW0) enabled a description of interactions involved in the recognition at the binding subsites (Figure 49). Compared to maltoheptaose, the glucosyl unit of L1, L2 and L3 bound at subsite -1 is quasi perfectly superposed with the glucosyl unit of maltoheptaose, establishing a dense network of hydrogen bonding interactions earlier described with amino acid residues D144, H187, R284,

D286, E/N328, H392, D393 and R509 (Mirza et al., 2001) (Figure 49 A and 49 B). Regarding the binding mode of L1, the B phenyl ring was found to fit nicely in subsite +1 with its 3'-hydroxyl group establishing the same hydrogen bonding interaction with E/N328 found for OH3 from maltoheptaose Glc (+1) (Figure 49 C). However, given the nearly-planar geometry of luteolin (Amat et al., 2008), the 1,4-benzopyrone deviated from subsite +2 as defined for maltoheptaose, although endocyclic oxygen from ring C mimics quite well the interosidic oxygen O1 between Glc (+1) and Glc (+2) of maltoheptaose. The 1,4-benzopyrone was found to stack onto I228 whose mutation by an alanine, cysteine or proline led to significant enhancement of luteolin mono-glucosylation (L1). Introduction of less hindered residue at position 228 could assist a better positioning of luteolin in the catalytic pocket. Mutations at nearby positions 226 and 229 may also favor binding of luteolin in a catalytically productive mode, although beneficial effects were less pronounced than for position 228 (Figure 46).

Production of L2 was found to be enhanced using single mutants from profile 2, with best results obtained for V331G mutant (Figure 46). Docking mode adopted by L2 revealed that the presence of the luteolin did not disturb binding of  $\alpha$ -D-Glc-(1 $\rightarrow$ 4)-D-Glc moiety at subsites -1 and +1 which matched remarkably the binding mode of crystallographic maltoheptaose at these subsites (Figure 49 D). In di-glucosylated luteolin (L2), the B phenyl ring was found to sit onto Glc (+2) binding subsite of maltoheptaose. Mutations at positions 226 and 331 were considered to facilitate accommodation of luteolin in productive conformation in the narrow catalytic pocket. Interestingly, mutation of R226 by aromatic or aliphatic residue facilitated the elongation of L2, with the higher production of tri-glucosylated form, L3. Compound L3 showed an unexpected  $\alpha$ -D-Glc-(1 $\rightarrow$ 6) unit at its non-reducing end. Inspection of L3 docking model indicated that the constrained topology of the amylosucrase active site would not be compatible with the accommodation in catalytic manner of  $\alpha$ -D-Glc-(1 $\rightarrow$ 4)- $\alpha$ -D-Glc-(1 $\rightarrow$ 4)-D-Glc-luteolin (Figure 49 B). Glucosylation of L2 by single mutants from profile 3 appears to be only achievable through a more flexible  $\alpha$ -(1 $\rightarrow$ 6) osidic linkage that provides enough conformational flexibility for planar luteolin to bind at subsites +3 and +4, establishing beneficial van der Waals interactions with hydrophobic residues at position 226 (Figure 49 E). Overall, the remarkable fit of L1, L2 and L3 binding modes with crystallographic bound maltoheptaose provides some hints regarding the increase of L1, L2 and L3 production by the different amylosucrase single mutants and the observation of the non-usual  $\alpha$ -(1 $\rightarrow$ 6) linkage in L3. Optimizing the reaction conditions allowed converting 13.7 mM of luteolin starting from 19mM of luteolin. Such a conversion level has never been reported before for a flavone. In a previous report, the maximal amount of luteolin monoglucoside that was synthesized did not exceed 3.6 mM. In addition, the dextransucrase from *L. mesenteroides* NRRL B-512F was not regioselective as both luteolin-3'-O- $\alpha$ -D-



glucopyranoside and luteolin-4'-O- $\alpha$ -D-glucopyranoside were produced and the reaction require the use of larger amounts of organic solvent (Bertrand et al., 2006).

Finally, the results presented herein demonstrate the potency of enzyme design to diversify the panel of glucosylated flavones that can be obtained with transglucosylases. Two novel structures (L2 and L3), never described before, have been disclosed. They correspond to novel  $\alpha$ -glucosylated forms, not found in nature. In addition, we have shown that the solubility and by consequence the bioavailability of these glucoderivatives of luteolin can be tremendously enhanced especially when three glucosyl units are transferred onto this compound. Such level of solubility may greatly facilitate the design of new formulations. The UV spectra of the molecules are little affected compared to that of luteolin. In contrast, antioxidant properties are lost. On one hand, this could be seen as a disadvantage but, on the other hand, glucosylation could also help to preserve the native structure of the active molecule from rapid oxidation, delayed their metabolization and facilitate their absorption. This could provide an efficient way for a new generation of polyphenol-based "prodrugs" with high bioavailability and delayed actions.

## IV. Conclusion

In conclusion, to access to a broader variety of flavonoid-based structures, enzyme-based processes relying on the use of transglucosylases are promising and competitive compared to chemical synthesis. Intensifying the screening of enzymes and acceptors is essential to generate novel structures. Combining such efforts with the power of enzyme engineering should help to improve the glycosylation efficiency and control the glycosylation pattern. This will undoubtedly expand libraries of flavonoid-based structures in the future and should also greatly facilitate the structure-property studies of these compounds.

## V. Experimental Section

### V.1. Bacterial Strains, Plasmids, and Chemicals

The plasmid pGEX-6P-3 (GE Healthcare Biosciences) in which the *Neisseria polysaccharea* Amylosucrase (*NpAS*) encoding gene was fused with a glutathione S-transferase (GST) tag was used for the construction of the *NpAS* single mutant library as reported in Champion et al., 2009. *Escherichia coli* JM109 cells were used as hosts for plasmid library transformation, gene expression, and large scale production of the mutants.

All chemicals were purchased from Sigma-Aldrich (Saint-Louis, MO, USA). Luteolin was supplied by Servier Technologies (Orléans, France).

## V.2. Semi-automated procedure for *NpAS* single mutant library screening

Storage microplates of the single mutants expressing clones were thawed and used to inoculate 96-well microplates containing, in each well, LB medium (150  $\mu$ L) supplemented with ampicillin (100  $\mu$ g.mL<sup>-1</sup>). After 24h growth at 30 °C under agitation (200 rpm), plates were duplicated into 96-Deep Well plates containing, in each well, 1 mL of autoinducible ZYM5052 medium (Studier, 2005) (containing  $\alpha$ -lactose (0.2%), glucose (0.05%), glycerol (0.5%)) to induce glutathion-S-transferase-amylosucrase fusion protein expression, supplemented with ampicillin (100  $\mu$ g.mL<sup>-1</sup>). Cultures were then grown for 24 h at 30 °C under agitation (200 rpm). The plates were centrifuged (20 min, 3000 g, 4 °C), cell pellets were resuspended in lysozyme (300  $\mu$ L; 0.5 mg.mL<sup>-1</sup>) and freezed at -80 °C for 8 to 12 h. After thawing at room temperature, supernatants were transferred onto filter microplates (Glass fiber membrane, PS, 0.25 mm pore) to be clarified by centrifugation (5 min, 2000 g, 4 °C). Filtrates were then transferred in 96-Deep Well Format plates for screening. Luteolin glucosylation reactions were carried out in 300  $\mu$ L final volume in the presence of sucrose (146 mM final) and luteolin (5 mM) and DMSO (3% final). Reactions were incubated at 30 °C for 20 h under agitation (inforS, 700 rpm) and stopped by heating at 95 °C for 5 min. Sucrose consumption and luteolin conversion were then evaluated for each mutant. The reproducibility of the protocol was estimated by determining the relative standard deviation values ( $RSD = (\text{standard deviation}/\text{mean}) \times 100$ ) at each measurable step of the screening procedure for a series of 288 assays corresponding to four 96 well-microplates inoculated with wild-type *NpAS* expressing clones.

## V.3. Analytical methods

Sucrose consumption was monitored by high performance liquid chromatography analysis as previously described in using a Dionex P 680 series pump equipped with an autosampler HTC PAL and a Shodex RI 101 series refractometer and a Biorad HPLC Carbohydrate Analysis column (HPX-87K column (300 mm  $\times$  7.8 mm)) maintained at 65 °C, using ultrapure water as eluent with a flow rate of 0.6 mL.min<sup>-1</sup>. Data acquisition and processing were performed using Chromeleon 6.80 data systems. Reaction samples were diluted in 4 volumes of ultrapure water and centrifuged (2 min, 10000 g) to remove insoluble compounds. The percentages of sucrose consumption (%Sc) were determined as follows:



$$\%Sc = \frac{[Sucrose]_{T_0} - [Sucrose]_{T_f}}{[Sucrose]_{T_0}} \times 100$$

Where  $[Sucrose]_{T_0}$  corresponds to the initial concentration of sucrose in  $\text{mol.kg}^{-1}$ , and  $[Sucrose]_{T_f}$  is the concentration of sucrose in  $\text{mol.kg}^{-1}$  at the end of the reaction.

Luteolin and its glucosylated forms were analyzed by LC-MS analysis using a DIONEX Ultima 3000 series chromatograph equipped with a Dionex UVD 340 UV/vis detector and coupled with a simple-quadrupole mass spectrometer (MSQ Plus, Thermo Scientific). HPLC analyses were performed using a reversed phase analytical column (Prontosil Eurobond, 3  $\mu\text{m}$ , 53 mm  $\times$  4.6 mm; Bischoff Chromatography, Germany) maintained at 40°C. Data acquisition and processing were performed using MSQ Plus 2.0 SP1 and Chromeleon 6.80 data systems. Samples were analyzed using the following gradient starting from 15% B to 25% B over 3 min and from 25% B to 50% B over 3.5 min (A: ultrapure water with formic acid (0.1%); B: acetonitrile with formic acid (0.1%)) at a flow rate of 0.5  $\text{mL.min}^{-1}$ . Detection of flavonoid compounds was monitored at 255 and 340 nm. The samples were diluted with DMSO (15 volumes) to solubilize luteolin. The amount of converted luteolin was calculated by summing the moles of glucosylated products identified from HPLC analysis after checking that the UV response factors of the various products were similar. Luteolin conversion was determined as the ratio of consumed Luteolin versus theoretical concentration at initial time (taking into account both the solubilized and the insoluble fraction). The HPLC system was coupled to the mass spectrometer equipped with electrospray ionization (ESI) ion source according to the following parameters: desolvation gas temperature (450°C), electrospray capillary voltage (5kV) in negative mode, source block temperature (150°C) and cone voltage (50 V; 110 V). Nitrogen was used as drying gas and nebulizing gas at flow rates approximately 80  $\text{L.h}^{-1}$ . The mass spectrometer scanned from  $m/z$  100 to 1,500.

#### V.4. Production of recombinant *NpAS* and enzyme assay

Production and purification of *NpAS* variants were performed as previously described (Potocki De Montalk et al., 1999). Since pure glutathion S-transferase- amylosucrase fusion protein possesses the same function and the same efficiency as pure *NpAS*, enzymes were purified to the *NpAS* fusion protein stage (96 kDa). All activity assays were performed at 30 °C in Tris,HCl buffer (50 mM), pH 7.0 using the dinitrosalicylic acid assay (DNS) as previously described in Potocki De Montalk et al., 1999.

## V.5. Design of Response Surface Methodology (RSM) Experiment

RSM was used to optimize luteolin production with variant I228A. Organic solvent (DMSO) concentration, luteolin amount ( $\text{mmol.L}^{-1}$  of reaction medium, taking into account soluble and insoluble fraction), and sucrose/luteolin ratio (R(S/L)) (coded X1, X2 and X3, respectively) were considered as important factors to generate a Box-Behnken design considering for each of them a low (-1) and high level (+1) values: DMSO (5%) and (10%); Luteolin ( $15 \text{ mmol.L}^{-1}$ ) and ( $25 \text{ mmol.L}^{-1}$ ); R(S/L) (15) and (25). The RSM experimental design was generated and analyzed using Minitab software (Release 16) and consisted in 15 experiments (Table S3).

Analyses of the experimental levels were performed using the response surface analysis module of Minitab software by fitting the following second order polynomial equation:

$$\hat{Y} = \beta_0 + \sum_{i=1}^I \beta_i X_i + \sum_{i=1}^I \beta_{ii} X_i^2 + \sum_i \sum_j \beta_{ij} X_i X_j$$

where  $\hat{Y}$  is the predicted response,  $I$  is the number of factors (3 in this study),  $\beta_0$  is the model constant,  $\beta_i$  is the linear coefficient associated to factor  $X_i$ ,  $\beta_{ii}$  is the quadratic coefficient associated to factor  $X_i^2$  and  $\beta_{ij}$  is the interaction coefficient between factors  $X_i$  and  $X_j$ .  $X_i$  is the factor variable in its coded form.

The following equation was used for coding the actual experimental levels of the factors in the range of (-1) to (+1):

$$X_{c,i} = \frac{[X_i - (\text{low.level} + \text{high.level})/2]}{(\text{high.level} - \text{low.level})/2}$$

With  $1 \leq i \leq I$ , where  $X_{c,i}$  is the variable expressed in coded unit.

Therefore the analysis of variance (ANOVA) table for the model, the estimates of each coefficient as well as their significance level could be provided and response surface curves could be drawn. The model also allowed determining values of each factor leading to the optimum of each criterion, corresponding to luteolin conversion and converted amount of luteolin. When optimal conditions were defined, experimental validations were performed to check consistency between the prediction and the experimental results.

## V.6. Production of luteolin glucosides (L1, L2 and L3)

Production of mono-glucosylated luteolin (L1) and di-glucosylated luteolin (L2) was carried out on preparative scale, with NpAS I228A mutant, using the optimized conditions defined from the RSM analyses. The reaction mixture (100 mL) containing Tris, HCl buffer

(50 mM; pH 7), DMSO (9%, v/v), sucrose (418 mM, 14.29 g), luteolin (19 mM, 0.65 g), and purified *NpAS* I228A mutant ( $0.8 \text{ U.mL}^{-1}$ ) was incubated at  $30^\circ\text{C}$  for 20 h yielding 342 mg and 108 mg of product L1 and L2, respectively. The mutant R226Y was preferred for L3 production. By applying the same optimized reaction culture conditions, 202 mg of L3 and 104 mg of L2 were additionally produced for characterization. Reactions were stopped by heating at  $95^\circ\text{C}$  for 10 min. The samples were concentrated by evaporation under vacuum. The luteolin glucosides were separated by preparative chromatography using Kromasil C18  $10 \mu\text{m}$  (1 kg) column, at  $200 \text{ mL.min}^{-1}$  with the following gradient starting from 5% B to 30% B over 30 min then at 100% B over 10 min (A: ultrapure water with trifluoroacetic acid (0.2%); B: acetonitrile with trifluoroacetic acid (0.2%)). Detection was monitored at 254 nm.

### V.7. Nuclear Magnetic Resonance (NMR) analysis

Luteolin, L1, L2 and L3 (10 mg) were dissolved in  $\text{CD}_3\text{OD}$  (500  $\mu\text{L}$ ) and dispensed in NMR tubes (5 mm). 1D and 2D-NMR spectra of luteolin and compounds L1 and L2 were recorded on an Avance 500 MHz spectrometer (Bruker) using 5 mm z-gradient TBI probe. The data were acquired and processed using TopSpin 3 software. The temperature was 298 K. Proton spectra were acquired by using a  $30^\circ$  pulse, 20 sweep width and 3.2 s acquisition time. A total of 16 scans were recorded and the relaxation delay between scans was 6.5 s. All NMR spectra of compound L3 were acquired using a Bruker Ascend 800 MHz spectrometer equipped with a 5 mm cryoprobe CQPCI.

### V.8. Water Solubility, UV spectra and antioxidant properties of luteoline glucosides

Luteolin glucosides, L1 and L2 (150 mg) and L3 (500 mg) were mixed with ultrapure water (1 g). The mixtures were let stand for 20 h at  $23^\circ\text{C}$  and then centrifuged to eliminate undissolved flavones. The composition of each supernatant was analyzed by HPLC in the conditions described above to determine the concentration of dissolved luteolin at saturation.

Luteolin and the glucosylated forms were dissolved at 65  $\mu\text{M}$  solutions in acetonitrile. The mixtures were filled into UV quartz cuvettes and UV spectra (220–450 nm) were recorded with a Carry 100 Bio UV–Vis Spectrometer (VARIAN, Inc.).

All DPPH (2,2-diphenyl-2-picryl-hydrazyl) free radical scavenging measurements were performed on a VersMax<sup>TM</sup> microplate reader (Molecular Devices®, California, USA), according to previously described methods (Molyneux, 2004). Daily-prepared methanolic DPPH solution (100  $\mu\text{L}$  of 400  $\mu\text{M}$ ) was mixed with luteolin, L1, L2 or L3 solutions (100  $\mu\text{L}$ ) at

different concentrations. The absorbance of the mixtures was recorded 75 min at 515 nm and 25°C. The percentage of remaining DPPH was plotted versus reaction time. Antioxidant activities were expressed by means of EC50, expressed as the molar ratio  $\text{mol}_{\text{antioxidant}}/\text{mol}_{\text{DPPH}}$ .

## V.9. Docking studies

Three-dimensional models of various mutants were generated using as template the coordinates of *N. polysaccharea* amylosucrase taken from the crystal structure with covalent  $\beta$ -glucosyl intermediate (PDB: 1S46). Mutations were introduced using the Biopolymer module of InsightII software package (Accelrys, Saint Louis, USA). Products L1, L2, and L3 were constructed starting from the covalent  $\beta$ -glucosyl moiety (PDB: 1S46) and elongated to give the entire molecular structure. Molecules were initially manually docked in the active site of the enzymes. The conformation of the mutated residue side chain was optimized by manually selecting a low-energy conformation from a side-chain rotamer library. Steric clashes (Van der Waals overlap) and non-bonded interaction energies (Coulombic and Lennard–Jones) were evaluated for the different side-chain conformations. All energy minimization calculations were performed using the CFF91 force field and (steepest descent algorithm, 10,000 iterations) (Maple et al., 1988). All drawings were performed using PyMOL software (DeLano, 2002).

## V.10. Characterization of luteolin glucosides L2 and L3

**L2:** Luteolin-4'-O- $\alpha$ -D-glucopyranosyl- $\alpha$ -(1 $\rightarrow$ 4)-glucopyranoside. MS: Anal. Calcd for C<sub>27</sub>H<sub>30</sub>O<sub>16</sub>: 610.3 [MH]<sup>+</sup>. Found: 607.6; NMR: <sup>1</sup>H and <sup>13</sup>C NMR data given in Table S2; Water solubility (23°C): 10.3 mM; UV/Vis  $\lambda_{\text{max}}$ : 250, 270, 330 nm.

**L3:** Luteolin-4'-O- $\alpha$ -D-glucopyranosyl- $\alpha$ -(1 $\rightarrow$ 4)-D-glucopyranosyl- $\alpha$ -(1 $\rightarrow$ 6)-D-glucopyranoside. LC/MS: Anal. Calcd for C<sub>33</sub>H<sub>40</sub>O<sub>21</sub>: 772.4 [MH]<sup>+</sup>. Found: 772.1; NMR: <sup>1</sup>H and <sup>13</sup>C NMR data given in Table S2; Water solubility (23°C): over 647 mM; UV/Vis  $\lambda_{\text{max}}$ : 250, 270, 330 nm.

## VI. Acknowledgements

We are grateful to MetaSys, the metabolomics & Fluxomics Center at the Laboratory for engineering of Biological Systems & Processes (Toulouse, France), for NMR experiments. We also thankfully acknowledge the assistance of S. Bozonnet to use the ICEO facility for

high-throughput screening of enzymes and the technical assistance of N. Monties for analytical experiments.

## VII. References

- Albenne, C., Potocki de Montalk, G., Monsan, P., Skov, L.K., Mirza, O., Gajhede, M., and Remaud-Simeon, M. (2002). Site-directed mutagenesis of key amino acids in the active site of amylosucrase from *Neisseria polysaccharea*. *Biol. Bratisl.* 57, 119–128.
- Albenne, C., Skov, L.K., Mirza, O., Gajhede, M., Feller, G., D'Amico, S., André, G., Potocki-Véronèse, G., Veen, B.A. van der, Monsan, P., et al. (2004). Molecular Basis of the Amylose-like Polymer Formation Catalyzed by *Neisseria polysaccharea* Amylosucrase. *J. Biol. Chem.* 279, 726–734.
- Amat, A., Sgamellotti, A., and Fantacci, S. (2008). Theoretical Study of the Structural and Electronic Properties of Luteolin and Apigenin Dyes. In *Computational Science and Its Applications – ICCSA 2008*, (Springer Berlin / Heidelberg), pp. 1141–1155.
- Bertrand, A., Morel, S., Lefoulon, F., Rolland, Y., Monsan, P., and Remaud-Simeon, M. (2006). *Leuconostoc mesenteroides* glucansucrase synthesis of flavonoid glucosides by acceptor reactions in aqueous-organic solvents. *Carbohydr. Res.* 341, 855–863.
- Cambon, E., Barbe, S., Pizzut-Serin, S., Remaud-Siméon, M., and André, I. (2014). Essential role of amino acid position 226 in oligosaccharide elongation by amylosucrase from *Neisseria polysaccharea*. *Biotechnol. Bioeng.*
- 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.
- Champion, E., Guérin, F., Moulis, C., Barbe, S., Tran, T.H., Morel, S., Descroix, K., Monsan, P., Mourey, L., Mulard, L.A., et al. (2012). Applying Pairwise Combinations of Amino Acid Mutations for Sorting Out Highly Efficient Glucosylation Tools for Chemo-Enzymatic Synthesis of Bacterial Oligosaccharides. *J. Am. Chem. Soc.* 134, 18677–18688.
- 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., et al. (2011). Biosynthesis of (+)-catechin glycosides using recombinant amylosucrase from *Deinococcus geothermalis* DSM 11300. *Enzyme Microb. Technol.* 49, 246–253.
- Cho, J.S., Yoo, S.S., Cheong, T.K., Kim, M.J., Kim, Y., and Park, K.H. (2000). Transglycosylation of neohesperidin dihydrochalcone by *Bacillus stearothermophilus* maltogenic amylase. *J. Agric. Food Chem.* 48, 152–154.
- Coutinho, P.M., Deleury, E., Davies, G.J., and Henrissat, B. (2003). An evolving hierarchical family classification for glycosyltransferases. *J. Mol. Biol.* 328, 307–317.
- Daudé, D., Champion, E., Morel, S., Guieysse, D., Remaud-Siméon, M., and André, I. (2013). Probing Substrate Promiscuity of Amylosucrase from *Neisseria polysaccharea*. *ChemCatChem* 5, 2288–2295.
- DeLano, W.L. (2002). The PyMOL Molecular Graphics System (San Carlos, CA: DeLano Scientific).
- Fischer, F., Zufferey, E., Bourgeois, J.-M., Héritier, J., and Micaux, F. (2011). UV-ABC screens of luteolin derivatives compared to edelweiss extract. *J. Photochem. Photobiol. B* 103, 8–15.
- Galleano, M., Verstraeten, S.V., Oteiza, P.I., and Fraga, C.G. (2010). Antioxidant actions of flavonoids: thermodynamic and kinetic analysis. *Arch. Biochem. Biophys.* 501, 23–30.
- Hyun, E.-K., Park, H.-Y., Kim, H.-J., Lee, J.-K., Kim, D., and Oh, D.-K. (2007). Production of epigallocatechin gallate 7-O-alpha-D-glucopyranoside (EGCG-G1) using the glucosyltransferase from *Leuconostoc mesenteroides*. *Biotechnol. Prog.* 23, 1082–1086.
- Kim, G.-E., Kang, H.-K., Seo, E.-S., Jung, S.-H., Park, J.-S., Kim, D.-H., Kim, D.-W., Ahn, S.-A., Sunwoo, C., and Kim, D. (2012). Glucosylation of the flavonoid, astragalin by *Leuconostoc mesenteroides* B-512FMCM dextranucrase acceptor reactions and characterization of the products. *Enzyme Microb. Technol.* 50, 50–56.
- Kometani, T., Nishimura, T., Nakae, T., Takii, H., and Okada, S. (1996). Synthesis of neohesperidin glycosides and naringin glycosides by cyclodextrin glucanotransferase from an alkalophilic *Bacillus* species. *Biosci. Biotechnol. Biochem.* 60, 645–649.
- Lee, S.J., Kim, J.C., Kim, M.J., Kitaoka, M., Park, C.S., Lee, S.Y., Ra, M.J., Moon, T.W., Robyt, J.F., and Park, K.H. (1999). Transglycosylation of naringin by *Bacillus stearothermophilus* Maltogenic amylase to give glycosylated naringin. *J. Agric. Food Chem.* 47, 3669–3674.

- Leemhuis, H., Pijning, T., Dobruchowska, J.M., van Leeuwen, S.S., Kralj, S., Dijkstra, B.W., and Dijkhuizen, L. (2013). Glucansucrases: Three-dimensional structures, reactions, mechanism,  $\alpha$ -glucan analysis and their implications in biotechnology and food applications. *J. Biotechnol.* 163, 250–272.
- Li, D., Park, S.-H., Shim, J.-H., Lee, H.-S., Tang, S.-Y., Park, C.-S., and Park, K.-H. (2004). In vitro enzymatic modification of puerarin to puerarin glycosides by maltogenic amylase. *Carbohydr. Res.* 339, 2789–2797.
- Lopez-Lazaro, M. (2009). Distribution and biological activities of the flavonoid luteolin. *Mini Rev. Med. Chem.* 9, 31–59.
- Maple, J.R., Dinur, U., and Hagler, A.T. (1988). Derivation of force fields for molecular mechanics and dynamics from ab initio energy surfaces. *Proc. Natl. Acad. Sci.* 85, 5350–5354.
- Miller, G.L. (1959). Use of Dinitrosalicylic Acid Reagent for Determination of Reducing Sugar. *Anal. Chem.* 31, 426–428.
- 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 (Mosc.)* 40, 9032–9039.
- Molyneux, P. (2004). The Use of the Stable Free Radical Diphenylpicrylhydrazyl (DPPH) for Estimating Antioxidant Activity. *Songklanakarin J Sci Technol* 26, 211–219.
- Palcic, M.M. (2011). Glycosyltransferases as biocatalysts. *Curr. Opin. Chem. Biol.* 15, 226–233.
- Lo Piparo, E., Scheib, H., Frei, N., Williamson, G., Grigorov, M., and Chou, C.J. (2008). Flavonoids for controlling starch digestion: structural requirements for inhibiting human  $\alpha$ -amylase. *J. Med. Chem.* 51, 3555–3561.
- Potocki De Montalk, G., 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.
- Quideau, S., Deffieux, D., Douat-Casassus, C., and Pouységu, L. (2011). Plant Polyphenols: Chemical Properties, Biological Activities, and Synthesis. *Angew. Chem. Int. Ed.* 50, 586–621.
- Studier, F.W. (2005). Protein production by auto-induction in high-density shaking cultures. *Protein Expr. Purif.* 41, 207–234.
- Tadera, K., Minami, Y., Takamatsu, K., and Matsuoka, T. (2006). Inhibition of  $\alpha$ -glucosidase and  $\alpha$ -amylase by flavonoids. *J. Nutr. Sci. Vitaminol. (Tokyo)* 52, 149–153.
- Tsimogiannis, D.I., and Oreopoulou, V. (2004). Free radical scavenging and antioxidant activity of 5,7,3',4'-hydroxy-substituted flavonoids. *Innov. Food Sci. Emerg. Technol.* 5, 523–528.
- Wang, A., Zhang, F., Huang, L., Yin, X., Li, H., Wang, Q., Zeng, Z., and Xie, T. (2010). New progress in biocatalysis and biotransformation of flavonoids. *J. Med. Plants Res.* 4, 847–856.
- Williams, L.K., Li, C., Withers, S.G., and Brayer, G.D. (2012). Order and disorder: differential structural impacts of myricetin and ethyl caffeate on human amylase, an antidiabetic target. *J. Med. Chem.* 55, 10177–10186.
- Woo, H.-J., Kang, H.-K., Nguyen, T.T.H., Kim, G.-E., Kim, Y.-M., Park, J.-S., Kim, D., Cha, J., Moon, Y.-H., Nam, S.-H., et al. (2012). Synthesis and characterization of ampelopsin glucosides using dextransucrase from *Leuconostoc mesenteroides* B-1299CB4: Glucosylation enhancing physicochemical properties. *Enzyme Microb. Technol.* 51, 311–318.

## VIII. Supplementary information

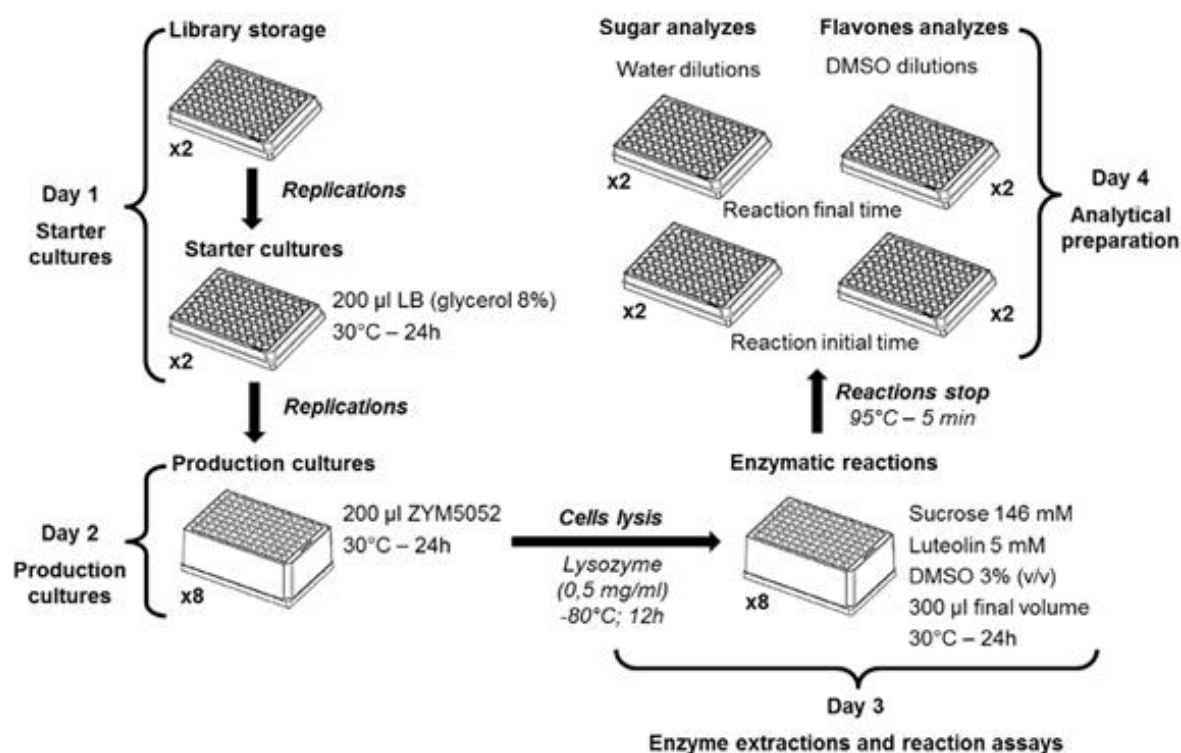


Figure S1. Method developed for screening *NpAS* single mutants for luteolin glucosylation.

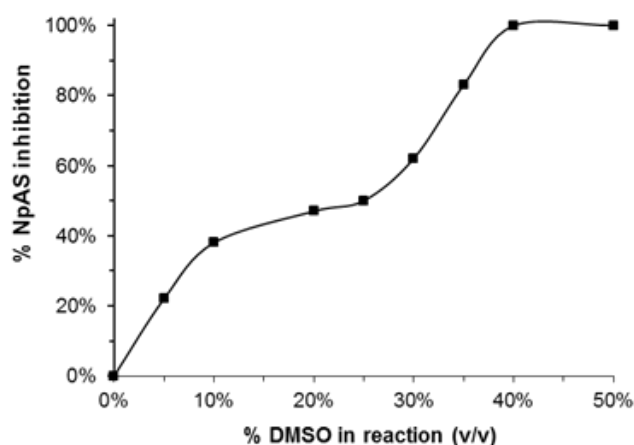
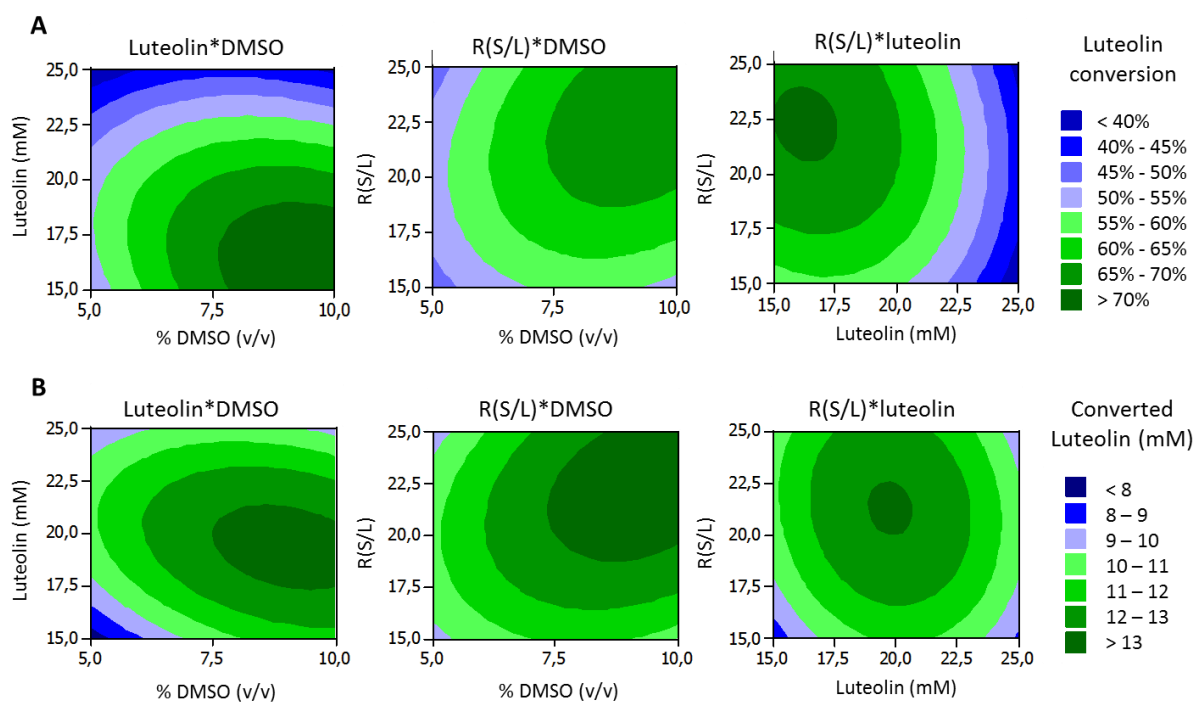


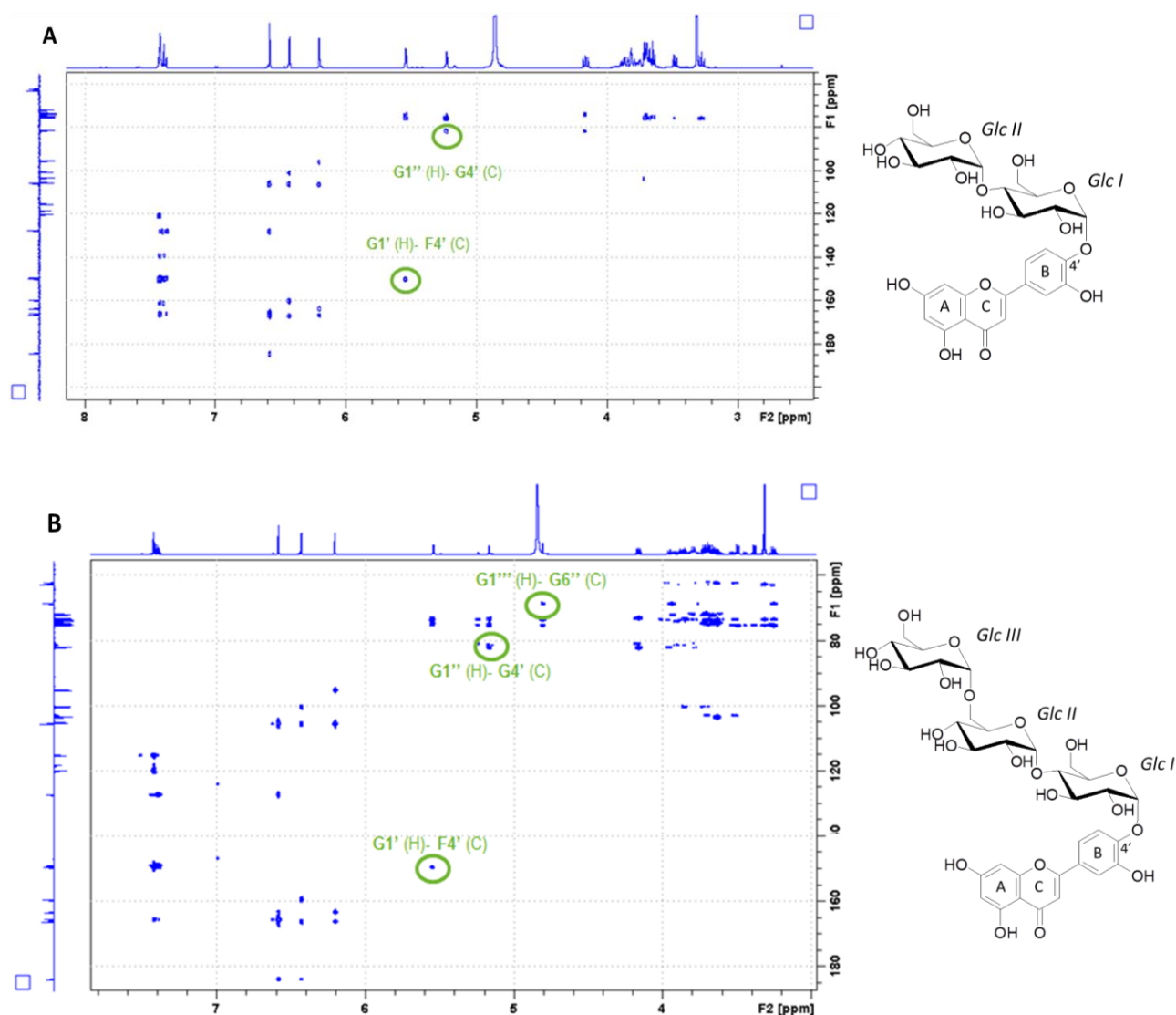
Figure S2. Inhibition of *NpAS* activity on sucrose in the presence of DMSO. Inhibition was measured using standard DNS assays and calculated according to the following formula:  $\%Inhibition_x = 100 \times (1 - (Activity\ x\% DMSO / Activity\ 0\% DMSO))$





**Figure S3. Response Surface Methodology 2D-plots showing the effect of different reaction parameters (luteolin quantity, mM; DMSO quantity, % (v/v); sucrose / luteolin molar ratio (R(S/L)), no units) added on the response: (A) luteolin conversion rate (%); (B) quantity of converted luteolin (mM).**





**Figure S4. NMR spectra of luteolin glucosides L2 and L3.** (A) HMBC NMR spectra of L2 luteolin-4'-O- $\alpha$ -D-glucopyranosyl- $\alpha$ (1 $\rightarrow$ 4)-D-glucopyranoside recorded in  $\text{CD}_3\text{OD}$  at 298 K on a Avance 500 MHz spectrometer (Bruker) (B) HMBC NMR spectra of L3 luteolin-4'-O- $\alpha$ -D-glucopyranosyl- $\alpha$ (1 $\rightarrow$ 4)-D-glucopyranosyl- $\alpha$ (1 $\rightarrow$ 6)-D-glucopyranoside recorded in  $\text{CD}_3\text{OD}$  at 298 K on a Ascend 800 MHz spectrometer (Bruker). F correspond to the flavone luteolin; G correspond to glucosyl units (" for GlcI, " for GlcII and "" for GlcIII)

**Table S1. 171 NpAS single mutant library bacteria growth and activities on sucrose only.**  
*Cultures were performed in deep-well microplates. Optical Density (OD) was measured at 600nm at the end of the 20h culture, in quadruplicates. Activities were measured in the soluble fraction of disrupted cells using standard DNS method (quadruplicates). NpAS mutants showing an activity below 60 U/L of culture (0.2 U.mL of soluble fraction) were considered as inactive.*

NpAS mutant	UOD 600nm	Activity (U.L <sup>-1</sup> of culture)	NpAS mutant	UOD 600nm	Activity (U.L <sup>-1</sup> of culture)	NpAS mutant	UOD 600nm	Activity (U.L <sup>-1</sup> of culture)
I228A	3.5 ±0.4	463 ±107	F290A	4.0 ±0.7	123 ±8	D394A	3.5 ±0.7	103 ±13
I228C	3.2 ±0.2	257 ±49	F290C	4.3 ±0.6	148 ±25	D394C	3.5 ±0.7	31 ±6
I228D	3.5 ±0.1	37 ±1	F290D	3.8 ±0.8	50 ±5	D394D	3.3 ±0.5	280 ±59
I228E	3.1 ±0.3	208 ±32	F290E	3.8 ±0.2	92 ±5	D394E	4.6 ±0.3	356 ±102
I228F	2.2 ±0.3	87 ±56	F290F	4.2 ±1.2	328 ±29	D394F	3.2 ±0.1	14 ±2
I228G	3.7 ±0.6	44 ±2	F290G	3.8 ±0.2	316 ±23	D394G	3.2 ±0.3	67 ±12
I228H	3.6 ±0.6	70 ±5	F290H	5.0 ±0.3	556 ±78	D394H	3.4 ±0.2	22 ±4
I228I	4.6 ±0.7	360 ±30	F290I	3.8 ±0.6	51 ±6	D394I	3.5 ±0.4	15 ±3
I228K	3.7 ±0.6	63 ±7	F290K	4.4 ±0.5	70 ±6	D394K	3.2 ±0.4	18 ±3
I228L	3.4 ±0.2	148 ±20	F290L	3.7 ±0.8	82 ±13	D394L	3.1 ±0.3	18 ±11
I228M	3.8 ±0.2	202 ±19	F290M	3.9 ±0.3	138 ±27	D394M	3.2 ±0.3	14 ±4
I228N	3.8 ±1.0	169 ±37	F290N	4.2 ±0.8	32 ±8	D394N	3.5 ±0.4	29 ±6
I228P	5.1 ±0.4	356 ±52	F290P	5.0 ±0.3	189 ±53	D394P	3.9 ±0.1	96 ±17
I228Q	4.3 ±0.4	443 ±24	F290Q	4.5 ±0.3	249 ±57	D394Q	3.9 ±0.1	98 ±16
I228R	4.7 ±0.1	173 ±13	F290R	4.6 ±0.1	189 ±34	D394R	4.0 ±0.2	86 ±9
I228S	4.5 ±0.1	245 ±55	F290S	4.5 ±0.4	221 ±21	D394S	4.2 ±0.2	111 ±8
I228T	4.3 ±0.1	253 ±27	F290T	4.8 ±0.2	336 ±31	D394T	3.8 ±0.2	86 ±2
I228V	4.5 ±0.2	239 ±79	F290V	4.8 ±0.1	156 ±22	D394V	4.0 ±0.2	75 ±9
I228W	4.6 ±0.2	151 ±25	F290W	4.9 ±0.2	302 ±49	D394W	3.7 ±0.4	79 ±11
I228Y	4.5 ±0.3	164 ±23	F290Y	4.8 ±0.3	213 ±23	D394Y	4.0 ±0.1	85 ±24
F229A	3.5 ±0.8	77 ±9	I330A	3.9 ±1.0	51 ±10	R446A	3.5 ±0.2	129 ±15
F229C	3.6 ±1.3	105 ±17	I330C	2.9 ±0.5	69 ±14	R446C	3.7 ±0.3	223 ±32
F229D	3.7 ±1.1	25 ±4	I330D	4.3 ±0.4	38 ±6	R446D	3.3 ±0.3	45 ±10
F229E	2.5 ±0.2	18 ±1	I330E	4.4 ±1.1	37 ±4	R446E	3.4 ±0.4	21 ±3
F229F	3.1 ±0.5	321 ±64	I330F	3.6 ±0.3	22 ±1	R446F	3.0 ±0.3	55 ±10
F229G	3.9 ±1.1	15 ±1	I330G	4.5 ±0.4	38 ±8	R446G	3.6 ±0.4	105 ±16
F229H	3.3 ±0.5	23 ±1	I330H	4.2 ±0.5	36 ±4	R446H	2.8 ±0.2	16 ±3
F229I	3.6 ±0.5	66 ±8	I330I	3.7 ±0.4	297 ±19	R446I	3.8 ±0.6	31 ±6
F229K	3.6 ±0.3	27 ±3	I330K	4.3 ±0.9	28 ±6	R446K	3.9 ±0.8	55 ±7

F229L	4.2 $\pm$ 0.7	70 $\pm$ 9	I330L	3.2 $\pm$ 0.2	11 $\pm$ 2	R446L	3.5 $\pm$ 0.7	51 $\pm$ 10
F229M	3.1 $\pm$ 0.6	354 $\pm$ 49	I330M	3.6 $\pm$ 0.2	82 $\pm$ 28	R446M	4.3 $\pm$ 0.9	92 $\pm$ 20
F229N	3.4 $\pm$ 0.2	100 $\pm$ 17	I330N	4.0 $\pm$ 0.8	58 $\pm$ 12	R446N	4.5 $\pm$ 0.9	115 $\pm$ 29
F229P	4.2 $\pm$ 0.3	214 $\pm$ 51	I330P	5.7 $\pm$ 0.2	150 $\pm$ 37	R446P	4.0 $\pm$ 0.1	136 $\pm$ 11
F229Q	4.4 $\pm$ 0.4	158 $\pm$ 15	I330Q	4.2 $\pm$ 0.2	117 $\pm$ 7	R446Q	4.3 $\pm$ 0.1	179 $\pm$ 27
F229R	4.4 $\pm$ 0.1	95 $\pm$ 12	I330R	4.4 $\pm$ 0.3	59 $\pm$ 12	R446R	4.3 $\pm$ 0.2	220 $\pm$ 38
F229S	4.4 $\pm$ 0.1	90 $\pm$ 6	I330S	4.8 $\pm$ 0.4	92 $\pm$ 14	R446S	0.4 $\pm$ 0.0	44 $\pm$ 11
F229T	4.5 $\pm$ 0.2	90 $\pm$ 11	I330T	4.4 $\pm$ 0.5	106 $\pm$ 8	R446T	4.4 $\pm$ 2.3	128 $\pm$ 17
F229V	4.3 $\pm$ 0.2	170 $\pm$ 44	I330V	4.4 $\pm$ 0.3	167 $\pm$ 16	R446V	4.3 $\pm$ 0.2	109 $\pm$ 9
F229W	4.2 $\pm$ 0.2	168 $\pm$ 43	I330W	4.6 $\pm$ 0.6	118 $\pm$ 17	R446W	4.5 $\pm$ 0.0	136 $\pm$ 15
F229Y	4.3 $\pm$ 0.3	243 $\pm$ 31	I330Y	5.3 $\pm$ 0.2	144 $\pm$ 15	R446Y	4.0 $\pm$ 0.1	145 $\pm$ 46
A289A	3.7 $\pm$ 0.9	410 $\pm$ 31	V331A	3.7 $\pm$ 0.4	429 $\pm$ 19	R226A	0.4 $\pm$ 0.0	55 $\pm$ 6
A289C	3.3 $\pm$ 0.4	539 $\pm$ 81	V331C	4.0 $\pm$ 0.3	738 $\pm$ 51	R226C	4.4 $\pm$ 0.2	500 $\pm$ 107
A289D	3.4 $\pm$ 0.4	104 $\pm$ 10	V331D	3.7 $\pm$ 0.3	198 $\pm$ 31	R226D	4.6 $\pm$ 0.2	372 $\pm$ 26
A289E	3.9 $\pm$ 0.4	71 $\pm$ 6	V331E	4.2 $\pm$ 0.9	391 $\pm$ 60	R226E	3.5 $\pm$ 0.1	253 $\pm$ 22
A289F	3.2 $\pm$ 0.2	25 $\pm$ 2	V331F	3.8 $\pm$ 0.8	137 $\pm$ 51	R226F	4.2 $\pm$ 0.2	420 $\pm$ 74
A289G	4.3 $\pm$ 0.8	285 $\pm$ 50	V331G	3.6 $\pm$ 0.4	324 $\pm$ 62	R226G	4.2 $\pm$ 0.2	361 $\pm$ 47
A289H	3.2 $\pm$ 0.4	19 $\pm$ 2	V331H	3.6 $\pm$ 0.4	207 $\pm$ 41	R226H	3.9 $\pm$ 0.1	617 $\pm$ 55
A289I	3.9 $\pm$ 0.6	188 $\pm$ 14	V331I	3.8 $\pm$ 1.1	132 $\pm$ 57	R226I	4.6 $\pm$ 0.3	458 $\pm$ 40
A289K	3.6 $\pm$ 1.1	26 $\pm$ 1	V331K	3.9 $\pm$ 0.4	130 $\pm$ 62	R226K	4.4 $\pm$ 0.1	382 $\pm$ 70
A289L	3.5 $\pm$ 0.5	27 $\pm$ 2	V331L	3.2 $\pm$ 0.5	190 $\pm$ 35	R226L	4.9 $\pm$ 0.2	342 $\pm$ 62
A289M	4.0 $\pm$ 0.4	157 $\pm$ 26	V331M	3.6 $\pm$ 0.4	155 $\pm$ 22	R226M	5.0 $\pm$ 0.2	255 $\pm$ 42
A289N	4.3 $\pm$ 1.0	908 $\pm$ 187	V331N	3.5 $\pm$ 0.3	276 $\pm$ 25	R226N	4.2 $\pm$ 0.2	572 $\pm$ 98
A289P	4.4 $\pm$ 0.1	1000 $\pm$ 231	V331P	4.2 $\pm$ 0.2	155 $\pm$ 35	R226P	4.5 $\pm$ 0.2	546 $\pm$ 153
A289Q	4.4 $\pm$ 0.2	299 $\pm$ 52	V331Q	4.1 $\pm$ 0.1	148 $\pm$ 18	R226Q	3.9 $\pm$ 0.2	296 $\pm$ 68
A289R	4.6 $\pm$ 0.1	87 $\pm$ 17	V331R	4.5 $\pm$ 0.1	122 $\pm$ 15	R226R	3.9 $\pm$ 0.3	219 $\pm$ 14
A289S	4.5 $\pm$ 0.2	319 $\pm$ 44	V331S	4.4 $\pm$ 0.2	148 $\pm$ 12	R226S	4.3 $\pm$ 0.3	484 $\pm$ 78
A289T	4.6 $\pm$ 0.3	354 $\pm$ 27	V331T	4.3 $\pm$ 0.3	228 $\pm$ 59	R226T	4.6 $\pm$ 0.3	360 $\pm$ 90
A289V	4.6 $\pm$ 0.2	311 $\pm$ 40	V331V	4.3 $\pm$ 0.1	178 $\pm$ 18	R226V	4.7 $\pm$ 0.2	310 $\pm$ 18
A289W	4.8 $\pm$ 0.1	176 $\pm$ 18	V331W	4.3 $\pm$ 0.2	151 $\pm$ 28	R226W	4.2 $\pm$ 0.1	634 $\pm$ 79
A289Y	4.8 $\pm$ 0.1	92 $\pm$ 14	V331Y	4.3 $\pm$ 0.1	185 $\pm$ 13	R226Y	4.2 $\pm$ 0.2	444 $\pm$ 64

**Table S2.**  $^{13}\text{C}$  and  $^1\text{H}$  NMR data for luteolin, L1, L2 and L3. (L1) luteolin-4'-O- $\alpha$ -D-glucopyranoside, (L2) luteolin-4'-O- $\alpha$ -D-glucopyranosyl- $\alpha$ (1 $\rightarrow$ 4)-D-glucopyranoside and (L3) luteolin-4'-O- $\alpha$ -D-glucopyranosyl- $\alpha$ (1 $\rightarrow$ 4)-D-glucopyranosyl- $\alpha$ (1 $\rightarrow$ 6)-D-glucopyranoside.

	Luteolin (in DMSO-d <sub>6</sub> )		L1 (in DMSO-d <sub>6</sub> )		L2 (in CD <sub>3</sub> OD)		L3 (in CD <sub>3</sub> OD)	
	$\delta_{\text{C}}$	$\delta_{\text{H}}$	$\delta_{\text{C}}$	$\delta_{\text{H}}$	$\delta_{\text{C}}$	$\delta_{\text{H}}$	$\delta_{\text{C}}$	$\delta_{\text{H}}$
2	163.9		163.7		165.48		165.46	
3	102.9	6.65 (s)	104.5	6.82 (s)	105.06	6.58 (s)	105.02	6.57 (s)
4	181.7		182.2		183.81		183.80	
5	161.5		161.9		163.33		163.22	
6	98.9	6.18 (d, 2.1)	99.4	6.21 (s)	100.29	6.20 (d, 2.1 Hz)	100.21	6.20 (d 2.1 Hz)
7	164.1		164.7		166.11		166.11	
8	93.9	6.43 (d, 2.1)	94.5	6.51 (s)	95.05	6.43 (d, 2.1 Hz)	95.04	6.43 (d 2.1 Hz)
9	157.3		157.8		159.40		159.37	
10	103.8		104.3		105.06		105.40	
1'	121.6		125.3		127.23		127.15	
2'	113.4	7.39 (d, 2.2)	114.0	7.50 (s)	115.10	7.43 (d, 1.8 Hz)	115.07	7.42 (d, 2.28 Hz)
3'	145.8		147.9		149.00		148.94	
4'	149.7		148.7		149.49		149.46	
5'	116.1	6.89 (d, 9.0)	117.2	7.35 (d, 8.3)	118.10	7.38 (d, 9.2 Hz)	118.03	7.39 (d 8.54 Hz)
6'	119.0	7.40 (dd, 9.0,2.2)	119.0	7.52 (d, 8.3)	119.73	7.42 (dd, 9.2 Hz, 2.1 Hz)	119.72	7.415 (dd, 8.54, 2.28 Hz)
1''			99.7	5.43 (d, 3.7)	100.22	5.51 (d 3.6 Hz)	100.15	5.53 (d 3.6 Hz)
2''			72.4	3.39 (m)	74.82	3.66 (dd, 9.3 Hz, 3.6 Hz)	72.91	3.72 (3.6 Hz, 9.9 Hz)
3''			73.6	3.73 (m)	74.27	4.12 (t, 9.3 Hz)	74.53	4.16 (9 Hz, 9.8 Hz)
4''			70.4	3.20 (m)	80.90	3.66 (d, 9.3 Hz)	81.91	3.63 (9 Hz, 9.80 Hz)
5''			74.4	3.51 (m)	73.42	3.72 (m)	73.47	3.78 (2.28, 4.80, 9)
6''			61.2	3.61 (m)	61.77	3.84 (dd, 12Hz, 4.2 Hz)	62.06	3.95 (2.04, 12.12)
6''				3.48 (m)		3.76 (dd, 12Hz, 1,8 Hz)		3.87 (4.80, 12.16)
1'''					102.93	5.19 (d, 3.6 Hz)	103.27	5.16 (d 3.84 Hz)
2'''					74.62	3.44 (dd, 9,3 Hz, 3,6 Hz)	74.25	3.49 (3.76, 9.68)
3'''					75.10	3.59 (d, 9 Hz)	75.14	3.66 (8.80, 9.76)
4'''					71.56	3.24 (t, 9,3 Hz)	71.97	3.25 (9.04, 10.00)
5'''					72.99	3.64 (m)	73.32	3.93 (2.1, 6.94, 10)
6'''					62.76	3.79 (dd, 11,7 Hz, 1,8 Hz)	68.45	3.73 (2.00, 10.64)
6'''						3.61 (dd, 12 Hz, 5,4Hz)		3.85 (7.00, 10.7)
1''''							100.05	4.80 (3.68 Hz)
2''''							73.66	3.38 (3.68, 9.76)
3''''							75.03	3.70 (9.12, 9.52)
4''''							71.73	3.31 (9.00, 9.56)
5''''							73.40	3.68
6''''							62.56	3.68
6''''								3.79

**Table S3. RSM and results of the Box-Behnken design with the three variables studied.** Multiple regression analysis polynomial equation with coefficients of each factor is also given for luteolin conversion rate (%) response and amount of converted luteolin (mM) response.

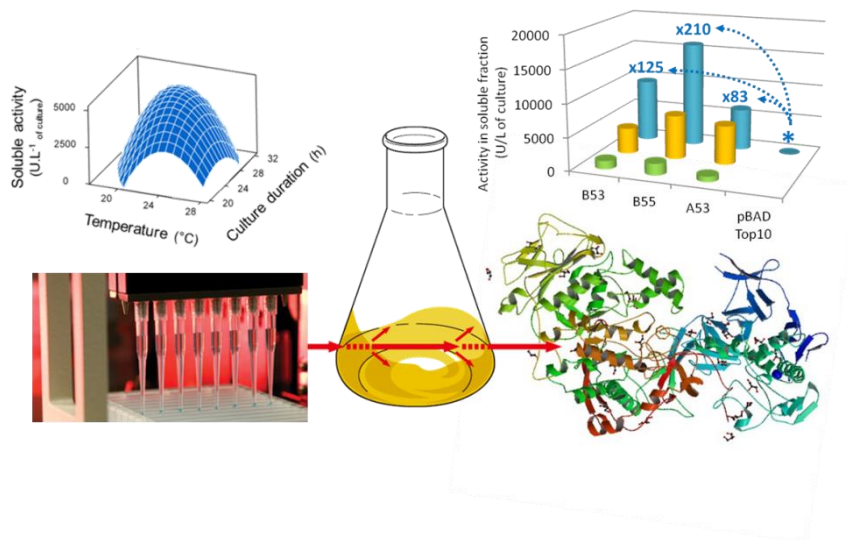
Standard order	% DMSO (v/v) ( $X_1$ )	Luteolin (mM) ( $X_2$ )	R(S/L) ( $X_3$ )	Luteolin conversion (%) <sup>[a]</sup>	Converted luteolin (mM) <sup>[b]</sup>
1	5	15	20	49.1%	5.3
2	10	15	20	64.9%	7.8
3	5	25	20	37.8%	7.9
4	10	25	20	32.3%	7.0
5	5	20	15	37.6%	5.8
6	10	20	15	51.7%	8.5
7	5	20	25	40.4%	6.9
8	10	20	25	69.8%	11.0
9	7.5	15	15	57.1%	6.6
10	7.5	25	15	31.7%	6.5
11	7.5	15	25	65.6%	8.0
12	7.5	25	25	30.4%	6.5
13	7.5	20	20	62.4%	9.8
14	7.5	20	20	61.0%	9.7
15	7.5	20	20	60.8%	9.6

[a] Luteolin conversion =  $-2.68421 + 0.189605X_1 + 0.176546X_2 + 0.0952530X_3 - 0.00919010X_1^2 - 0.00380782X_2^2 - 0.00227886X_3^2 - 0.00418049X_1X_2 + 0.00290137X_1X_3 - 0.000951113X_2X_3$

[b] Luteolin converted =  $-71.3517 + 3.29548X_1 + 4.93199X_2 + 1.99097X_3 - 0.167051X_1^2 - 0.102299X_2^2 - 0.0497648X_3^2 - 0.0739547X_1X_2 + 0.0580275X_1X_3 - 0.0155371X_2X_3$

# Chapter III

## Optimizing the production of an $\alpha$ -(1 $\rightarrow$ 2) branching sucrase in *Escherichia coli* using statistical design



Yannick Malbert <sup>‡</sup>, Marlène Vuillemin <sup>‡</sup>, Sandrine Laguerre, Magali Remaud-Siméon and Claire Moulis

<sup>‡</sup> Both authors contributed equally

Appl Microbiol Biotechnol. (2014)



---

The results described in the preceding chapter have underlined that a structurally-guided approach was efficient to generate promiscuous variants improved for luteolin glucosylation.

To further explore the promiscuity of this *NpAS* library, we decided to attempt the glucosylation of two other flavonoids, diosmetin and quercetin, which are very poorly recognized by the wild-type *NpAS* (conversion rates of 0% and 0.6%, respectively). Unfortunately, none of the *NpAS* mutants were able to glucosylate diosmetin. The best *NpAS* mutant (R226I) only converted 16% of quercetin. This represents a considerable improvement but was not sufficient to ensure a good production level of glucosides.

Another sucrose-utilizing transglucosylase, the  $\alpha$ -(1→2) branching sucrase was preferred for the second part of our work. In the presence of luteolin, quercetin or diosmetin acceptors, this enzyme showed interesting conversion rates around 40%, 60% and 2%, respectively, making it a good candidate for evolution. This protein also displayed other interesting features:

- $\alpha$ -(1→2) linkage specificity different from that of *NpAS*.
- Naturally design for acceptor reaction
- Less hindered active site than that of *NpAS*, which might be beneficial for flavonoid accommodation.
- Higher stability in DMSO than that of *NpAS*

However,  $\alpha$ -(1→2) branching sucrase engineering had never been previously attempted and the recombinant enzyme was almost exclusively produced in inclusion bodies. This was not compatible with the implementation of a screening assay. Thus, a prerequisite to enzyme engineering was the optimization of the  $\alpha$ -(1→2) branching sucrase heterologous production in *E. coli*. To achieve this goal, we first constructed and screened different expression systems. The best hits enzyme productions were further optimized using experimental design and Response Surface Methodology, whose results are described in the following chapter.

---



## I. Introduction

*Escherichia coli* is the most commonly used host for heterologous gene expression (Gordon et al., 2008; Koehn and Hunt, 2009). Despite the deep and continuously increasing knowledge about *E. coli* genetics and physiology, overproduction of some recombinant proteins is still a challenge. In particular, the formation of insoluble protein aggregates as inclusion bodies is reported to be the major phenomenon limiting heterologous expression. Such aggregates are often made up from recombinant protein misfolding, but the mechanism behind their formation remains unclear (Baneyx, 1999; Upadhyay et al., 2012; Villaverde and Carrió, 2003)

Numerous *E. coli* strains and/or expression vectors were specifically designed to improve protein overexpression and solubility, leading sometimes to significant success (Jana and Deb, 2005; Terpe, 2003; Waugh, 2005). However, improvements are highly protein-dependent and none of these systems can be considered as universal enhancers of protein expression or solubility. Decreasing culture temperature was also described as an efficient way to diminish protein aggregation (Song et al., 2012). Aside, auto-induction methods of protein expression based on diauxic growth was demonstrated to be quite efficient to improve expression, as they allow a tightly controlled induction by host metabolism while minimizing culture handling (Blommel et al., 2008; Studier, 2009; Tyler et al., 2005).

Lots of combinations may be tested to find the best conditions for protein overexpression and the statistical design of experiments is an efficient mean to identify the most influencing factors, and provide reliable predictions of production levels (Baş and Boyacı, 2007; Nawani and Kapadnis, 2005; Papanephoytu and Kontopidis, 2012; Swalley et al., 2006).

Our case study concerns an  $\alpha$ -(1→2) branching sucrose named GBD-CD2, a truncated form of the glucansucrose DSR-E isolated from *Leuconostoc citreum* B-1299. This enzyme belongs to the glycoside-hydrolase family 70 (Lombard et al., 2013) and exclusively catalyzes the formation of  $\alpha$ -(1→2) glucosyl ramifications onto  $\alpha$ -(1→6)-linked glucan acceptors (Brison et al., 2009; Fabre et al., 2004). Due to its unique  $\alpha$ -(1→2) linkage specificity, GBD-CD2 holds a great potential for the production of novel functional ingredients with attractive prebiotic properties (Sarbin et al., 2011, 2013; Serino et al., 2011). Recently, the 3D-structure of a N-terminally reduced variant named  $\Delta N_{123}$ -GBD-CD2 (123 kDa) was solved at 1.90 Å (Brison et al., 2012). This truncated protein was mainly produced in inclusion bodies by *E. coli*, the level of soluble enzyme activity reaching only 76 U.L<sup>-1</sup> of culture.

Time-consuming and tedious steps of denaturation/renaturation of the aggregated enzyme were required to finally recover 2300 U.L<sup>-1</sup> of culture (Brison et al., 2009). Such a mode of enzyme production is inappropriate for high-scale productions, as those in microplate format required for structure-function relationship studies or enzyme engineering approaches. To improve the soluble expression of this  $\alpha$ -(1→2) branching sucrose, it was challenging to test a Response Surface Methodology (RSM).

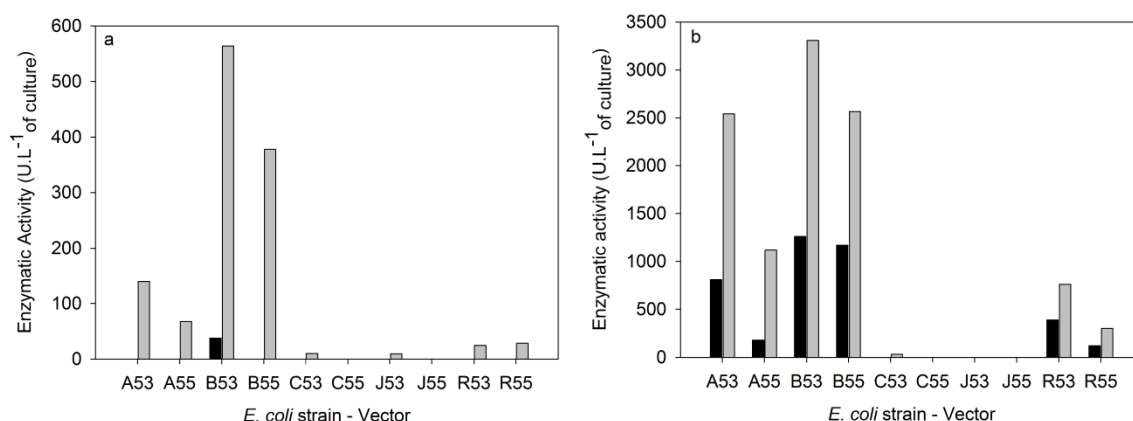
In the present study, we first focused our efforts on a pre-selection of suitable bacterial host/expression vector combinations. A particular attention was paid on the choice of small affinity tags located at both N and C-terminal ends of the target enzyme. Screening of culture conditions (including temperature, culture duration, auto-induction conditions and controlled addition of glycerol) and evaluation of their impact on soluble enzyme production were then performed through RSM. Remarkably improved yields of enzymes were reached, that will greatly facilitate future studies and uses of this branching sucrose.

## II. Results

### II.1. Pre-selection of the bacterial host and expression vector

The gene coding for the  $\alpha$ -(1→2) branching sucrose  $\Delta N_{123}$ -GBD-CD2 was cloned into two different expression vectors (pET-53 or pET-55) under the control of the T7 promoter, and in fusion with a 6xHis tag and a Strep tag II encoding sequences at 5' and 3'-ends, respectively in pET-53, or *vice versa* in pET-55. Four different *E. coli* strains (ArcticExpress DE3, BL21 Star DE3, C41 DE3 and Rosetta DE3) were tested as hosts and grown in 96-Deep Well format microplates at 20 °C for 14 h or 22 h.

For all combinations, a higher level of enzymatic activity was recovered after 22 h culture (Figure 50). In all cases, the activity values were at least twice higher in the crude extract obtained after sonication than those measured in the soluble fraction (crude extract supernatant) only. This indicates that most of the enzyme was still produced in inclusion bodies. The three best combinations for soluble enzyme production were ArcticExpress DE3 and BL21 Star DE3 strains transformed with pET-53, named A53 and B53 respectively, or BL21 Star DE3 strain transformed with pET-55, named B55. They achieved soluble activity production yields of 810, 1260 and 1120 U.L<sup>-1</sup> of culture respectively. These values are at least ten times higher than that reported by Brison et al, which was only of 76 U.L<sup>-1</sup> of culture using TOP10 *E. coli* cells and pBAD expression vector (Brison et al., 2009). Those three expression systems were selected for further optimization.



**Figure 50. Selection of the best expression systems.** Measured activity (U.L<sup>-1</sup> of culture) in soluble fraction (black columns) and crude extract (grey columns) in the different expression systems (*E. coli* host strains: A for ArcticExpress DE3, B for BL21 Star DE3, C for C41 DE3 and R for Rosetta DE3. Expression vectors: 53 for pET-53 and 55 for pET-55) for auto-induction time of: (a) 14 hours and (b) 22 hours

## II.2. Optimizing the production of soluble $\Delta N_{123}$ -GBD-CD2

For each selected expression system, the optimization of the soluble enzyme production was carried out using a Response Surface Methodology (RSM). A Box-Behnken was designed to investigate the effects of five factors on soluble enzyme production *i.e.* culture duration, temperature, as well as inducer ( $\alpha$ -lactose), repressor (glucose), and glycerol concentrations. A preliminary set of experiments was carried out to determine the boundaries of the glycerol,  $\alpha$ -lactose and glucose concentrations of this Box-Behnken design.

### II.2.1. Model fitting and statistical analysis

The results of the second-order response surface model fitting in the form of ANOVA and p-values of each linear, quadratic, or interaction terms are reported in Online Resource 1, table S6.

For A53 expression system, linear terms glycerol (G), culture temperature (T) and duration (D) showed a p-value lower than 5%, indicating that those three factors influence significantly the production of soluble enzyme. Quadratic terms were also statistically significant, except for repressor (Re<sup>2</sup>) and glycerol (G<sup>2</sup>). Three interaction terms, *i.e.* repressor\*glycerol (Re\*G), repressor\*culture duration (Re\*D) and temperature\*duration (T\*D) exhibited a significant influence on enzyme production.

Regarding B53 expression system, only T linear term (temperature of culture) had a significant impact on the production of soluble enzyme. Its quadratic effect was also statistically significant (p-value < 0.000). No other quadratic effect was significant.

Inducer\*repressor (I\*Re) and repressor\*culture duration (Re\*D) were the only two cross-product terms influencing the production of soluble enzyme.

Analysis of RSM for B55 system revealed that only the linear terms T and D (i.e., temperature and time of culture, respectively) exhibited low p-values. Temperature ( $T^2$ ) and culture duration ( $D^2$ ) quadratic terms were statistically relevant. The interaction terms, inducer\*repressor (I\*Re), glycerol\*culture duration (G\*D) and temperature\*culture duration ( $T^*D$ ), had the largest effect on the soluble enzyme production.

The three models were simplified to keep only statistically significant factors (Table 12). Then, the values of the soluble activity were fitted to the following second-order polynomial equations **(P1)**, **(P2)** and **(P3)**, respectively for A53, B53 and B55 expression systems.

**Table 12. Response surface quadratic model analysis of  $\Delta N_{123}$ -GBD-CD2 soluble activity level obtained with A53, B53 and B55 productions, after model simplification.**

a. Polynomial function regression coefficients

Factors <sup>a</sup>	Regression coefficients			Standard error			p-value <sup>b</sup>		
	A53	B53	B55	A53	B53	B55	A53	B53	B55
Constant	-1.64*10 <sup>5</sup>	-5.85*10 <sup>4</sup>	-1.90*10 <sup>5</sup>	7385.62	1452	7910.9	0.000	0.000	0.000
I	2855	-597.6	-601	915.17	466.9	381.4	0.002	0.203	0.117
Re	1.10*10 <sup>5</sup>	-5.34*10 <sup>4</sup>	-7.08*10 <sup>4</sup>	2.64*10 <sup>4</sup>	2.04*10 <sup>4</sup>	2.06*10 <sup>4</sup>	0.000	0.010	0.001
G	420	-70.1	-3022	293.78	103.5	1397.2	0.156	0.499	0.033
T	9367	5438.3	1.21*10 <sup>4</sup>	354.91	117.6	359.7	0.000	0.000	0.000
D	4294	-16.0	4529	283.45	20	313.2	0.000	0.425	0.000
I <sup>2</sup>	-994	-	-	362.63	-	-	0.007	-	-
Re <sup>2</sup>	-	-	-	-	-	-	-	-	-
G <sup>2</sup>	-	-	-	-	-	-	-	-	-
T <sup>2</sup>	-164	-118.4	-210	4.48	2.5	4.5	0.000	0.000	0.000
D <sup>2</sup>	-49	-	-47	2.85	-	3.2	0.000	-	0.000
I*Re	-	3.58*10 <sup>4</sup>	3.32*10 <sup>4</sup>	-	1.64*10 <sup>4</sup>	1.35*10 <sup>4</sup>	-	0.031	0.015
I*G	-	-	-	-	-	-	-	-	-
I*T	-	-	-	-	-	-	-	-	-
I*D	-	-	-	-	-	-	-	-	-
Re*G	-4.49*10 <sup>4</sup>	-	-	1.05*10 <sup>4</sup>	-	-	0.000	-	-
Re*T	-	-	-	-	-	-	-	-	-
Re*D	-3344	1383.9	-	923.5	716.2	-	0.000	0.056 <sup>c</sup>	-
G*T	-	-	-	-	-	-	-	-	-
G*D	-	-	105	-	-	53.4	-	-	0.052 <sup>c</sup>
T*D	-66	-	-94	8.04	-	8.3	0.000	-	0.000

<sup>a</sup> Factors coded as following : Inducer (I), repressor (Re), glycerol (G), temperature (T), culture duration (D).

<sup>b</sup> p-value < 0.05 indicates a statistical significance of the factor.

<sup>c</sup> p-value very closed to 0.05 leading us to consider the corresponding factor as statistically significant.

## b. Analysis of variance (ANOVA)

Source	Sum of Square			Degree of Freedom			F-value			p-value <sup>b</sup>		
	A53	B53	B55	A53	B53	B55	A53	B53	B55	A53	B53	B55
Model	3.31*10 <sup>8</sup>	2.54*10 <sup>8</sup>	6.13*10 <sup>8</sup>	11	8	10	158.85	276.73	241.67	0.000	0.000	0.000
Residual	1.74*10 <sup>7</sup>	1.33*10 <sup>7</sup>	2.79*10 <sup>7</sup>	92	116	110	-	-	-	-	-	-
Lack of fit	8.59*10 <sup>6</sup>	2.57*10 <sup>6</sup>	1.08*10 <sup>7</sup>	26	31	28	2.46	0.66	1.84	0.002	0.907	0.018
Pure error	8.86*10 <sup>6</sup>	1.07*10 <sup>7</sup>	1.72*10 <sup>7</sup>	66	85	82	-	-	-	-	-	-
Total	3.49*10 <sup>8</sup>	2.68*10 <sup>8</sup>	6.41*10 <sup>8</sup>	103	124	120	-	-	-	-	-	-
R <sup>2</sup>	0.950	0.950	0.957	-	-	-	-	-	-	-	-	-
Adj R <sup>2</sup>	0.944	0.947	0.953	-	-	-	-	-	-	-	-	-

<sup>b</sup> p-value < 0.05 indicates a statistical significance of the factor.

(P1) Soluble activity ( $U \cdot L^{-1}_{of\ culture}$ ) =  $-1.64 \times 10^5 + 2855 \times I + 1.10 \times 10^5 \times Re + 420 \times G + 9367 \times T + 4294 \times D - 992 \times I^2 - 164 \times T^2 - 49 \times D^2 - 4.49 \times 10^4 \times Re * G - 3344 \times Re * D - 66 \times T * D$

(P2) Soluble activity ( $U \cdot L^{-1}_{of\ culture}$ ) =  $-5.85 \times 10^4 - 597.6 \times I - 5.34 \times 10^4 \times Re - 70.1 \times G + 5438.3 \times T - 16 \times D - 118.4 \times T^2 + 43.58 \times 10^4 \times Re * I + 1383.9 \times Re * D$

(P3) Soluble activity ( $U \cdot L^{-1}_{of\ culture}$ ) =  $-1.90 \times 10^5 - 601 \times I - 7.08 \times 10^4 \times Re - 3022 \times G + 1.21 \times 10^4 \times T + 4529 \times D - 47 \times D^2 + 3.32 \times 10^4 \times Re * I + 105 \times G * D - 94 \times T * D$

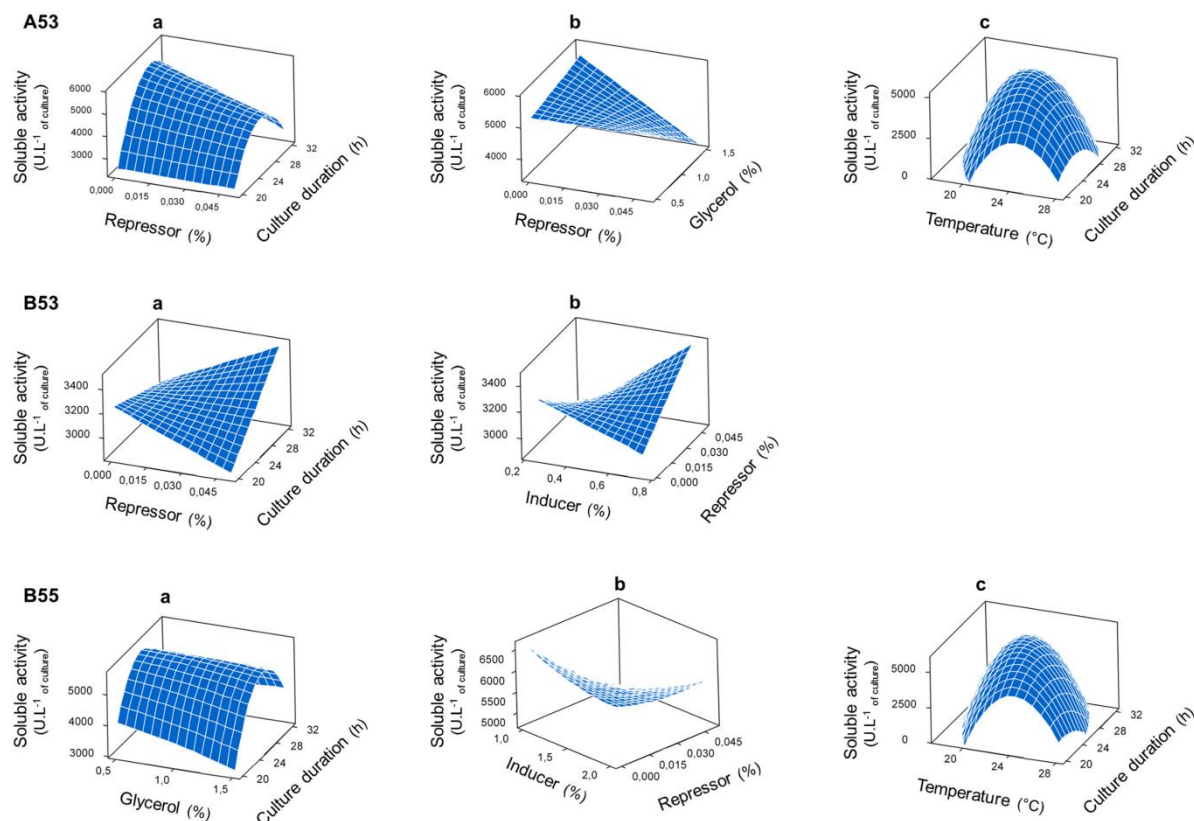
R<sup>2</sup> coefficients of 95%, 95% and 96%, and adjusted R<sup>2</sup> coefficients of 94%, 95% and 95% were evaluated for the systems A53, B53 and B55, respectively. They revealed a satisfactory correlation between the experimental values and the predicted ones.

### II.2.2. 3D response surface plots and experimental validation

Effects of cross-term products factors on the soluble activity level were further examined using the 3D response surface plots (Figure 51).

Regarding A53, temperature and culture duration are critical factors that have to be tightly adjusted to reach a maximum production at 23 °C and 28 h (Figure 51, A53c). Curiously, the soluble activity values are maximal in the absence of repressor (Figure 51, A53a and A53b). Surface plot as function of repressor and culture duration confirmed the 28 h optimal culture duration (Figure 51, A53a). Finally, soluble activity values were slightly increased with addition of glycerol up to 1.25% (w/v) (Figure 51, A53b). According to the model, a predicted production of 5810 U of soluble enzyme per liter of culture could be obtained with optimal conditions fixed at 1.40% (w/v) of  $\alpha$ -lactose, 1.25% (w/v) of glycerol, no addition of glucose, a temperature of 23 °C and a culture duration of 28 hours. Experimentally, 4800  $U \cdot L^{-1}_{of\ culture}$  of soluble enzyme activity was recovered. This is in accordance with the predicted value, considering that the 17% standard deviation calculated here is usually observed during microtiter plate productions (Emond et al., 2007). A 5.9-fold increase of soluble enzyme production was thus obtained compared to the preliminary

assays without optimization (810 U.L.<sup>-1</sup> of culture, Figure 50 b). Compared to the previous values reported in (Brison et al., 2009), it represents a 63-fold improvement.



**Figure 51. Surface plots (3D) of the activity recovered in the soluble fraction for A53 productions, as function of: (a) repressor and culture duration; (b) repressor and glycerol; (c) temperature and culture duration; for B53 productions, as function of: (a) repressor and culture duration; (b) inducer and repressor; and for B55 productions, as function of: (a) glycerol and culture duration; (b) inducer and repressor; (c) temperature and culture duration**

With B53 system, the response surface shows that the soluble activity value increased with addition of both inducer and repressor (Figure 51, B53b) and also with the increase of repressor concentration and culture duration together (Figure 51, B53a). The highest value of soluble enzyme activity (3750 U.L.<sup>-1</sup> of culture) was predicted with 0.75% (w/v) inducer, 0.05% (w/v) repressor, 1.50% (w/v) glycerol, a temperature of 23 °C and a culture duration of 32 hours. The experimental validation allowed a production of 3060 U.L.<sup>-1</sup> of culture, which matches well with the predicted value (standard deviation of 18%). It represents a 2.5-fold increase compared to the preliminary experiment (1260 U.L.<sup>-1</sup> of culture, Figure 50 b) and a 40-fold increase compared to previous studies (Brison et al., 2009).

Finally, for B55 system, the highest value of soluble activity is predicted with a temperature of 23 °C and culture duration of 26.5 hours (Figure 51, B55c). Even if glycerol\*culture duration interaction term exhibited a significant p-value, no relevant impact of glycerol amount was observed on the 3-D response surface plot at the optimal time of 26.5

hours (Figure 51, B55a). Interestingly, inducer and repressor concentrations were linked (Figure 51, B55b), as observed for B53. However, the best expression was achieved when both inducer and repressor reached their minimal concentrations. An optimal production of 6420 U.L<sup>-1</sup><sub>of culture</sub> is predicted with inducer and glycerol concentrations of 1.00% (w/v) each, no addition of glucose (repressor), a temperature of 23 °C and a culture duration of 26.5 h. This prediction was experimentally validated with a measured enzymatic activity of 5740 U.L<sup>-1</sup><sub>of culture</sub>, corresponding to a standard deviation value of only 11%. This represents a 4.9-fold increase compared to the soluble production obtained before culture optimization (1170 U.L<sup>-1</sup><sub>of culture</sub>; Figure 50 b) and a 76-fold compared to previous studies (Brison et al., 2009).

### II.3. Large-scale productions

For each expression system (A53, B53 and B55), 1 L productions of  $\Delta N_{123}$ -GBD-CD2 were performed in a 5 L Erlenmeyer flasks, using the conditions determined from the model prediction (Table 13). For all host/vector combination, cultures in Erlenmeyer flasks always favored  $\Delta N_{123}$ -GBD-CD2 production, compared to deep-well format microplate ones.

**Table 13 Predicted and measured  $\Delta N_{123}$ -GBD-CD2 soluble activity level with optimal conditions from RSM analyses, for A53, B53 and B55 expression systems.**

Expression system	I (%)	Re (%)	G (%)	T (°C)	D (h)	Predicted activity <sup>a</sup> (U.L <sup>-1</sup> <sub>of culture</sub> )	Measured activity <sup>b</sup> (U.L <sup>-1</sup> <sub>of culture</sub> )
A53	1.40	0.00	1.25	23	28	5810	4800
B53	0.75	0.05	1.50	23	32	3750	3060
B55	1.00	0.00	1.00	23	26.5	6420	5740

Factors are coded as following: Inducer (I), repressor (Re), glycerol (G), temperature (T), culture duration (D)

<sup>a</sup> Activity predicted by RSM analysis for 1 mL productions in microplates.

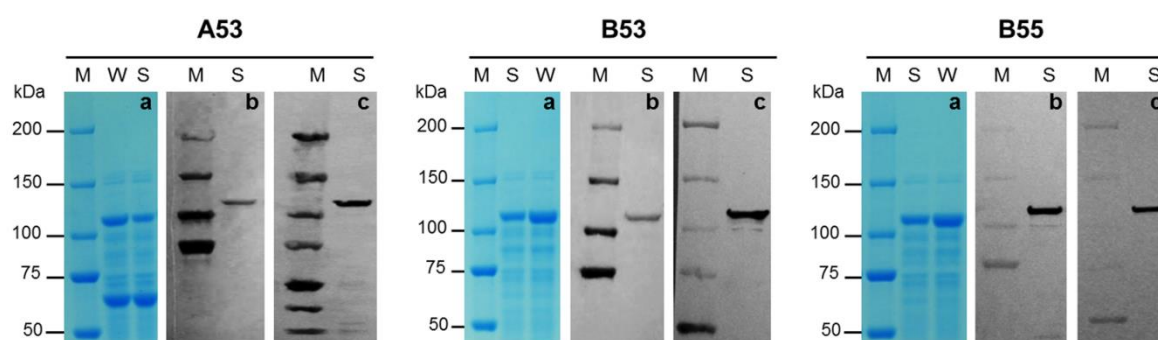
<sup>b</sup> 1 mL production in microplates for experimental validation. Measured activity in optimal condition predicted by RSM analysis

Productions with A53 and B53 expression systems reached 5590 and 9530 U.L<sup>-1</sup><sub>of culture</sub> respectively. A maximal soluble activity value of 12530 U.L<sup>-1</sup><sub>of culture</sub> was obtained with B55 expression system. This represents an impressive 165-fold increase compared to the values obtained with TOP10 *E. coli* cells and pBAD expression vector (Brison et al., 2009).



## II.4. Protein expression pattern

Identical amounts of whole cell lysates and soluble fractions obtained with A53, B53 and B55 systems were loaded in SDS-PAGE (Figure 52). A band estimated around 120 kDa, corresponding to  $\Delta N_{123}$ -GBD-CD2 molecular weight, was clearly observed in all lanes, confirming enzyme overexpression. In all cases, a larger amount of  $\Delta N_{123}$ -GBD-CD2 was observed in lanes corresponding to the whole cells extracts, indicating that there was still a slight production of insoluble enzyme. For A53, bands around 50 kDa corresponded to chaperons overexpressed in ArcticExpress DE3 host strain.



**Figure 52. SDS-PAGE and western blot analyses of  $\Delta N_{123}$ -GBD-CD2 enzyme expression in A53, B53 and B55.** (a) SDS-PAGE analyses with W: whole cells extract; S: soluble fraction; M: Precision Plus Protein All Blue Standards (Biorad). (b) Western blot analyses using 6xHis tag detection. (c) Western blot analyses using Strep tag II detection

Western blot analyzes were performed using 6xHis tag (Figure 52 b) and Strep tag II (Figure 52 c). For all preparations, entire enzymes bearing the two tags were the most preponderant proteins. For all A53 and B53 productions, slight bands of lower molecular weight were detected by anti- Strep tag II, indicating N-terminal degradations (Figure 52 c, A53, B53) but no trace of degradation was observed with the 6xHis tag detection, indicating that C-terminal extremity was preserved. Concerning B55 expression system, no degradation was detected at protein C-terminal end. Only traces of N-terminal truncated forms were revealed by anti 6xHis tag antibodies (Figure 52 b-c, B55). Globally, from Western blot results, degradations seemed to occur preferentially at  $\Delta N_{123}$ -GBD-CD2 N-terminal end, but remain a minor phenomenon.

## II.5. Affinity purification of recombinant protein

The first purification assays rapidly showed that highly overexpressed chaperons in ArcticExpress DE3 *E. coli* extracts troubled the automated purification process used in this study (data not shown). Purification assays were then focused only on protein extracts produced by *E. coli* BL21 Star DE3 (expression systems B53 and B55).

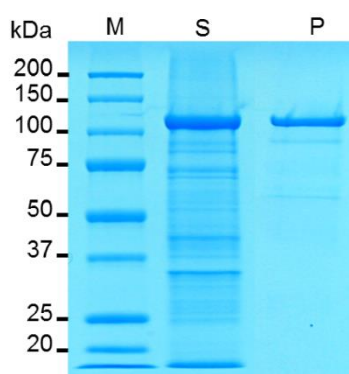


**Table 14 Purification of  $\Delta N_{123}$ -GBD-CD2 expressed in pET-53 and pET-55 constructions from BL21 Star DE3 overexpression scale-up experiments**

		Activity <sup>a</sup> (U)	Protein (mg)	Specific activity (U.mg <sup>-1</sup> )	Purification factor	Yield (%)
<b>B53</b>	Soluble fraction	227	74.8	3.0	1	100
	Fixation outflow	11	60.9	0.2	0.1	5
	Wash outflow	22	3.9	5.6	1.9	10
	Purified fraction	108	3.9	27.7	9.1	48
<b>B55</b>	Soluble fraction	343	74.2	4.6	1	100
	Fixation outflow	21	70.1	0.3	0.1	6
	Wash outflow	6	2.0	3.0	0.6	2
	Purified fraction	39	1.3	30.0	6.5	11

<sup>a</sup> One unit of  $\Delta N_{123}$ -GBD-CD2 enzyme corresponds to the amount of enzyme that catalyzes the production of 1  $\mu$ mol of fructose per min under the assay conditions.

Affinity purification of B53 extracts was performed thanks to the N-terminal 6xHis fused tag, as no degradation occurred at the C-terminal extremity of the protein. Conversely, purification of B55 extracts was carried out through Strep tag II for the same reasons. The specific activities, purification factors and yields are reported in Table 14.

**Figure 53. SDS-PAGE of  $\Delta N_{123}$ -GBD-CD2 expressed in B53 expression system in soluble extract (S) and purified by 6xHis tag affinity chromatography**

The best purification factor of nine with a yield of 48% was achieved with  $\Delta N_{123}$ -GBD-CD2 expressed in B53 expression system. SDS-PAGE gel (Figure 53) displayed very little contaminations by native *E. coli* proteins or  $\Delta N_{123}$ -GBD-CD2 degraded forms, resulting in satisfying enzyme preparation.

### III. Discussion

When confronted to a difficult-to-produce protein, one of the first recommended reflexes is to test several bacterial host strains to maximize chances to produce recombinant protein in soluble fraction. However, as there is no “universal rule” to avoid the formation of inclusion bodies, the option of optimizing recombinant protein production by mean of factorial design is also attractive. In this study, we combined those approaches to enhance the production of the  $\alpha$ -(1 $\rightarrow$ 2) branching enzyme named  $\Delta N_{123}$ -GBD-CD2. We showed that the choice of the host is effectively quite relevant. By varying *E. coli* strain from TOP10 to BL21 Star DE3 enabled a huge increase in activity recovered in soluble fraction (from 0 to 1260 U.L<sup>-1</sup><sub>of culture</sub>, for the construction with pET-53 vector, Figure 50 b).

The expression vector is also a key parameter. For this purpose, the Gateway cloning system is advantageous as it permits to rapidly and easily insert the gene of interest into several destination vectors, under the control of different promoters and in fusion with different types of sequence encoding tags. Fusion-tag and their location at N or C-terminal extremity of the target protein are generally described as affecting soluble enzyme production. However, this is highly protein-dependent and remains difficult to predict, as experienced by other groups (Dümmler et al., 2005; Noguère et al., 2012; Peti and Page, 2007). Fusion to well-known solubility enhancers such as thioredoxin or SUMO tags (Peroutka III et al., 2011; Waugh, 2005) was not retained due to their relative large size (both about 12 kDa). Two destination pET-vectors allowing fusion with small tags at N and C-term were preferred. The position of the tags at N or C-terms influenced the level of the soluble enzyme production, as seen in the differences of activity recovered when using pET-53 or pET-55 expression vectors. On a general basis, 6xHis tag fused to the N-terminal end of  $\Delta N_{123}$ -GBD-CD2 sequence favoured protein expression (Figure 50). The presence of two tags was also found very useful to evaluate the amount of non-degraded proteins, which is important for subsequent purification. The rapid nickel affinity chromatography set-up in this study resulted in satisfactory protein yield and purification factor, for such a high molecular weight protein (123 kDa). Regarding to A53 expression system, the amount of produced chaperones was higher than that of the protein of interest. Bolanos-Garcia and Davies have already shown that subsequent purification quality can decline due to their predominance in cell extracts (Bolanos-Garcia and Davies, 2006). Thus, we recommend avoiding this kind of host if further affinity purification is needed.

The factorial design developed here in microtiter plates was fast and efficient to optimize the soluble enzyme production. Among the five studied factors (temperature, culture duration, inducer ( $\alpha$ -lactose), repressor (glucose) and glycerol concentration), the

temperature and culture duration were the two most important ones, as they showed the strongest effects on soluble enzyme yields (linear, quadratic and/or cross term interactions, table 12). Decreasing culture temperature slows down the cell machinery, probably diminishing protein misfolding and inclusion bodies formation, as already shown in previous works (Feller et al., 1998; Song et al., 2012). However, we showed that drastic decrease of the culture temperature (down to 18 °C) resulted in huge decrease of soluble enzyme production. For all considered expression systems, 23 °C was the temperature of choice for a maximal production. Most of the time, this factor is tightly linked to culture duration, as reported for A53 and B55 expression systems. This highlights the importance of using experimental designs, such as response surface methodologies, that consider cross term interactions to establish the best compromise. A less intense effect of glycerol,  $\alpha$ -lactose and glucose concentrations was observed throughout the tested factor space. Those factors were more statistically significant during the first experimental design performed with lower concentrations (data not shown). In this second better-defined Box-Behnken design, the effect of those factors reached a plateau, indicating an overall tolerance to minor variations. However, it was advantageous to not add glucose (repressor) in the ZYM auto-inducible medium (optimal values reached with 0% added glucose for A53 and B55, and with 0.05% for B53). Indeed, glucose, the preferred source of carbon for *E. coli*, favors acetic acid synthesis (as an end product of glycolysis) that may interact with bacteria metabolism to the detriment of recombinant protein production (Ko et al., 1995). This effect was previously observed for three recombinant proteins expressed in *E. coli* B834, where better productions were achieved in small scale cultures when glucose concentration had been decreased from 0.050% to 0.015% (Blommel et al., 2008). Compared to classical ZYM5052 medium composition, glycerol concentration had to be at least doubled to maximize protein soluble production for the three studied expression systems (from 0.5% to 1.25%, 1.5% and 1% for A53, B53 and B55 respectively). Whereas it is commonly used in growth media, glycerol effects on heterologous protein overexpression still remain unclear. Besides its effects on metabolism when used as carbon source (Lin, 1976), glycerol may also play a role as a protective agent limiting protein aggregation. Interestingly, when using *E. coli* BL21 Star DE3 as host strain, significant linear, quadratic or term interaction effects were different for the two expression systems. This behavior dissimilarity further underlines the strong influence of the expression vector on the soluble enzyme production. This is particularly true if they allow fusions to affinity or solubility tags that could influence the resulting fold of the target protein (Diaz Ricci and Hernández, 2000).

As it was previously described, notably by (Kimata et al., 2012), levels of protein expression are often reduced in small scale cultivation format compared to production in

Erlenmeyer flasks, probably due to aeration and agitation parameters that are more difficult to control. However, this study shows that quite large amounts of recombinant protein can be produced when key-parameters are tightly controlled. To this end, RSM approach is quite efficient. Moreover, scaling-up the culture conditions allowed to produce higher amounts of soluble enzymes that were never reported before for any member of the GH70 family, which are all large enzymes having a tendency to form inclusion bodies in *E. coli*.  $\Delta N_{123}$ -GBD-CD2 is dedicated to the synthesis of gluco-oligosaccharides with controlled amount of  $\alpha$ -(1→2) linkages. This specificity is really attractive, especially for the trendy market of prebiotic compounds. The optimized production of 12 530 U.L<sup>-1</sup><sub>of culture</sub> achieved in this study (165-fold increase compared to first trials) may substantially facilitate production process, especially by avoiding denaturation and refolding step previously used (Brison et al., 2012). Moreover, the excellent expression in deep-well format microplates of  $\Delta N_{123}$ -GBD-CD2 glucansucrase (5740 U.L<sup>-1</sup><sub>of culture</sub> with B55 expression system) allows development of new enzyme engineering strategies to expand  $\Delta N_{123}$ -GBD-CD2 applications.

Finally, beyond the specific field of glucansucrases, we hope that this general strategy, which has been designed to overcome common expression problems, can be applied to many other proteins.

## IV. Materials and methods

### IV.1. Bacterial strains and expression vectors

Four *E. coli* strains, based on T7 expression systems, were selected: ArcticExpress DE3 (Agilent Technologies), BL21 Star DE3 (Life Technologies), C41 DE3 (Lucigen) and Rosetta DE3 (Novagen). ArcticExpress DE3 strain contains chaperonins Cpn60 and Cpn10 from psychrophilic bacterium *Oleispira Antarctica*. BL21 Star DE3 is a derivative of the *lon* and *ompT* deficient BL21 strain that carries a mutated *rne* gene encoding a truncated RNase E enzyme, to improve mRNA stability. C41 DE3 cells are commercialized for their efficiency in expression of toxic proteins, whereas Rosetta host strain is a BL21 derivative which contains a plasmid encoding for additional tRNAs. *E. coli* TOP10 cells were used for cloning experiments, (Life technologies). Genotypes of these different strains are presented in Online Resource 1, table S4.

Two expression vectors were used: the Gateway<sup>®</sup> Nova pET-53-DEST<sup>™</sup> DNA and Gateway<sup>®</sup> Nova pET-55-DEST<sup>™</sup> DNA (Novagen). The pET-53 destination vector allows cloning of the desired gene in fusion with an N-terminal 6xHis tag and a C-terminal Strep tag

II coding sequence. Conversely, the pET-55 destination vector allows gene cloning between an N-terminal Strep tag II and a C-terminal 6xHis tag encoding sequence.

## IV.2. Molecular cloning / Construction of expression clones

$\delta n_{123}$ -*gbd-cd2* gene (accession number: AJ430204) was amplified by PCR from pBAD-TOPO-TA- $\Delta N_{123}$ -*gbd-cd2* plasmid template (Brison et al. 2012) using the two following primers: Forward primer: 5'- CACCCAAGCAGGTCACTATATC -3'; Reverse primer: 5'- AGGTAATGTTGATTTATCACCATCAAGC -3'. The addition of CACC sequence (underlined) to the 5'-forward primer allowed correct insertion of the gene into the pENTR/D-TOPO<sup>®</sup> vector (Life Technologies). From a positive entry clone, LR recombination (Gateway<sup>®</sup> LR Clonase<sup>®</sup> II enzyme mix, Life technologies) was performed with pET-53-DEST and pET-55-DEST (Novagen) destination vectors. Expression clones were selected on LB agar plates supplemented with 100  $\mu\text{g.mL}^{-1}$  of ampicillin. Plasmids were extracted with Sigma-Aldrich GenElute HP Plasmid Miniprep kit, verified by restriction analyses and sequencing (GATC Biotech).

## IV.3. Growth media

For cloning experiments and starter cultures, cells were cultivated in Luria-Bertani (LB) medium. All starter culture media contained 100  $\mu\text{g.mL}^{-1}$  of ampicillin, supplemented with 12.5  $\mu\text{g.mL}^{-1}$  of chloramphenicol for expression in Rosetta DE3 host strains, or 20  $\mu\text{g.mL}^{-1}$  of gentamycin for expression in ArcticExpress DE3 cells. The auto-induction medium ZYM5052 used for the recombinant protein expression was prepared as described in Studier's work (Studier, 2005) and was supplemented with 100  $\mu\text{g.mL}^{-1}$  of ampicillin. Concentrations of glycerol, glucose (repressor) and  $\alpha$ -lactose (inducer) varied depending on the experimental design conditions.

## IV.4. Response surface design and statistical analyses

For each of the three strain/construction couples, a Box-Behnken design (Box and Behnken, 1960) with 5 factors was used. Table 15 details factor's levels. The response surface designs were generated for each strain/construction couple using Minitab software (Release 16) and consisted in 3 blocks of 45 experiments. They are listed in Online Resource 1, Table S5.

**Table 15. Experimental level of factors affecting the enzyme  $\Delta N_{123}$ -GBD-CD2 production for each host / vector combination. A53 for ArcticExpress DE3 *E. coli* cells transformed with pET-53. B53 for BL21 Star DE3 transformed with pET-53 expression plasmid. B55 for BL21 Star *E. coli* DE3 cells transformed with pET-55.**

<i>E. coli</i> strain / construction	A53		B53		B55	
Factor	Low level	High level	Low level	High level	Low level	High level
Lactose (I) (w/v)	0.75%	1.75%	0.25%	0.75%	1.00%	2.00%
Glucose (Re) (w/v)	0.00%	0.05%	0.00%	0.05%	0.00%	0.05%
Glycerol (G) (w/v)	0.50%	1.50%	0.50%	1.50%	0.50%	1.50%
Temperature (T)	18°C	28°C	18°C	28°C	18°C	28°C
Culture duration (D)	20h	32h	20h	32h	20h	32h

The analyses were performed using the response surface analysis module of Minitab software, by fitting the following second order polynomial equation:

$$\hat{Y} = \beta_0 + \sum_{i=1}^I \beta_i X_i + \sum_{i=1}^I \beta_{ii} X_i^2 + \sum_i \sum_j \beta_{ij} X_i X_j$$

where  $\hat{Y}$  is the predicted response,  $I$  is the number of factors (5 in this study),  $\beta_0$  is the model constant,  $\beta_i$  is the linear coefficient associated to factor  $X_i$ ,  $\beta_{ii}$  is the quadratic coefficient associated to factor  $X_i$  and  $\beta_{ij}$  is the interaction coefficient between factors  $X_i$  and  $X_j$ . To achieve clarity, factors,  $X_i$  will be coded as following: Inducer (I), Repressor (Re), Glycerol (G), Temperature (T), culture Duration (D).

The analysis of the variance (ANOVA) table, the estimates of each coefficient as well as their significance level were determined. The models were simplified by keeping terms showing p-values lower than 0.05 (except for first order terms which were kept as long as an interaction or a quadratic term including this factor was statistically significant). The model also allowed determining each factor value leading to the optimal production of soluble  $\Delta N_{123}$ -GBD-CD2 enzyme (expressed in  $U \cdot L^{-1}_{\text{of culture}}$ ). The response surface plots of the predicted model were obtained by depicting two significant factors within experimental range, and keeping the other factors at their central level.

## IV.5. Productions of recombinant $\Delta N_{123}$ -GBD-CD2

### IV.5.1. Starter cultures

All Starter cultures were performed in the same conditions. 50 mL of LB medium supplemented with ampicillin ( $100 \mu\text{g} \cdot \text{mL}^{-1}$ ) were inoculated with 300  $\mu\text{L}$  of *E. coli* BL21 Star

DE3 or ArcticExpress DE3 (supplemented with 100  $\mu\text{g.mL}^{-1}$  gentamycin) cells freshly transformed, and incubated overnight at 37 °C under agitation (200 rpm).

### IV.5.2. Culture in microplates

The experimental design experiments were performed in 96-Deep Well Format plates using 1 mL final culture volume. Before sterilization, plate wells were filled with 375  $\mu\text{L}$  aqueous solutions of glucose,  $\alpha$ -lactose and glycerol, according to Box-Benken design and using a pipetting robot (TECAN, Maennedorf Switzerland). Then, 525  $\mu\text{L}$  of sterile ZY2X (20  $\text{g.L}^{-1}$  tryptone, 10  $\text{g.L}^{-1}$  yeast extract) containing 200  $\mu\text{g.mL}^{-1}$  of ampicillin were added, using a pipetting robot Biomek2000 (Beckman, Fullerton, CA) under sterile conditions. The starter cultures were diluted to reach an OD (600 nm) of 0.5 and 100  $\mu\text{L}$  of the dilutions were used to inoculate the culture medium with the Beckman Coulter Biomek2000 automaton. The plates were incubated at desired temperatures under horizontal agitation (700 rpm) and centrifuged (20 min, 2200 g, 4 °C) after different periods of incubation. The cell pellets were then washed with 500  $\mu\text{L}$  of physiological saline water, resuspended in 300  $\mu\text{L}$  of buffered lysozyme (50 mM sodium acetate buffer pH 5.75, lysozyme 0.5  $\text{mg.mL}^{-1}$ ), followed by freezing O/N at -80 °C. After thawing at 4 °C, the plates were centrifuged (20 min, 2200 g, 4 °C) and the supernatants, containing recombinant  $\Delta\text{N}_{123}$ -GBD-CD2, were transferred into novel microplates.

### IV.5.3. Cultures in 5L-Erlenmeyer flask

One liter of optimized auto-inducible ZYM5052 medium prepared according to the conditions defined from RSM and containing 100  $\mu\text{g.mL}^{-1}$  ampicillin was inoculated at an OD (600 nm) of 0.05 with the starter culture, and incubated at temperature preconized from experimental design analyses. Cells were harvested by centrifugation, resuspended in 50 mM sodium acetate buffer (pH 5.75) at a final OD (600 nm) of 30, and disrupted by sonication. The amount of  $\Delta\text{N}_{123}$ -GBD-CD2 enzymes (123 kDa) recovered in the soluble fraction after centrifugation of the crude cell extract (15000 g, 30 min) was determined and expressed in units per liter of culture ( $\text{U.L}^{-1}$  of culture).

## IV.6. Enzymatic assays

All assays were performed at 30 °C in 50 mM sodium acetate, pH 5.75. Enzyme activity was determined by measuring the initial velocity of released fructose from 292 mM sucrose. Fructose concentration was determined using the dinitrosalicylic acid (DNS) method



(Miller, 1959). One unit of  $\Delta N_{123}$ -GBD-CD2 enzyme corresponds to the amount of enzyme that catalyses the production of 1  $\mu$ mol of fructose per min under the assay conditions.

## IV.7. Protein expression analysis

### IV.7.1. Electrophoresis gel

NuPAGE<sup>®</sup> Tris-Acetate gels 3-8%, 1.5 mm (Life technologies) were used to analyse protein expression and solubility. The enzymatic samples (soluble fraction and crude extract) were 10-time diluted and 15  $\mu$ l were loaded onto a precast gel. Protein marker Precision Plus, Protein All Blue (Bio-Rad) was used as standard. The separation was performed in a NuPAGE<sup>®</sup> Tris-Acetate SDS Running buffer (Life Technology) for one hour with a 150 V constant voltage. Proteins were stained with a PAGE BLUE protein staining solution (Thermo Scientific).

### IV.7.2. Western blots

After NuPAGE<sup>®</sup> gel migration, proteins were transferred onto a nitrocellulose membrane (0.45  $\mu$ m Protran<sup>®</sup> BA85 from GE Healthcare, Life Sciences – Whatman<sup>™</sup>) with Bio-Rad Mini Trans-Blot Cell at 100 V during 1 hour using a transfer buffer (3 g.L<sup>-1</sup> Tris, 14.4 g.L<sup>-1</sup> glycine, 20% (v/v) ethanol).

The 6xHis tag detection was carried out as follow. The nitrocellulose membrane was blocked overnight at 4 °C under gentle agitation with PBSMT solution: PBS 1X (NaCl 1.4 M, Na<sub>2</sub>HPO<sub>4</sub> 100 mM, KH<sub>2</sub>HPO<sub>4</sub> 18 mM, KCl 27 mM, pH7.3), 5% (w/v) skim milk, 0.5% (v/v) tween 20. Primary monoclonal antibodies (Mouse Anti-6xHis antibodies, Life technologies) were 1:5000-diluted in PBSMT. After incubation with primary antibodies for 1 hour at room temperature, the blotting membrane was washed 3 $\times$ 15 min with PBS 1X, 1% (w/v) skim milk, 0.5% (v/v) tween 20. The membrane was then incubated in the same solution with 1:5000 diluted Alkaline Phosphatase-conjugated secondary antibodies (Anti-Mouse IgG Alkaline Phosphatase from goat, Sigma-Aldrich), during one hour at room temperature. Then, the membrane was washed 3 $\times$ 15 min with PBS 1X, 0.5% (v/v) tween 20. The tag was detected with the AP chromogenic substrate BCIP/NBT (BCIP<sup>®</sup>/NBT Liquid Substrate System, Sigma-Aldrich).

Regarding Strep tag II detection, the nitrocellulose membrane was blocked overnight at 4 °C under gentle agitation with PBS blocking buffer (PBS 1X, 3% (w/v) BSA, 0.5% (v/v) tween 20). After 3 $\times$ 5 min washes with PBST (PBS 1X, 0.1% tween 20), the membrane was incubated with AP conjugated antibodies for one hour, at room temperature. Strep Tactin



conjugated to alkaline phosphatase antibodies (IBA technology) were diluted 1:4000 in PBST solution. The membrane was then washed 2x1 min with PBST solution, followed by a final wash in PBS. Finally, the membrane was incubated in reaction buffer (100 mM NaCl, 5 mM MgCl<sub>2</sub>, 100 mM Tris-HCl, pH 8.8, 10% BCIP/NBT) for band detection.

## **IV.8. Affinity purification of recombinant protein**

All purification assays were performed with ÄKTAexpress protein-purification system at 8 °C (GE Healthcare).

### **IV.8.1. Affinity 6xHis tag chromatography.**

The soluble fraction of cell lysate was adjusted to 500 mM NaCl and pH 7.4 and filtered using a 0.22  $\mu$ m cartridge. The cleared lysate was applied onto a 1 mL HisTrap HP<sup>®</sup> (GE Healthcare) column formerly equilibrated with the binding buffer (20 mM sodium phosphate buffer pH 7.4, 500 mM NaCl). Proteins were eluted by applying a 25-minute long gradient of 10 to 500 mM imidazole in binding buffer. 3-mL fractions of the eluate were desalted by loading onto 10-DG column (Biorad), pre-equilibrated with 50 mM sodium acetate buffer pH 5.75.

### **IV.8.2. Affinity Strep tag II chromatography**

The soluble fraction of cell lysate was adjusted to 280 mM NaCl and pH 7.3. After filtration on a 0.22  $\mu$ m cartridge, the cleared lysate was applied onto a 5 mL StrepTrap HP<sup>®</sup> (GE Healthcare) column formerly equilibrated with the binding buffer (50 mM PBS 1X pH 7.3, 280 mM NaCl). Proteins were eluted by applying a 100-minute long gradient of 0 to 2.5 mM D-desthiobiotin in binding buffer. The collected fractions were desalted onto 10-DG column (Biorad) pre-equilibrated with 50 mM sodium acetate buffer pH 5.75.

## **V. Acknowledgments**

We thankfully acknowledge the assistance of S. Bozonnet and S. Pizzut-Serin in using the ICEO facility. This work was supported by the French National Research Agency (ANR-10-ALIA-0003 Oenopolys 2011-2013).

## VI. References

- Baneyx, F. (1999). Recombinant protein expression in *Escherichia coli*. *Curr. Opin. Biotechnol.* 10, 411–421.
- Baş, D., and Boyacı, İ.H. (2007). Modeling and optimization I: Usability of response surface methodology. *J. Food Eng.* 78, 836–845.
- Blommel, P.G., Becker, K.J., Duvnjak, P., and Fox, B.G. (2008). Enhanced Bacterial Protein Expression During Auto-Induction Obtained by Alteration of Lac Repressor Dosage and Medium Composition. *Biotechnol. Prog.* 23, 585–598.
- Bolanos-Garcia, V.M., and Davies, O.R. (2006). Structural analysis and classification of native proteins from *E. coli* commonly co-purified by immobilised metal affinity chromatography. *Biochim. Biophys. Acta BBA - Gen. Subj.* 1760, 1304–1313.
- Box, G.E.P., and Behnken, D.W. (1960). Simplex-Sum Designs: A Class of Second Order Rotatable Designs Derivable From Those of First Order. *Ann. Math. Stat.* 31, 838–864.
- Brison, Y., Fabre, E., Moulis, C., Portais, J.-C., Monsan, P., and Remaud-Siméon, M. (2009). Synthesis of dextrans with controlled amounts of  $\alpha$ -1,2 linkages using the transglucosidase GBD-CD2. *Appl. Microbiol. Biotechnol.* 86, 545–554.
- Brison, Y., Pijning, T., Malbert, Y., Fabre, E., Mourey, L., Morel, S., Potocki-Veronese, G., Monsan, P., Tranier, S., Remaud-Simeon, M., et al. (2012). Functional and Structural Characterization of  $\alpha$ -1,2 Branching Sucrase Derived from DSR-E Glucansucrase. *J. Biol. Chem.* 287, 7915–7924.
- Diaz Ricci, J.C., and Hernández, M.E. (2000). Plasmid Effects on *Escherichia coli* Metabolism. *Crit. Rev. Biotechnol.* 20, 79–108.
- Dümmler, A., Lawrence, A.-M., and de Marco, A. (2005). Simplified screening for the detection of soluble fusion constructs expressed in *E. coli* using a modular set of vectors. *Microb. Cell Factories* 4, 34.
- Emond, S., Potocki-Veronese, G., Mondon, P., Bouayadi, K., Kharrat, H., Monsan, P., and Remaud-Simeon, M. (2007). Optimized and Automated Protocols for High-Throughput Screening of Amylosucrase Libraries. *J. Biomol. Screen.* 12, 715–723.
- Fabre, E., Bozonnet, S., Arcache, A., Willemot, R.-M., Vignon, M., Monsan, P., and Remaud-Simeon, M. (2004). Role of the Two Catalytic Domains of DSR-E Dextranucrase and Their Involvement in the Formation of Highly  $\alpha$ -1,2 Branched Dextran. *J. Bacteriol.* 187, 296–303.
- Feller, G., Bussy, O.L., and Gerday, C. (1998). Expression of Psychrophilic Genes in Mesophilic Hosts: Assessment of the Folding State of a Recombinant  $\alpha$ -Amylase. *Appl. Environ. Microbiol.* 64, 1163–1165.
- Gordon, E., Horsefield, R., Swarts, H.G.P., de Pont, J.J.H.M., Neutze, R., and Snijder, A. (2008). Effective high-throughput overproduction of membrane proteins in *Escherichia coli*. *Protein Expr. Purif.* 62, 1–8.
- Jana, S., and Deb, J.K. (2005). Strategies for efficient production of heterologous proteins in *Escherichia coli*. *Appl. Microbiol. Biotechnol.* 67, 289–298.
- Kimata, K., Yamaguchi, M., Saito, Y., Hata, H., Miyake, K., Yamane, T., Nakagawa, Y., Yano, A., Ito, K., and Kawarasaki, Y. (2012). High cell-density expression system: A novel method for extracellular production of difficult-to-express proteins. *J. Biosci. Bioeng.* 113, 154–159.
- Ko, Y.-F., Bentley, W.E., and Weigand, W.A. (1995). The effect of cellular energetics on foreign protein production. *Appl. Biochem. Biotechnol.* 50, 145–159.
- Koehn, J., and Hunt, I. (2009). High-Throughput Protein Production (HTPP): A Review of Enabling Technologies to Expedite Protein Production. In *High Throughput Protein Expression and Purification*, S.A. Doyle, ed. (Totowa, NJ: Humana Press), pp. 1–18.
- Lin, E.C.C. (1976). Glycerol Dissimilation and its Regulation in Bacteria. *Annu. Rev. Microbiol.* 30, 535–578.

- Lombard, V., Golaconda Ramulu, H., Drula, E., Coutinho, P.M., and Henrissat, B. (2013). The carbohydrate-active enzymes database (CAZy) in 2013. *Nucleic Acids Res.*
- Miller, G.L. (1959). Use of Dinitrosalicylic Acid Reagent for Determination of Reducing Sugar. *Anal. Chem.* 31, 426–428.
- Nawani, N.N., and Kapadnis, B.P. (2005). Optimization of chitinase production using statistics based experimental designs. *Process Biochem.* 40, 651–660.
- Noguère, C., Larsson, A.M., Guyot, J.-C., and Bignon, C. (2012). Fractional factorial approach combining 4 *Escherichia coli* strains, 3 culture media, 3 expression temperatures and 5 N-terminal fusion tags for screening the soluble expression of recombinant proteins. *Protein Expr. Purif.* 84, 204–213.
- Papaneophytou, C.P., and Kontopidis, G.A. (2012). Optimization of TNF- $\alpha$  overexpression in *Escherichia coli* using response surface methodology: Purification of the protein and oligomerization studies. *Protein Expr. Purif.* 86, 35–44.
- Peroutka III, R.J., Orcutt, S.J., Strickler, J.E., and Butt, T.R. (2011). SUMO Fusion Technology for Enhanced Protein Expression and Purification in Prokaryotes and Eukaryotes. In *Heterologous Gene Expression in E.coli*, T.C. Evans, and M.-Q. Xu, eds. (Totowa, NJ: Humana Press), pp. 15–30.
- Peti, W., and Page, R. (2007). Strategies to maximize heterologous protein expression in *Escherichia coli* with minimal cost. *Protein Expr. Purif.* 51, 1–10.
- Sarbini, S.R., Kolida, S., Naeye, T., Einerhand, A., Brison, Y., Remaud-Simeon, M., Monsan, P., Gibson, G.R., and Rastall, R.A. (2011). In Vitro Fermentation of Linear and -1,2-Branched Dextrins by the Human Fecal Microbiota. *Appl. Environ. Microbiol.* 77, 5307–5315.
- Sarbini, S.R., Kolida, S., Naeye, T., Einerhand, A.W., Gibson, G.R., and Rastall, R.A. (2013). The prebiotic effect of  $\alpha$ -1,2 branched, low molecular weight dextran in the batch and continuous faecal fermentation system. *J. Funct. Foods* 5, 1938–1946.
- Serino, M., Luche, E., Gres, S., Baylac, A., Berge, M., Cenac, C., Waget, A., Klopp, P., Iacovoni, J., Klopp, C., et al. (2011). Metabolic adaptation to a high-fat diet is associated with a change in the gut microbiota. *Gut* 61, 543–553.
- Song, J.M., An, Y.J., Kang, M.H., Lee, Y.-H., and Cha, S.-S. (2012). Cultivation at 6–10°C is an effective strategy to overcome the insolubility of recombinant proteins in *Escherichia coli*. *Protein Expr. Purif.* 82, 297–301.
- Studier, F.W. (2005). Protein production by auto-induction in high-density shaking cultures. *Protein Expr. Purif.* 41, 207–234.
- Studier, F.W. (2009). High density growth of T7 expression strains with auto-induction option. U.S. Patent Application 7,560,264.
- Swalley, S.E., Fulghum, J.R., and Chambers, S.P. (2006). Screening factors effecting a response in soluble protein expression: Formalized approach using design of experiments. *Anal. Biochem.* 351, 122–127.
- Terpe, K. (2003). Overview of tag protein fusions: from molecular and biochemical fundamentals to commercial systems. *Appl. Microbiol. Biotechnol.* 60, 523–533.
- Tyler, R.C., Sreenath, H.K., Singh, S., Aceti, D.J., Bingman, C.A., Markley, J.L., and Fox, B.G. (2005). Auto-induction medium for the production of [U-15N]- and [U-13C, U-15N]-labeled proteins for NMR screening and structure determination. *Protein Expr. Purif.* 40, 268–278.
- Upadhyay, A.K., Murmu, A., Singh, A., and Panda, A.K. (2012). Kinetics of Inclusion Body Formation and Its Correlation with the Characteristics of Protein Aggregates in *Escherichia coli*. *PLoS ONE* 7, e33951.
- Villaverde, A., and Carrió, M.M. (2003). Protein aggregation in recombinant bacteria: biological role of inclusion bodies. *Biotechnol. Lett.* 25, 1385–1395.
- Waugh, D.S. (2005). Making the most of affinity tags. *Trends Biotechnol.* 23, 316–320.

## VII. Supplementary material

**Table S4. *E. coli* strain genotypes used in this study for recombinant protein expression and for cloning experiments.**

<i>E. coli</i> expression strains	Genotype
Arctic Express DE3	<i>E. coli</i> B, F <sup>-</sup> , ompT, hsdS, (rB <sup>-</sup> mB <sup>-</sup> ), dcm <sup>+</sup> , TetR, gal $\lambda$ (DE3) endA Hte [cpn10 cpn60 GentR].
BL21 star DE3	F <sup>-</sup> , ompT, hsdSB, (rB-mB <sup>-</sup> ), gal, dcm, rne131 (DE3).
C41 DE3	F <sup>-</sup> , ompT, hsdSB, (rB <sup>-</sup> mB <sup>-</sup> ), gal, dcm (DE3).
JM109 DE3	endA1, recA1, gyrA96, thi, hsdR17 (rk <sup>-</sup> , mk <sup>+</sup> ), relA1, supE44, $\lambda$ <sup>-</sup> , $\Delta$ (lac-proAB), [F <sup>+</sup> , traD36, proAB, lacIqZ $\Delta$ M15], IDE3.
Rosetta DE3	F <sup>-</sup> , ompT, hsdSB, (rB <sup>-</sup> mB <sup>-</sup> ), gal, dcm, (DE3), pRARE, (CamR).
<i>E. coli</i> strain for cloning experiments	Genotype
TOP10	F <sup>-</sup> , mcrA, $\Delta$ (mrr-hsdRMS-mcrBC), $\phi$ 80lacZ $\Delta$ M15, $\Delta$ lacX74, recA1, araD139, $\Delta$ (araleu)7697, galU, galK, rpsL, (StrR), endA1, nupG

**Table S5. Experimental design (Box-bekhen design) for each expression system (host / vector).****a. ArcticExpress DE3 *E. coli* strain / pET-53 expression vector (A53)**

Sample number (StdOrder)	I (%)	Re (%)	G (%)	T (°C)	D (h)	Soluble activity (U.L <sup>-1</sup> of culture)
1	0.75	0.000	1.0	23	26	7313*
2	1.75	0.000	1.0	23	26	5720
3	0.75	0.050	1.0	23	26	3780
4	1.75	0.050	1.0	23	26	4401
5	1.25	0.025	0.5	18	26	154
6	1.25	0.025	1.5	18	26	118
7	1.25	0.025	0.5	28	26	824*
8	1.25	0.025	1.5	28	26	1091
9	1.25	0.000	1.0	23	20	3359
10	1.25	0.050	1.0	23	20	2360
11	1.25	0.000	1.0	23	32	7268*
12	1.25	0.050	1.0	23	32	2668
13	0.75	0.025	0.5	23	26	4347
14	1.75	0.025	0.5	23	26	4111*
15	0.75	0.025	1.5	23	26	3680
16	1.75	0.025	1.5	23	26	4038
17	1.25	0.025	1.0	18	20	88*
18	1.25	0.025	1.0	28	20	731
19	1.25	0.025	1.0	18	32	2216*
20	1.25	0.025	1.0	28	32	1866*
21	1.25	0.000	0.5	23	26	4937
22	1.25	0.050	0.5	23	26	5289
23	1.25	0.000	1.5	23	26	5250
24	1.25	0.050	1.5	23	26	3781
25	0.75	0.025	1.0	18	26	227*
26	1.75	0.025	1.0	18	26	174*
27	0.75	0.025	1.0	28	26	896
28	1.75	0.025	1.0	28	26	938
29	1.25	0.025	0.5	23	20	3122
30	1.25	0.025	1.5	23	20	2286
31	1.25	0.025	0.5	23	32	4693
32	1.25	0.025	1.5	23	32	3105
33	0.75	0.025	1.0	23	20	2623
34	1.75	0.025	1.0	23	20	2354
35	0.75	0.025	1.0	23	32	4014*
36	1.75	0.025	1.0	23	32	4010
37	1.25	0.000	1.0	18	26	221*
38	1.25	0.050	1.0	18	26	175*
39	1.25	0.000	1.0	28	26	902*
40	1.25	0.050	1.0	28	26	859
41	1.25	0.025	1.0	23	26	5092
42	1.25	0.025	1.0	23	26	4547
43	1.25	0.025	1.0	23	26	2745*
44	1.25	0.025	1.0	23	26	4188
45	1.25	0.025	1.0	23	26	5011
46	1.25	0.025	1.0	23	26	5201
47	0.75	0.000	1.0	23	26	5822
48	1.75	0.000	1.0	23	26	6248
49	0.75	0.050	1.0	23	26	3219
50	1.75	0.050	1.0	23	26	4537
51	1.25	0.025	0.5	18	26	124
52	1.25	0.025	1.5	18	26	149*
53	1.25	0.025	0.5	28	26	1*
54	1.25	0.025	1.5	28	26	1046
55	1.25	0.000	1.0	23	20	1605*
56	1.25	0.050	1.0	23	20	2231
57	1.25	0.000	1.0	23	32	4861
58	1.25	0.050	1.0	23	32	2821
59	0.75	0.025	0.5	23	26	1768*
60	1.75	0.025	0.5	23	26	5442
61	0.75	0.025	1.5	23	26	3328
62	1.75	0.025	1.5	23	26	5443*
63	1.25	0.025	1.0	18	20	96*
64	1.25	0.025	1.0	28	20	498
65	1.25	0.025	1.0	18	32	1448
66	1.25	0.025	1.0	28	32	979*
67	1.25	0.000	0.5	23	26	4769
68	1.25	0.050	0.5	23	26	4292*
69	1.25	0.000	1.5	23	26	5119
70	1.25	0.050	1.5	23	26	3605
71	0.75	0.025	1.0	18	26	121
72	1.75	0.025	1.0	18	26	134
73	0.75	0.025	1.0	28	26	952

74	1.75	0.025	1.0	28	26	897
75	1.25	0.025	0.5	23	20	1841*
76	1.25	0.025	1.5	23	20	1332*
77	1.25	0.025	0.5	23	32	4149
78	1.25	0.025	1.5	23	32	2654
79	0.75	0.025	1.0	23	20	1684
80	1.75	0.025	1.0	23	20	2146
81	0.75	0.025	1.0	23	32	2696
82	1.75	0.025	1.0	23	32	3756
83	1.25	0.000	1.0	18	26	144*
84	1.25	0.050	1.0	18	26	132
85	1.25	0.000	1.0	28	26	968*
86	1.25	0.050	1.0	28	26	1182
87	1.25	0.025	1.0	23	26	5322
88	1.25	0.025	1.0	23	26	4681
89	1.25	0.025	1.0	23	26	4630
90	1.25	0.025	1.0	23	26	4141
91	1.25	0.025	1.0	23	26	4565
92	1.25	0.025	1.0	23	26	4737
93	0.75	0.000	1.0	23	26	5870
94	1.75	0.000	1.0	23	26	6751*
95	0.75	0.050	1.0	23	26	4222
96	1.75	0.050	1.0	23	26	4332
97	1.25	0.025	0.5	18	26	141
98	1.25	0.025	1.5	18	26	121
99	1.25	0.025	0.5	28	26	1586
100	1.25	0.025	1.5	28	26	1232
101	1.25	0.000	1.0	23	20	2463
102	1.25	0.050	1.0	23	20	2138
103	1.25	0.000	1.0	23	32	6155
104	1.25	0.050	1.0	23	32	2841
105	0.75	0.025	0.5	23	26	4998
106	1.75	0.025	0.5	23	26	6072*
107	0.75	0.025	1.5	23	26	4534
108	1.75	0.025	1.5	23	26	4648
109	1.25	0.025	1.0	18	20	103*
110	1.25	0.025	1.0	28	20	660
111	1.25	0.025	1.0	18	32	1623
112	1.25	0.025	1.0	28	32	1401*
113	1.25	0.000	0.5	23	26	4706
114	1.25	0.050	0.5	23	26	5787
115	1.25	0.000	1.5	23	26	5524
116	1.25	0.050	1.5	23	26	3705
117	0.75	0.025	1.0	18	26	129
118	1.75	0.025	1.0	18	26	132
119	0.75	0.025	1.0	28	26	1312*
120	1.75	0.025	1.0	28	26	1333*
121	1.25	0.025	0.5	23	20	2652
122	1.25	0.025	1.5	23	20	2002
123	1.25	0.025	0.5	23	32	6078*
124	1.25	0.025	1.5	23	32	3364
125	0.75	0.025	1.0	23	20	2079
126	1.75	0.025	1.0	23	20	1953
127	0.75	0.025	1.0	23	32	3317
128	1.75	0.025	1.0	23	32	3715
129	1.25	0.000	1.0	18	26	165*
130	1.25	0.050	1.0	18	26	125
131	1.25	0.000	1.0	28	26	1537
132	1.25	0.050	1.0	28	26	1104
133	1.25	0.025	1.0	23	26	5990
134	1.25	0.025	1.0	23	26	5363
135	1.25	0.025	1.0	23	26	4734
136	1.25	0.025	1.0	23	26	4817
137	1.25	0.025	1.0	23	26	5781
138	1.25	0.025	1.0	23	26	4693

Experimental conditions for inducer (I), repressor (Re) and glycerol (G) concentrations expressed as a w/v percentage, temperature (T) and culture duration (D).

\* Samples with a soluble activity considered as aberrant (compared to the corresponding triplicates) and not taken into account to generate the model.

b. BL21 star DE3 *E. coli* strain / pET-53 expression vector (B53)

Sample number (StdOrder)	I (%)	Re (%)	G (%)	T (°C)	D (h)	Soluble activity (U.L <sup>-1</sup> of culture)
1	0.25	0.000	1.0	23	26	3484
2	0.75	0.000	1.0	23	26	2818
3	0.25	0.050	1.0	23	26	3323
4	0.75	0.050	1.0	23	26	3193
5	0.50	0.025	0.5	18	26	295
6	0.50	0.025	1.5	18	26	149
7	0.50	0.025	0.5	28	26	199
8	0.50	0.025	1.5	28	26	58
9	0.50	0.000	1.0	23	20	4641*
10	0.50	0.050	1.0	23	20	2925
11	0.50	0.000	1.0	23	32	3099
12	0.50	0.050	1.0	23	32	2357*
13	0.25	0.025	0.5	23	26	2481
14	0.75	0.025	0.5	23	26	2308
15	0.25	0.025	1.5	23	26	3720
16	0.75	0.025	1.5	23	26	3051
17	0.50	0.025	1.0	18	20	107
18	0.50	0.025	1.0	28	20	560
19	0.50	0.025	1.0	18	32	2535*
20	0.50	0.025	1.0	28	32	221
21	0.50	0.000	0.5	23	26	2642
22	0.50	0.050	0.5	23	26	3286
23	0.50	0.000	1.5	23	26	3100
24	0.50	0.050	1.5	23	26	3461
25	0.25	0.025	1.0	18	26	284
26	0.75	0.025	1.0	18	26	307
27	0.25	0.025	1.0	28	26	64
28	0.75	0.025	1.0	28	26	46
29	0.50	0.025	0.5	23	20	3701
30	0.50	0.025	1.5	23	20	3151
31	0.50	0.025	0.5	23	32	1744*
32	0.50	0.025	1.5	23	32	2115*
33	0.25	0.025	1.0	23	20	3373
34	0.75	0.025	1.0	23	20	3358
35	0.25	0.025	1.0	23	32	2842
36	0.75	0.025	1.0	23	32	2239*
37	0.50	0.000	1.0	18	26	315
38	0.50	0.050	1.0	18	26	269
39	0.50	0.000	1.0	28	26	175
40	0.50	0.050	1.0	28	26	50
41	0.50	0.025	1.0	23	26	3253
42	0.50	0.025	1.0	23	26	3167
43	0.50	0.025	1.0	23	26	1989*
44	0.50	0.025	1.0	23	26	2866
45	0.50	0.025	1.0	23	26	3421
46	0.50	0.025	1.0	23	26	3727
47	0.25	0.000	1.0	23	26	3150
48	0.75	0.000	1.0	23	26	3070
49	0.25	0.050	1.0	23	26	2441
50	0.75	0.050	1.0	23	26	3741
51	0.50	0.025	0.5	18	26	257
52	0.50	0.025	1.5	18	26	190
53	0.50	0.025	0.5	28	26	298
54	0.50	0.025	1.5	28	26	53
55	0.50	0.000	1.0	23	20	3087
56	0.50	0.050	1.0	23	20	2448
57	0.50	0.000	1.0	23	32	3113
58	0.50	0.050	1.0	23	32	2948
59	0.25	0.025	0.5	23	26	570*
60	0.75	0.025	0.5	23	26	3883
61	0.25	0.025	1.5	23	26	2300
62	0.75	0.025	1.5	23	26	3679
63	0.50	0.025	1.0	18	20	115
64	0.50	0.025	1.0	28	20	350
65	0.50	0.025	1.0	18	32	2294*
66	0.50	0.025	1.0	28	32	175
67	0.50	0.000	0.5	23	26	2843
68	0.50	0.050	0.5	23	26	3304
69	0.50	0.000	1.5	23	26	2346
70	0.50	0.050	1.5	23	26	3057
71	0.25	0.025	1.0	18	26	181
72	0.75	0.025	1.0	18	26	212
73	0.25	0.025	1.0	28	26	60
74	0.75	0.025	1.0	28	26	67
75	0.50	0.025	0.5	23	20	2779
76	0.50	0.025	1.5	23	20	2380
77	0.50	0.025	0.5	23	32	2739
78	0.50	0.025	1.5	23	32	2852

79	0.25	0.025	1.0	23	20	2072
80	0.75	0.025	1.0	23	20	2871
81	0.25	0.025	1.0	23	32	3026
82	0.75	0.025	1.0	23	32	3503
83	0.50	0.000	1.0	18	26	236
84	0.50	0.050	1.0	18	26	176
85	0.50	0.000	1.0	28	26	195
86	0.50	0.050	1.0	28	26	75
87	0.50	0.025	1.0	23	26	2698
88	0.50	0.025	1.0	23	26	3140
89	0.50	0.025	1.0	23	26	2852
90	0.50	0.025	1.0	23	26	2950
91	0.50	0.025	1.0	23	26	3163
92	0.50	0.025	1.0	23	26	2805
93	0.25	0.000	1.0	23	26	3186
94	0.75	0.000	1.0	23	26	4291*
95	0.25	0.050	1.0	23	26	3218
96	0.75	0.050	1.0	23	26	3796
97	0.50	0.025	0.5	18	26	273
98	0.50	0.025	1.5	18	26	158
99	0.50	0.025	0.5	28	26	322
100	0.50	0.025	1.5	28	26	75
101	0.50	0.000	1.0	23	20	3767
102	0.50	0.050	1.0	23	20	2563
103	0.50	0.000	1.0	23	32	3527
104	0.50	0.050	1.0	23	32	3632
105	0.25	0.025	0.5	23	26	3167
106	0.75	0.025	0.5	23	26	3494
107	0.25	0.025	1.5	23	26	3774
108	0.75	0.025	1.5	23	26	4028*
109	0.50	0.025	1.0	18	20	124
110	0.50	0.025	1.0	28	20	352
111	0.50	0.025	1.0	18	32	2872*
112	0.50	0.025	1.0	28	32	47
113	0.50	0.000	0.5	23	26	3424
114	0.50	0.050	0.5	23	26	3176
115	0.50	0.000	1.5	23	26	2965
116	0.50	0.050	1.5	23	26	4386*
117	0.25	0.025	1.0	18	26	190
118	0.75	0.025	1.0	18	26	220
119	0.25	0.025	1.0	28	26	82
120	0.75	0.025	1.0	28	26	139
121	0.50	0.025	0.5	23	20	3626
122	0.50	0.025	1.5	23	20	2989
123	0.50	0.025	0.5	23	32	3768
124	0.50	0.025	1.5	23	32	3931
125	0.25	0.025	1.0	23	20	2995
126	0.75	0.025	1.0	23	20	2799
127	0.25	0.025	1.0	23	32	3633
128	0.75	0.025	1.0	23	32	3630
129	0.50	0.000	1.0	18	26	285
130	0.50	0.050	1.0	18	26	152
131	0.50	0.000	1.0	28	26	317
132	0.50	0.050	1.0	28	26	77
133	0.50	0.025	1.0	23	26	3481
134	0.50	0.025	1.0	23	26	3575
135	0.50	0.025	1.0	23	26	2854
136	0.50	0.025	1.0	23	26	2961
137	0.50	0.025	1.0	23	26	3273
138	0.50	0.025	1.0	23	26	3840

Experimental conditions for inducer (I), repressor (Re) and glycerol (G) concentrations expressed as a w/v percentage, temperature (T) and culture duration (D).

\* Samples with a soluble activity considered as aberrant (compared to the corresponding triplicates) and not taken into account to generate the model.



c. BL21 star DE3 *E. coli* strain / pET-55 expression vector (B55)

Sample number (StdOrder)	I (%)	Re (%)	G (%)	T (°C)	D (h)	Soluble activity (U.L <sup>-1</sup> of culture)
1	1.0	0.000	1.0	23	26	7755*
2	2.0	0.000	1.0	23	26	58967
3	1.0	0.050	1.0	23	26	4311
4	2.0	0.050	1.0	23	26	5442
5	1.5	0.025	0.5	18	26	384
6	1.5	0.025	1.5	18	26	198
7	1.5	0.025	0.5	28	26	156
8	1.5	0.025	1.5	28	26	225
9	1.5	0.000	1.0	23	20	5005
10	1.5	0.050	1.0	23	20	3937
11	1.5	0.000	1.0	23	32	5135
12	1.5	0.050	1.0	23	32	3008
13	1.0	0.025	0.5	23	26	4867
14	2.0	0.025	0.5	23	26	5339
15	1.0	0.025	1.5	23	26	4993
16	2.0	0.025	1.5	23	26	5323
17	1.5	0.025	1.0	18	20	112*
18	1.5	0.025	1.0	28	20	1129
19	1.5	0.025	1.0	18	32	2065
20	1.5	0.025	1.0	28	32	313*
21	1.5	0.000	0.5	23	26	70578
22	1.5	0.050	0.5	23	26	4958
23	1.5	0.000	1.5	23	26	6611
24	1.5	0.050	1.5	23	26	4380
25	1.0	0.025	1.0	18	26	574
26	2.0	0.025	1.0	18	26	517
27	1.0	0.025	1.0	28	26	245
28	2.0	0.025	1.0	28	26	396
29	1.5	0.025	0.5	23	20	4840
30	1.5	0.025	1.5	23	20	3625
31	1.5	0.025	0.5	23	32	3866
32	1.5	0.025	1.5	23	32	3310
33	1.0	0.025	1.0	23	20	4693*
34	2.0	0.025	1.0	23	20	3949
35	1.0	0.025	1.0	23	32	4041
36	2.0	0.025	1.0	23	32	3725
37	1.5	0.000	1.0	18	26	584
38	1.5	0.050	1.0	18	26	4301
39	1.5	0.000	1.0	28	26	324
40	1.5	0.050	1.0	28	26	204
41	1.5	0.025	1.0	23	26	6133
42	1.5	0.025	1.0	23	26	4839
43	1.5	0.025	1.0	23	26	4896
44	1.5	0.025	1.0	23	26	4912
45	1.5	0.025	1.0	23	26	5560
46	1.5	0.025	1.0	23	26	5933
47	1.0	0.000	1.0	23	26	6721
48	2.0	0.000	1.0	23	26	6875
49	1.0	0.050	1.0	23	26	3190*
50	2.0	0.050	1.0	23	26	5688
51	1.5	0.025	0.5	18	26	442
52	1.5	0.025	1.5	18	26	286
53	1.5	0.025	0.5	28	26	237
54	1.5	0.025	1.5	28	26	286
55	1.5	0.000	1.0	23	20	3152
56	1.5	0.050	1.0	23	20	3368
57	1.5	0.000	1.0	23	32	4686
58	1.5	0.050	1.0	23	32	4030
59	1.0	0.025	0.5	23	26	1482*
60	2.0	0.025	0.5	23	26	6688
61	1.0	0.025	1.5	23	26	2720*
62	2.0	0.025	1.5	23	26	6490
63	1.5	0.025	1.0	18	20	119*
64	1.5	0.025	1.0	28	20	975
65	1.5	0.025	1.0	18	32	1823
66	1.5	0.025	1.0	28	32	257*
67	1.5	0.000	0.5	23	26	5449
68	1.5	0.050	0.5	23	26	4172
69	1.5	0.000	1.5	23	26	5656
70	1.5	0.050	1.5	23	26	3849
71	1.0	0.025	1.0	18	26	303
72	2.0	0.025	1.0	18	26	416
73	1.0	0.025	1.0	28	26	277
74	2.0	0.025	1.0	28	26	221
75	1.5	0.025	0.5	23	20	2860*
76	1.5	0.025	1.5	23	20	2659
77	1.5	0.025	0.5	23	32	3648
78	1.5	0.025	1.5	23	32	3621

79	1.0	0.025	1.0	23	20	2324*
80	2.0	0.025	1.0	23	20	3159
81	1.0	0.025	1.0	23	32	3366
82	2.0	0.025	1.0	23	32	4194
83	1.5	0.000	1.0	18	26	468
84	1.5	0.050	1.0	18	26	292
85	1.5	0.000	1.0	28	26	300
86	1.5	0.050	1.0	28	26	331
87	1.5	0.025	1.0	23	26	5834
88	1.5	0.025	1.0	23	26	5145
89	1.5	0.025	1.0	23	26	5080
90	1.5	0.025	1.0	23	26	4925
91	1.5	0.025	1.0	23	26	5895
92	1.5	0.025	1.0	23	26	5659
93	1.0	0.000	1.0	23	26	7041
94	2.0	0.000	1.0	23	26	7311*
95	1.0	0.050	1.0	23	26	4938
96	2.0	0.050	1.0	23	26	6041
97	1.5	0.025	0.5	18	26	461
98	1.5	0.025	1.5	18	26	217
99	1.5	0.025	0.5	28	26	218
100	1.5	0.025	1.5	28	26	278
101	1.5	0.000	1.0	23	20	4077
102	1.5	0.050	1.0	23	20	3078
103	1.5	0.000	1.0	23	32	5920
104	1.5	0.050	1.0	23	32	4099
105	1.0	0.025	0.5	23	26	6069
106	2.0	0.025	0.5	23	26	7574*
107	1.0	0.025	1.5	23	26	5918
108	2.0	0.025	1.5	23	26	5801
109	1.5	0.025	1.0	18	20	129*
110	1.5	0.025	1.0	28	20	969
111	1.5	0.025	1.0	18	32	1930
112	1.5	0.025	1.0	28	32	206*
113	1.5	0.000	0.5	23	26	6139
114	1.5	0.050	0.5	23	26	6664*
115	1.5	0.000	1.5	23	26	6421
116	1.5	0.050	1.5	23	26	4653
117	1.0	0.025	1.0	18	26	254
118	2.0	0.025	1.0	18	26	372
119	1.0	0.025	1.0	28	26	339
120	2.0	0.025	1.0	28	26	322
121	1.5	0.025	0.5	23	20	4676
122	1.5	0.025	1.5	23	20	3146
123	1.5	0.025	0.5	23	32	5214*
124	1.5	0.025	1.5	23	32	4080
125	1.0	0.025	1.0	23	20	3511
126	2.0	0.025	1.0	23	20	3373
127	1.0	0.025	1.0	23	32	4315
128	2.0	0.025	1.0	23	32	4735
129	1.5	0.000	1.0	18	26	460
130	1.5	0.050	1.0	18	26	232
131	1.5	0.000	1.0	28	26	438
132	1.5	0.050	1.0	28	26	299
133	1.5	0.025	1.0	23	26	6264
134	1.5	0.025	1.0	23	26	5231
135	1.5	0.025	1.0	23	26	5238
136	1.5	0.025	1.0	23	26	4921
137	1.5	0.025	1.0	23	26	6174
138	1.5	0.025	1.0	23	26	6462

Experimental conditions for inducer (I), repressor (Re) and glycerol (G) concentrations expressed as a w/v percentage, temperature (T) and culture duration (D).

\*Samples with a soluble activity considered as aberrant (compared to the corresponding triplicates) and not taken into account to generate the model.

**Table S6. Response surface quadratic model analysis of soluble enzyme production with ArcticExpress DE3 *E. coli* strain / pET-53 expression vector (A53), BL21 star DE3 *E. coli* strain / pET-53 expression vector (B53) and BL21 star DE3 *E. coli* strain / pET-53 expression vector (B55) productions****a. Polynomial function regression coefficients**

Coded factors <sup>a</sup>	Regression coefficients			Standard error			p-value <sup>b</sup>		
	A53	B53	B55	A53	B53	B55	A53	B53	B55
Constant	-1.62*10 <sup>5</sup>	-6.06*10 <sup>4</sup>	-1.88*10 <sup>5</sup>	7782.97	3812	8962	0.000	0.000	0.000
I	1233	-176.5	-3308	2106.58	3035	2543	0.560	0.954	0.196
Re	56073	-5.04*10 <sup>4</sup>	-3.40*10 <sup>4</sup>	46350.2	3.00*10 <sup>4</sup>	4.49*10 <sup>4</sup>	0.230	0.095	0.452
G	4178	-878.5	-2901	2072.20	1538	2368	0.047	0.569	0.223
T	9239	5604.7	1.21*10 <sup>4</sup>	360.49	168	381	0.000	0.000	0.000
D	4232	12.9	4519	284.08	143	336	0.000	0.929	0.000
I <sup>2</sup>	-1094	-399.8	542	376.26	1137	433	0.005	0.726	0.213
Re <sup>2</sup>	2.73*10 <sup>5</sup>	-5.68*10 <sup>4</sup>	2.20*10 <sup>5</sup>	1.50*10 <sup>5</sup>	1.14*10 <sup>5</sup>	1.68*10 <sup>5</sup>	0.074	0.618	0.192
G <sup>2</sup>	-626	-115.6	-251	283.64	285	424	0.107	0.686	0.555
T <sup>2</sup>	-162	-117.8	-210	4.52	3	5	0.000	0.000	0.000
D <sup>2</sup>	-49	1.1	-46	2.83	2	3	0.000	0.585	0.000
I*Re	1.14*10 <sup>4</sup>	3.64*10 <sup>4</sup>	3.16*10 <sup>4</sup>	10827.1	1.68*10 <sup>4</sup>	1.36*10 <sup>4</sup>	0.294	0.033	0.022
I*G	22	-682.2	35	610.58	886	682	0.971	0.443	0.959
I*T	-1	-2.5	-3	60.19	80	59	0.981	0.975	0.957
I*D	60	27.3	43	43.01	70	55	0.164	0.697	0.432
Re*G	-4.45*10 <sup>4</sup>	2547.7	-1.08*10 <sup>4</sup>	10322.7	8402	1.23*10 <sup>4</sup>	0.000	0.762	0.381
Re*T	1024	-164.0	221	1484	801	1175	0.492	0.838	0.851
Re*D	-3342	1417.0	-1529	90.26	738	979	0.000	0.058	0.121
G*T	-42	-10.2	25	59.02	40	59	0.483	0.800	0.666
G*D	-59	60.3	106	47.58	37	54	0.216	0.104	0.053
T*D	-65	-6.9	-96	8.00	4	9	0.000	0.098	0.000

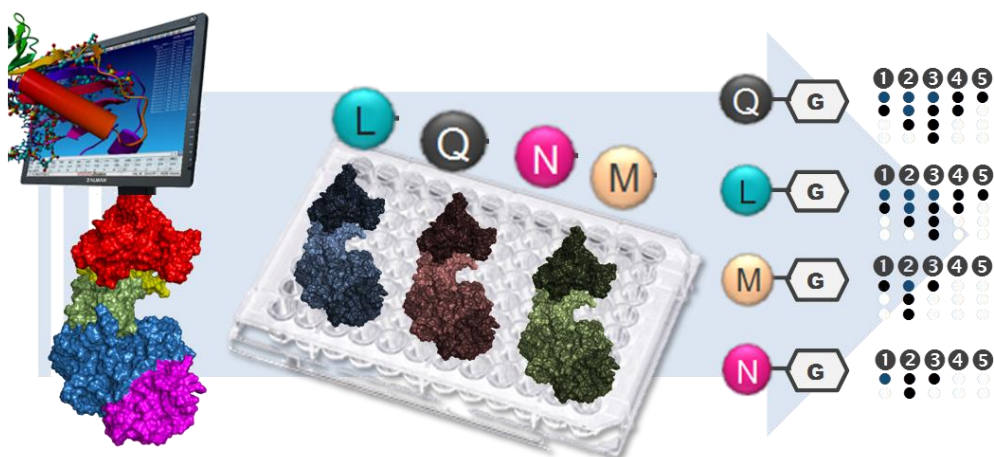
<sup>a</sup> Factors coded as following : Inducer (I), repressor (Re), glycerol (G), temperature (T), culture time (D).<sup>b</sup> p-value < 0.05 indicates a statistical significance of the factor.**b. Analysis of variance (ANOVA)**

Coded source	Sum of Square			Degree of Freedom			F-value			p-value <sup>b</sup>		
	A53	B53	B55	A53	B53	B55	A53	B53	B55	A53	B53	B55
Model	3.34*10 <sup>8</sup>	2.55*10 <sup>8</sup>	6.16*10 <sup>8</sup>	20	20	20	92.13	106.04	118.97	0.000	0.000	0.000
Residual	1.50*10 <sup>7</sup>	1.25*10 <sup>7</sup>	2.59*10 <sup>7</sup>	83	104	100						
Lack of fit	6.18*10 <sup>6</sup>	1.76*10 <sup>6</sup>	8.71*10 <sup>6</sup>	17	19	18	2.71	0.73	2.31	0.002	0.777	0.006
Pure error	8.86*10 <sup>6</sup>	1.07*10 <sup>7</sup>	1.72*10 <sup>7</sup>	66	85	82						
Total	3.49*10 <sup>9</sup>	2.68*10 <sup>8</sup>	6.41*10 <sup>8</sup>	103	124	120						
R <sup>2</sup>	0.957	0.953	0.960									
Adj R <sup>2</sup>	0.947	0.944	0.952									

<sup>b</sup> p-value < 0.05 indicates a statistical significance of the factor.

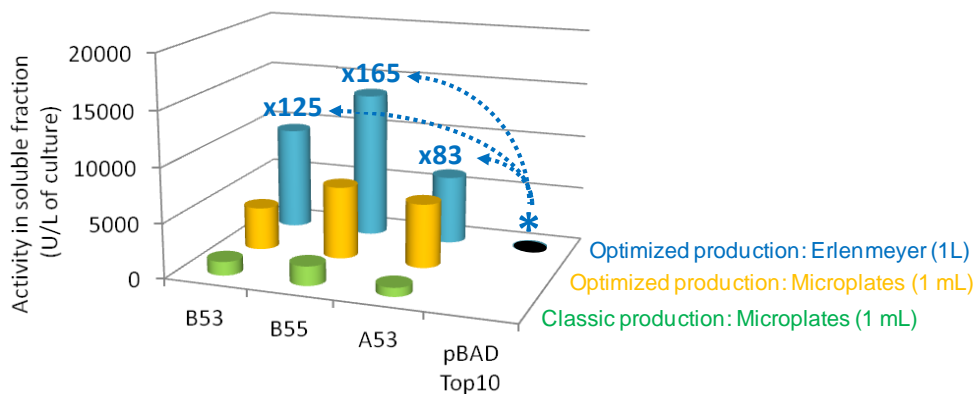
# Chapter IV

## Rational design of an $\alpha$ -(1→2) branching sucrase for flavonoid glucoside diversification





As shown in chapter III, the production of soluble and active  $\alpha$ -(1 $\rightarrow$ 2) branching sucrose was remarkably improved (Figure 54). The best production levels of soluble enzyme were obtained with *E. coli* BL21 Star DE3 cells transformed with the pET-55 vector, for both microplates and Erlenmeyer productions, as shown on Figure 54.



**Figure 54.** Activity recovered in the soluble fraction for production with *E. coli* BL21 Star DE3 cells transformed with the pET-53 and pET-55 vector (B53 and B55, respectively) and *E. coli* ArticExpress cells transformed with the pET-53 vector (A53), in microplates and Erlenmeyer. Optimized production were conducted using the optimal conditions predicted by RSM analyzes whereas classic productions correspond to the use of classic ZYM-5052 medium. Improvements reached for 1L-production are calculated using the production yield obtained in previous production of the enzyme with *E. coli* TOP10 cells transformed with the pBAD vector.

Notably very good yields were also obtained with B53. In addition, the enzyme produced was more efficiently purified and separated from its degraded forms (6xHis affinity chromatography) than the enzyme produced with B55 (StreptII-tag affinity chromatography)

Therefore, we selected this B53 expression system to produce the enzymes of the  $\alpha$ -(1 $\rightarrow$ 2) branching sucrose libraries. A structurally-guided approach based on flavonoid docking was applied to construct two libraries of around 3,000 variants. Screening assays were implemented. The promiscuity of the best hits was further evaluated by testing more than 5 different molecular structures. The achievements are presented and discussed detailed in the following chapter.

## I. Introduction

Bioavailability of flavonoids might be increased by glucosylation, but this phenomenon is also highly dependent on their molecular structures. With the view of diversifying the panel of accessible flavonoid glucosides, a mutant library of *N. polysaccharea* amylosucrase (*NpAS*) was recently screened, leading to the identification of mutants remarkably improved for luteolin glucosylation (Malbert et al., 2014a; Thesis Chapter II). This library was also screened for the glucosylation of two other flavonoids, diosmetin and quercetin. However, no mutant able to glucosylate diosmetin could be found, and the best amylosucrase mutant identified for quercetin glucosylation achieved a conversion rate of only 16 %.

Another transglucosylase, the  $\alpha$ -(1→2) branching sucrase, a member of the GH70 family, was thus envisaged for the glucosylation of these compounds. Compared to amylosucrases, this transglucosylase is naturally designed to branch very efficiently  $\alpha$ -(1→6) dextrans. It possesses an active site that is way less hindered than that of GH13 amylosucrases, what may be of particular interest to accommodate large acceptor molecules. In addition, this enzyme is unable to produce polymers using sole sucrose. This represents a clear advantage for the downstream process. Finally, as the glucosidic linkage specificity of amylosucrase and that of the  $\alpha$ -(1→2) branching sucrase are different, various glucosylation patterns and thus, novel products, could be expected.

Preliminary studies revealed that the  $\alpha$ -(1→2) branching sucrase ( $\Delta N_{123}$ -GBD-CD2) was able to glucosylate luteolin, quercetin and diosmetin at 5 mM, with conversion rates of 40%, 60% and 2%, respectively. Moreover, the three-dimensional structure of this enzyme,  $\Delta N_{123}$ -GBD-CD2, was recently solved (Brison et al., 2012), thus opening the way to molecular modeling studies aiming at improving flavonoid glucosylation through semi-rational engineering. To allow the development of screening assays in microtiter plates, we first optimized the recombinant  $\alpha$ -(1→2) branching sucrase production in microtiter plates and defined conditions that improved by 40-fold the levels of soluble enzyme production for B53 expression system (Malbert et al., 2014b; Thesis Chapter III).

In this work, the focus was placed on the semi-rational design of  $\alpha$ -(1→2) branching sucrase mutants improved for flavonoid glucosylation. Molecular modeling studies were carried out to identify amino acid residues of the catalytic pocket that could be targeted by mutagenesis to favor flavonoid recognition. Prior to that, a rapid solid state pH-based screening method had to be developed to first discriminate and select mutants still active on sucrose before to assay their

capacity to glucosylate quercetin and diosmetin. A set of variants, constituting a platform of enzymes improved for flavonoid glucosylation, was retained and evaluated for the glucosylation of a set of six distinct flavonoids, namely quercetin, luteolin, morin, naringenin, apigenin and chrysin. Altogether these results were used to get insight on the relations between amino acid mutations, glucosylation performances and product profiles in order to provide useful information for further design of flavonoid glucosylation tools.

## II. Results and discussion

### II.1. Semi-rational design and $\Delta N_{123}$ -GBD-CD2 mutant libraries creation

To identify the amino acid residues of the active site that could be targeted by mutagenesis to improve the recognition of flavonoids, we first built a three-dimensional model of the enzyme using the atomic coordinates of the  $\alpha$ -(1→2) branching sucrase ( $\Delta N_{123}$ -GBD-CD2) (PDB accession code 3TTQ; Brison et al., 2012). Next, we used as a guide the crystallographic structure of GTF-180 from *Lactobacillus reuteri* in complex with sucrose (PDB accession code 3HZ3; Vujičić-Žagar et al., 2010). We manually docked in the active site the donor substrate, sucrose. The glucosyl unit from sucrose was then extracted to build all potential mono-glucosylated products derived from luteolin, this latter being considered here as a representative flavonoid model. The docked molecules were mono-glucosylated at each of the four reactive hydroxyl groups of the luteolin, meaning positions C3' or C4' from the B-ring or at positions C5 or C7 from the A-ring (Figure 55).

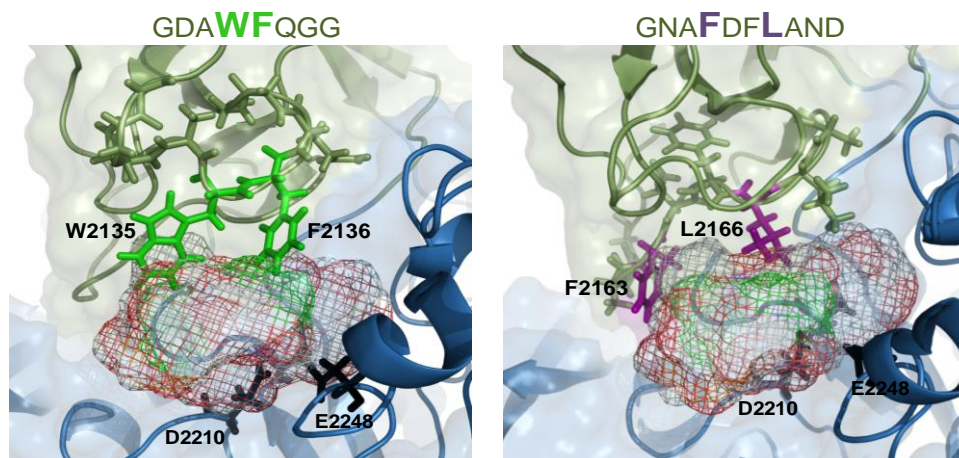
Comparative analysis of the binding modes of sucrose (donor) and luteolin glucosides (products) enabled to map all key residues involved in either i) the catalytic machinery (D2210, E2248, D2322), ii) the specific recognition of the glucosyl moiety transferred from donor sucrose and bound at subsite -1 (H2321, D2643, Y2650, F2641, Q2694), iii) the amino acid residues involved in recognition of fructosyl moiety from donor sucrose at subsite +1 (F2136, L2166, F2214), and iv) the amino acid residues from subsites +1, +2, +3 potentially involved in the recognition of luteolin, that could be targeted by mutagenesis but which were not essential for catalysis (W2135, F2136, F2163, L2166, A2211, D2213, F2214, G2324, Q2326).

Due to the distinct glucosylation patterns, the flavonoid glucoside molecules were docked in very different orientation in the active site, and therefore, the interactions established between



the molecules and the amino acid residues forming the active site differed significantly from one molecule to another. As a result, we chose to adopt an engineering strategy that targeted four key amino residues from subsites +1, +2, and +3, which did not affect catalytic mechanism or sucrose binding, and which were found systematically involved in interactions, although of different nature, with the mono-glucosylated luteolin molecules. The idea was then to authorize random substitution of these key selected positions by amino acids that could enable the formation of either hydrogen bonds or van der Waals interactions with the luteolin pattern to increase the affinity toward the non-natural acceptor, or unclutter the active site in order to make enough space for accommodation of the bulky luteolin conjugated system.

From a detailed analysis of molecular modeling studies combined with a sequence analysis of amino acid conservation within GH70 family, four amino acids were selected, W2135, F2136, F2163 and L2166, to proceed to the construction of small libraries of mutants of the  $\alpha$ -(1 $\rightarrow$ 2) branching sucrase. Previously, we demonstrated the benefits of pairwise combinations to create or enhance specificity toward non-natural acceptors of GH13 amylosucrase (Champion et al., 2012). Given the location of these amino acids in a conserved motif, characteristic of GH family 70, upstream the helix  $\alpha$ 3 of the  $\alpha$ -(1 $\rightarrow$ 2) branching sucrase, we opted for a similar strategy involving the construction of two sub-libraries. The first library targeted residues W2135 and F2136 while the second library focused on residues F2163 and L2166. In these double-mutant libraries, pairwise combinations of polar, charged and hydrophobic residues at the positions were authorized.



**Figure 55.** View of  $\alpha$ -(1 $\rightarrow$ 2) branching sucrase active site pocket showing the 2 amino acid mutagenesis targets of each library (light green and purple) and the adjacent ones (green). The nucleophile (D2210) and the acid/base (E2248) catalysts are shown in black. Luteolins mono-glucosylated at each of the 4 reactive hydroxyls are docked in the active site and represented as a mesh surface.

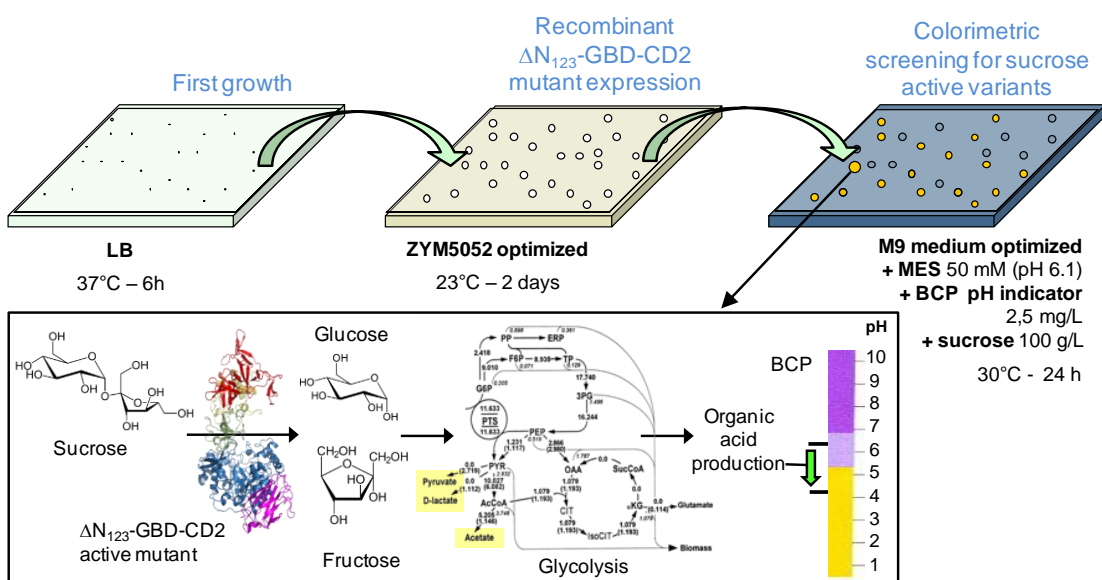
The two libraries were subsequently constructed by mutating the targeted positions, W2135 and F2136 or F2163 and L2166, using a saturation mutagenesis approach. To reduce codon degeneracy, we used NDT codons for saturation mutagenesis, which results in the exclusion of some amino acids but conversely enables to explore more extensively the generated diversity while reducing screening efforts (Reetz et al., 2008).

## II.2. Setting up the primary solid-state colorimetric screening assay

In parallel to the libraries design, we developed a rapid and efficient primary screening assay to select mutants still able to use sucrose.

### II.2.1. General principle of the screening method

The screening method (Figure 56) was inspired from Champion et al. (2010). The recombinant *E. coli* BL21 Star clones encoding enzyme mutants were first plated on LB medium supplemented with sucrose. This strain is derived from *E. coli* K12 and is unable to use sucrose as substrate. When transformed with a plasmid encoding active  $\alpha$ -(1 $\rightarrow$ 2) branching sucrase, the recombinant enzyme hydrolyzes sucrose and releases fructose and glucose. The monosaccharides can then enter into the glycolytic pathway to liberate acids and induce a pH decrease that can be detected, at least theoretically with an appropriate pH colorimetric indicator.



**Figure 56. Solid-state pH-based colorimetric assay for screening of sucrose active mutants.**

To set up the screening assay, we had first to check that in the absence of an active  $\alpha$ -(1→2) branching sucrase, no acid production occurs and no change in the pH indicator color is observed.

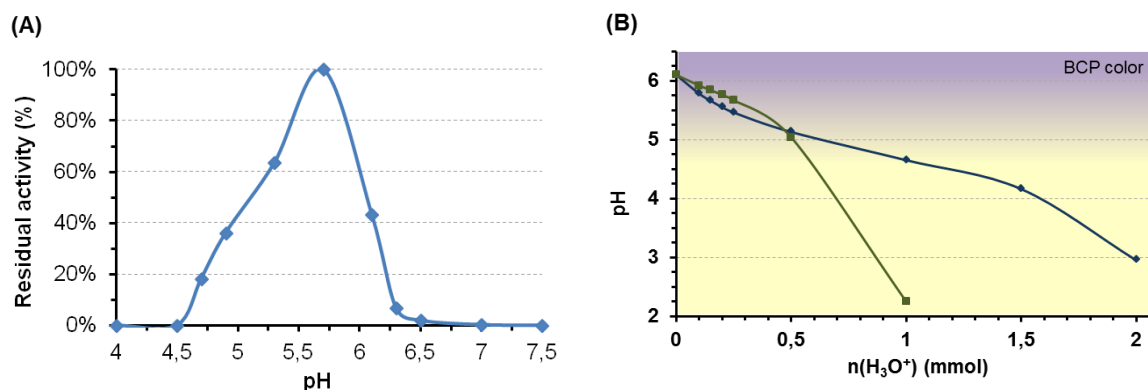
### II.2.2. Development and validation of the screening method

Several criteria had first to be fulfilled to ensure that the screening assay was achievable.

- The recombinant enzyme production on solid medium had to be optimized. In particular, we had to check that the culture medium supplemented with lactose, glucose or glycerol, optimized for liquid production, did not provoke any non-specific acidification that could lead to false positive identification.
- An appropriate pH indicator allowing detection of active clones at pH values that were not detrimental for enzyme activity had also to be identified.

To ensure a sufficient enzyme production avoiding a non-specific acidification due to the consumption of lactose, glucose and glycerol added to the solid-state expression medium, a series of assays were conducted. The amount of lactose (inducer), glucose (repressor) and glycerol were first adjusted to 0.5%, 0.01%, and 0.5% (w/v), respectively, for the expression step. These amounts were further decreased to 0% glucose, 0.25% lactose and 0.15% glycerol for the colorimetric visualization step. Such values are significantly lower than those previously defined for optimal production in liquid medium (Malbert et al., 2014b; Thesis Chapter III).

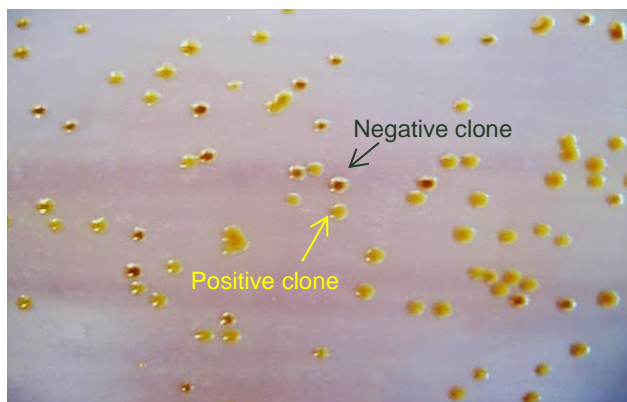
Concerning the resistance of the  $\alpha$ -(1→2) branching sucrase to various pH values, we analyzed the effect of different pH values on the enzyme activity after 24 h storage at 4°C (Figure 57 A). It clearly appeared that neutral pH, which is optimal for *E. coli* growth, was not suitable for the  $\alpha$ -(1→2) branching sucrase activity. Indeed, above pH 6.5, the enzyme was totally inactivated. In addition, for pH values below 5.0, *E. coli* growth and metabolism is known to be highly limited (Presser et al., 1997). The best pH compromise for growth, expression and colorimetric visualization was found to be of 6.1. At this pH value, 40% of the enzyme activity was retained. This also led to the selection of bromocresol purple pH indicator (BCP) which is characterized by a color change between 5.2 and 6.8.



**Figure 57. (A) Residual activity of the  $\alpha$ -(1→2) branching sucrase as a function of pH. Residual activities were calculated by comparison to activities determined before enzyme storage. (B) pH and BCP color variation versus  $\text{H}_3\text{O}^+$  addition. Blue and green curves correspond to sodium acetate and MES buffer at 50 mM and initial pH value of 6.1, respectively.**

Sodium acetate buffers at pH 5.4 or 5.75 are the usual buffer used for  $\Delta\text{N}_{123}\text{-GD-CD2}$  enzyme (Brison et al., 2009). However, with a  $\text{pK}_a$  of 4.6, its buffering capacity at pH values below 6.1 prevented the detection of rapid color change and differences between clones expressing active and non-active  $\alpha$ -(1→2) branching sucrase (Figure 57, B). In addition, acetate concentrations above 20 mM are toxic for *E. coli*, and damaging for biomass production, production of recombinant proteins and stability of intracellular proteins. Acetate is also known to specifically inhibit the consumption of both glucose and fructose (Holms, 1996). According to these considerations and the unsuccessful attempts of screening assays using 5 mM to 50 mM of sodium acetate, a MES buffer at 50 mM was assayed. Its  $\text{pK}_a$  of 6.15 well matched the pH values needed to rapidly identify the indicator color change (Figure 57). Moreover, previous studies reported that the use of MES buffer had no significant impact on bacteria metabolism and *E. coli* growth (Tucker et al., 2002).

To set up this buffer and validate the pH-based screening assay, we applied the method to a mix of freshly transformed *E. coli* BL21 Star expressing either wild-type enzyme or an inactive mutant (E2248Q). Upon acid release in the medium, the color of some colonies turned from purple-green (pH 6.1) to yellow (pH < 5.2), revealing the ability of the active clones to cleave sucrose and induce a pH drop (Figure 58). We picked 8 purple-green clones (negative) and 8 yellow clones (positive) and extracted the plasmidic DNA for sequencing. All purple-green clones corresponded to inactive mutant expressing clones, whereas all the yellow ones harbored a plasmid containing the  $\alpha$ -(1→2) branching sucrase encoding gene, thus validating the screening method.



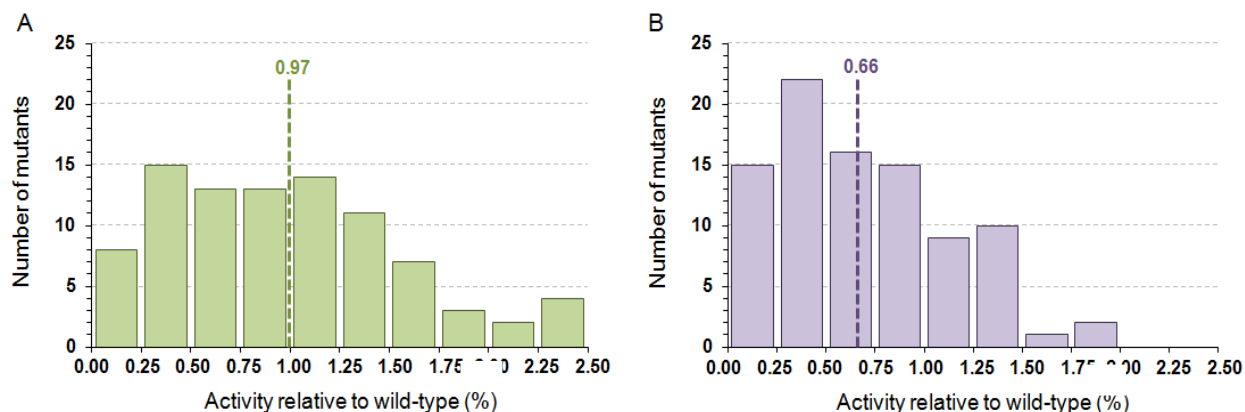
**Figure 58. Solid-state pH-based colorimetric screening methods validation with a mix of clones expressing wild-type or inactive  $\alpha$ -(1 $\rightarrow$ 2) branching sucrase (mutant E2248Q).** Yellow colonies correspond to positive clones, whereas purple-green ones are negative clones.

### II.2.3. Screening of the libraries

The procedure was subsequently used to screen the two libraries. The ratio of active clones was determined for each library, and a set of active mutants was picked out to further analyze their capacity to glucosylate flavonoids.

For each library, we screened around 1,500 recombinant clones to ensure a high coverage of the variant space (430 required for 95% coverage). For the library F2135-W2136, 19% of the screened clones were still active on sucrose and turned yellow. Concerning the library F2163-L2166, only 16% of the clones were still active. For each library, 92 of the active clones were randomly picked out and cultured in 96-well microplates. To confirm and evaluate the activity of the selected clones on sucrose, DNS assays were carried out on enzymatic soluble extracts. The activities were determined with respect to that of the wild-type in order to get better insight in the impact of mutations on sucrose uptake (Figure 59).

All the selected clones were confirmed to be active on sucrose. Library F2163-L2166 appeared negatively affected by mutations for sucrose activity, as the average relative activity value of the selected mutants was 0.66 (Figure 59). In contrast, in library W2135-F2136, most of the mutant showed an activity close to that of the wild-type. These two positions seem less sensitive to mutations than the F2163 and L2166 ones. Moreover, six clones presented a two-fold increase of their sucrase activity, when compared to the wild-type. Nevertheless, sucrase activity cannot be directly correlated to acceptor glucosylation.



**Figure 59. Distribution of the mutant activity (relative to wild-type) according to their activity on sucrose.** (A) library W2135-F2136 (green) and (B) library F2163-L2166 (purple). Green and purple dotted lines represent the average activity calculated for library W2135-F2136 and F2163-L2166, respectively.

Thus, all the 184 selected mutants were tested for their capacity to glucosylate flavonoids.

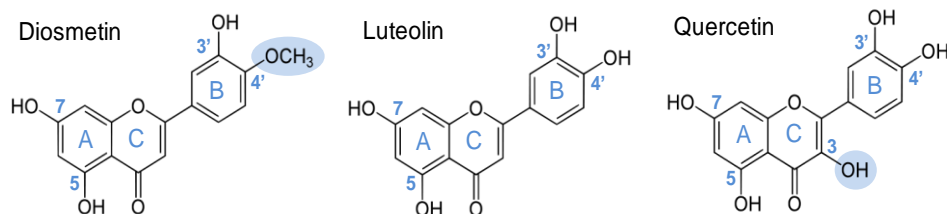
### II.3. Secondary screening: diosmetin and quercetin glucosylation

All the 92 clones of each library were grown on ZYM5052 optimized medium, in microplates, and lysed after growth. Acceptor reaction in the presence of diosmetin or quercetin, and sucrose were first carried out with the lysates in microplate format (Figure 60). The acceptor reaction products were analyzed by LC-MS and the glucosylation yields of diosmetin or quercetin calculate.

Of note, these two flavonoids were not or poorly glucosylated by other tested glucansucrases, even though they are structurally close to luteolin. Indeed, when previously used as acceptor with *NpAS* library (Malbert et al., 2014a; Thesis Chapter II), no glucosylation occurred for diosmetin and the best *NpAS* mutant converted only 16% of the quercetin at 5 mM (data not shown). Moreover, previous studies showed that only about 23% of quercetin at 2 mM could be glucosylated, using a glucansucrase from *L. mesenteroides* B-1299 (Moon et al., 2007) and around 4%, using the alternansucrase from *L. mesenteroides* NRRL B-23192 (Bertrand et al., 2006). In addition, our preliminary assays showed that the  $\alpha$ -(1 $\rightarrow$ 2) branching sucrose glucosylated diosmetin and quercetin at 5 mM with conversion rates around 2% and 60%, respectively.



### II.3.1. Analysis of the conversion rates



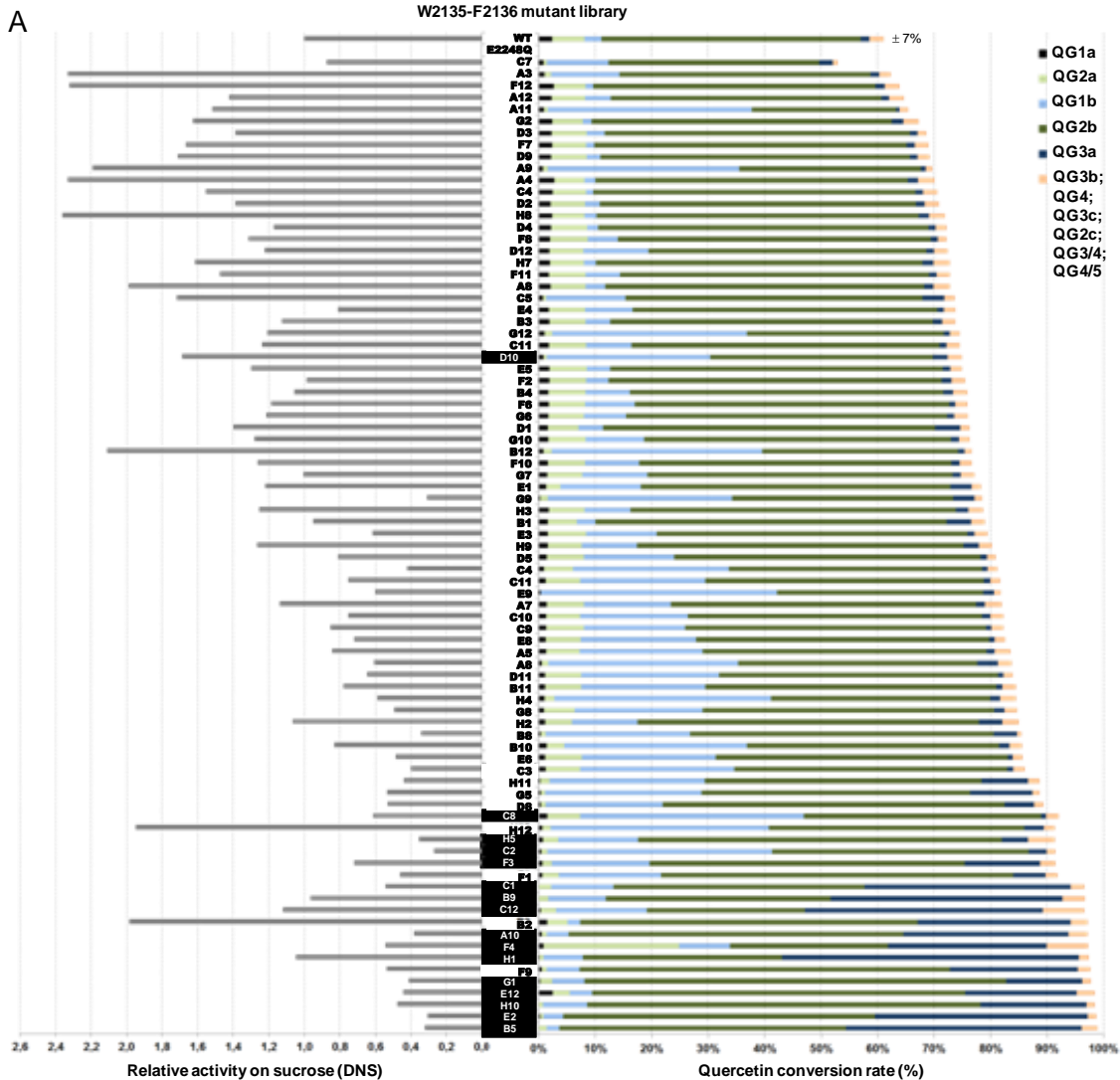
**Figure 60. Molecular structure of luteolin, diosmetin and quercetin.** Blue circles highlights the structural differences with luteolin.

The conversion of diosmetin never exceeded 5%. Even if some mutants are slightly improved compared to the wild-type enzyme, they remained unable to efficiently glucosylate diosmetin. The results obtained for quercetin glucosylation were more conclusive and reported in Figure 61. In the screening conditions, the  $\alpha$ -(1 $\rightarrow$ 2) branching sucrase glucosylated 60% ( $\pm$  7%) of the quercetin. For library W2135-F2136, 80% of the 92 mutants converted more than 70 % of quercetin. For library F2163-L2166, only 41% of the mutants showed an improved conversion rate of quercetin, compared to the wild-type enzyme. The library that was the less affected on sucrase activity also provided the highest number of improved mutants for quercetin glucosylation.

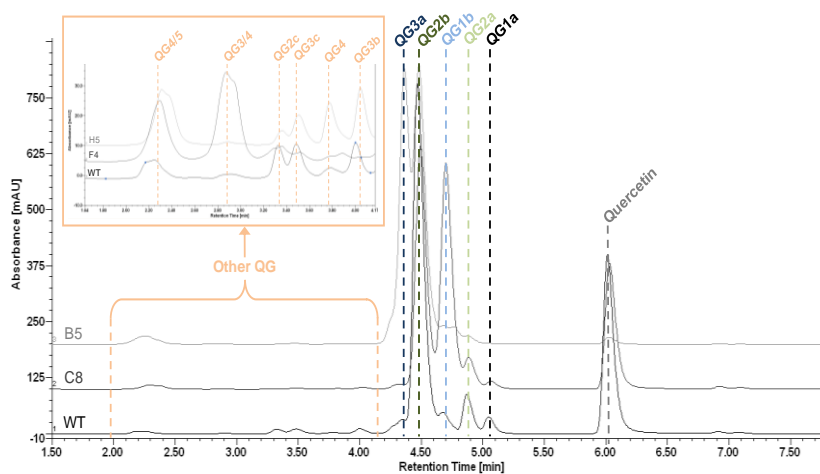
### II.3.2. Analysis of the product profile and sub-selection of the mutants

The profile of the acceptor reaction products obtained from quercetin with wild-type  $\alpha$ -(1 $\rightarrow$ 2) branching sucrase was analyzed by LC-MS. The main synthesized product corresponds to a quercetin diglucoside (QG2b) (Figure 61). Two other quercetin diglucosides (QG2a and QG2c), together with two monoglucosides, and three triglucosides, were also identified in smaller amounts. This demonstrates that the wild-type enzyme can glucosylate quercetin onto, at least, two hydroxylated position. Quercetin glucosides harboring up to five glucosyl units were also detected.

The products synthesized with the mutants isolated from each library were compared to those produced with the wild type enzyme. Mutants enabling the highest conversion of quercetin were selected together with the mutants whose glycosylation profile diverged from that of the wild-type.



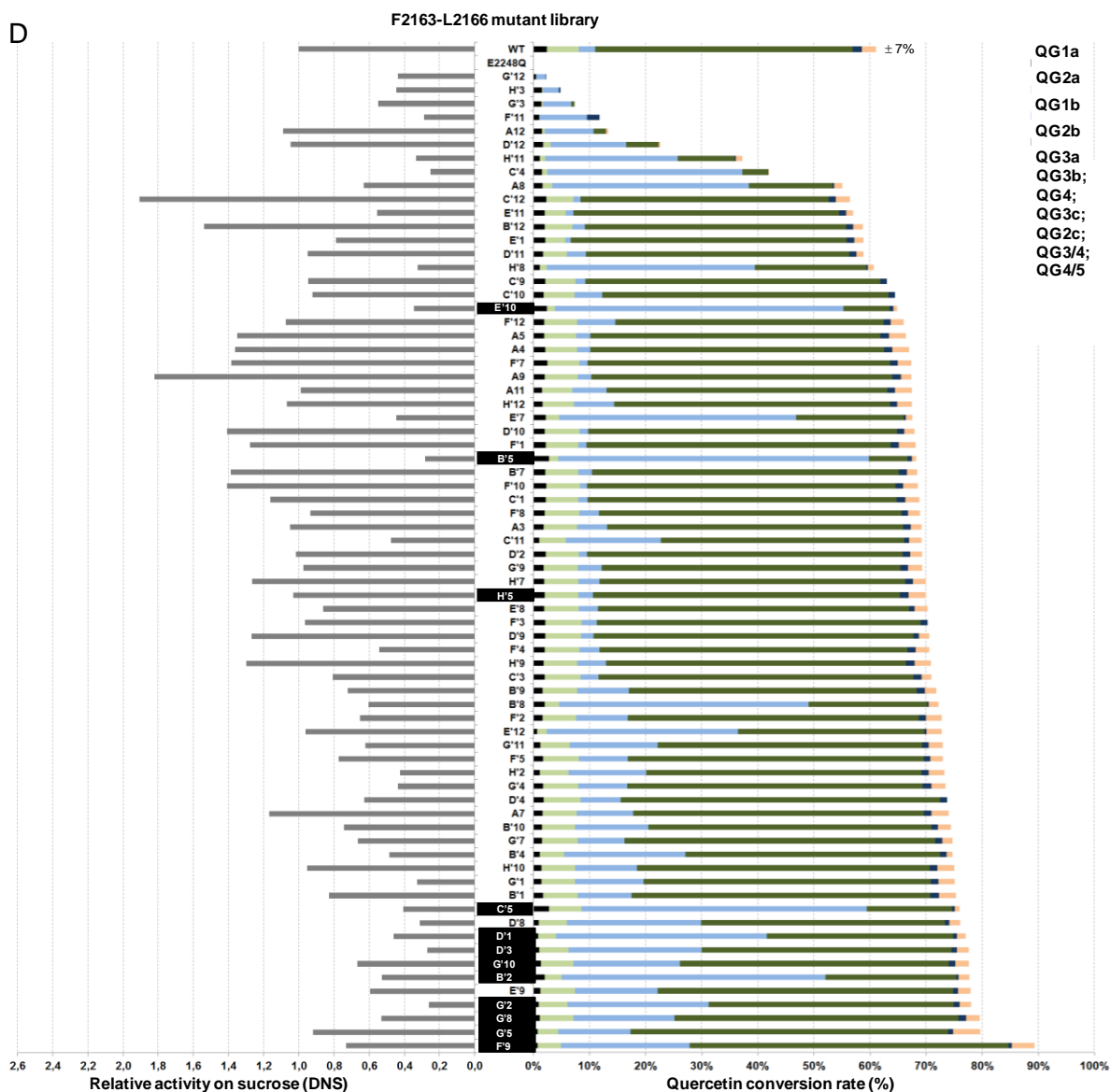
**B**



**C**

Quercetin glucoside	m/z detected	Glucosyl units
QG1a	462.8	1
QG2a	625.0	2
QG1b	462.9	1
QG2b	625.0	2
QG3a	787.1	3
QG3b	787.0	3
QG4	948.9	4
QG3c	787.1	3
QG2c	625.0	2
QG3/4	787.2; 948.9	3; 4
QG4/5	948.9; 1111.5	4; 5





**Figure 61. Results of the secondary screening of  $\alpha$ -(1→2) branching sucrose mutant libraries, for their capacity to glucosylate quercetin.** (A) Relative activities on sucrose (gray bars) and quercetin conversion rates for mutants belonging to library W2135-F2136. The proportion of each glucoside derivative is represented in a specific color. Selected mutants are highlighted by a black box. (B) Quercetin acceptor reaction product profiles obtained with wild-type (WT) and interesting mutants. (C) Mass spectrometry results for each products and determination of the number of glucosyl units. Quercetin glucosides are named according to the number of glucosyl units linked. (D) Relative activities on sucrose (gray bars) and quercetin conversion rates with mutants of library F2163-L2166. The proportion of each glucoside derivative is represented in a specific color. Selected mutants are highlighted by a black box.

In library W2135-F2136, we first retained 15 mutants that converted quercetin with conversion rate above 90% (Figure 61). Among them, the mutant F4 exhibited a particularly interesting profile. This mutant synthesized the highest amount of QG2a (4 times more than the wild-type enzyme and at least 2.5 times more than the other mutants). It was also the best producer of quercetin tetra- and penta- glucosides (QG3/4 and QG4/5). Mutants B5, C12, B9, H1 and C1 showed an improved production of quercetin triglucoside QG3a. The proportion of this compound represents 40 to 50% of the glucosylated products versus 2% only of the products synthesized with the wild-type enzyme. Mutant D10 was also retained from this library because they synthesized higher amounts of QG1b monoglucoside compared to the wild-type enzyme. Thus, 16 mutants were selected for library W2135-F2136.

Only 12 mutants were selected from the second library. None of the screened mutant was able to glucosylate quercetin with conversion rate higher than 90%. We selected 9 of the best mutants for which quercetin conversion ranged from 75% to 89%. In this library, no mutants were improved for the production of quercetin derivatives harboring three or more glucosyl units. Finally, we retained 3 mutants (B'5, C'5 and E'10) that catalyzed the synthesis of QG1b monoglucoside as the main product.

The 28 retained mutants were sequenced (Figure 62). Mutants B9 and C12, exhibiting equivalent quercetin glucosylation profiles and conversion rates, presented the same mutations, W2135S and F2136L (Figure 61 A). In this library, substitutions by hydrophobic amino acids were slightly more frequent than substitution by polar residues. When looking at the positions individually, only 66% of the mutants exhibit a mutation at position F2136, whereas all of them were mutated at W2135 position, indicating that position W2135 is tolerant to mutations and such mutations favor quercetin glucosylation. Interestingly, the presence of two aromatic amino acids is not favorable for quercetin glucosylation.

Considering library F2163-L2166, among the 12 selected mutants, 4 sequences correspond to the wild-type (D'3, G'2, G'8, and G'10). This result was not very surprising. Indeed, the mutants of library F2163-L2166 were more affected in their sucrase activity than those from library W2135-F2136, indicating a lowest tolerance to mutations. In addition, their quercetin conversion profiles and rates only slightly differed from that of the wild type. D'1 mutations could not be determined due to a sequencing problem.

Position L2166 was found mutated only in two variants and replaced by a structurally close isoleucine amino acid residue. This might signify that leucine is the best amino-acid residue at this position for the accommodation of both quercetin and sucrose in the active site.

Two efficient mutants F'9 and H'5 exhibited an unexpected mutation A2162E and D2164E, respectively, which were not targeted in our constructions.

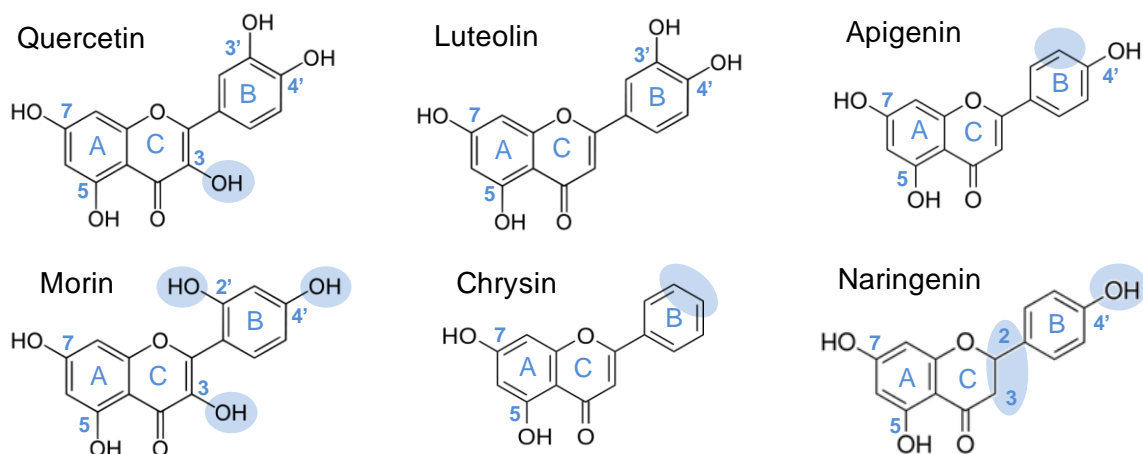
Finally, a total of 23 different mutants (B9 and C12 correspond to the same mutant) were retained, excluding the clones that encoded wild-type protein.



**Figure 62.  $\alpha$ -(1 $\rightarrow$ 2) branching sucrose mutant sequences.** Mutant names and mutated residues of library W2135-F2136 are presented in green. Mutant names and mutated residues of library F2163-L2166 are presented in purple. For each library a weblogo representation of the sequence containing the mutations is presented. Letter height is proportional to the amino acid distribution among the sequenced library.

## II.4. $\Delta N_{123}$ -GBD-CD2 best mutants as a flavonoid general glucosylation platform

The set of 23 mutants was further used to glucosylate six structurally different flavonoids (Figure 63). Quercetin was chosen to validate the results obtained in the preliminary screening. Morin also belongs to the flavonol subclass and possesses two non-vicinal hydroxyl groups on the B-ring. Compared to luteolin, quercetin is substituted with a hydroxyl group at the C3 position of its C-ring. Apigenin and chrysin also belong to the flavone subclass like luteolin but chrysin is not substituted with hydroxyl groups on its B-ring and apigenin is only substituted with one hydroxyl group at the C4' position of the B-ring.



**Figure 63. Molecular structure of the flavonoids tested as acceptor for the 23 selected mutants isolated from the preliminary screening assays.** *Their structural differences are highlighted with blue circles.*

Naringenin was selected to study glucosylation of a flavanone, with less planar structure than chrysin, due to the lack of C2,3-double bond. Comparison of the glycosylation pattern of these compounds should allow acquiring information on the relationships between flavonoid structures and their ability to be glycosylated using the wild-type enzyme and the set of isolated mutants. This set can be considered as a neutral platform of variants with regard to sucrose consumption and to mutant capacity to glucosylate hydroxyl groups of the diosmetin.

The reactions were all performed in the same conditions. The flavonoid global conversion rates are reported in Figure 64.

## II.4.1. Conversion rate analyses

### II.4.1.a. Reproducibility of the screening assays

We first examined the reproducibility of the screening assay by comparing the activity values (on sucrose alone) and the quercetin conversion determined from the first screening assay (entire library) or the second one, noted Q1 and Q2, respectively (Table 16).

**Table 16. Comparison of the activity values and the quercetin conversion rates of the two screening assays.**

Mutant	B5	B9	C12	H10	A10	E2	E12	F4	H1	H5	G1	C2	F3	B'2	D10
<sup>a</sup> Relative activities Q2/Q1 ratio	0,9	0,8	0,8	0,7	0,7	0,8	0,7	0,9	0,6	0,7	0,5	1,6	0,6	0,9	0,9
<sup>b</sup> Conversion rates Q2/Q1 ratio	0,9	0,9	0,9	0,9	0,9	0,9	0,9	0,8	0,8	0,8	0,7	0,8	0,8	0,9	0,9

Mutant	G'5	F'9	C8	D'1	H'5	B'5	C'5	E'10	C1	WT
<sup>a</sup> Relative activities Q2/Q1 ratio	0,8	0,8	0,8	0,6	0,8	0,6	0,4	0,7	0,3	-
<sup>b</sup> Conversion rates Q2/Q1 ratio	0,8	0,7	0,8	0,8	0,8	0,7	0,5	0,6	0,4	0,9

<sup>a</sup> Relative activity Q2/Q1 ratio = ([activity mutant Q2]\*[activity wild-type Q1]) / [activity mutant Q1]\*[activity wild-type Q2])

<sup>b</sup> Conversion rates Q2/Q1 ratio = [conversion rate Q2] / [conversion rate Q1]

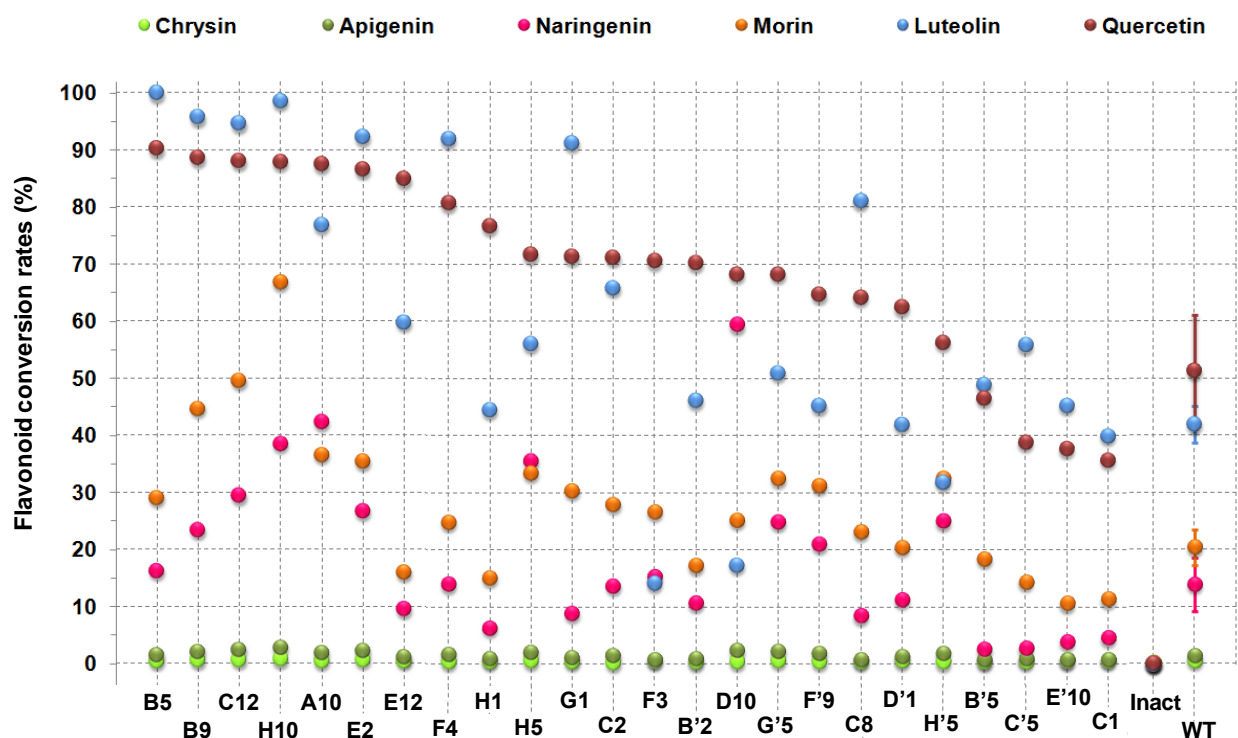
Regarding the mutants, most of the quercetin conversion and activity values determined from Q1 screening assay were higher than those determined from Q2. These variations may be linked to expression variability in *E. coli* BL21Star and render difficult the comparison between mutants. Fortunately, the acceptor reactions conducted with the six flavonoid acceptors were all carried out using the same enzymatic extract, thus eliminating the variability coming from the enzyme extract.

### II.4.1.b. Comparison of flavonoid conversion

First, the standard deviations values calculated for the wild-type enzyme, from 10 replicates, are within the 15% variation usually accepted for such type of screening assays (Figure 64).

Quercetin and luteolin were better converted than any other flavonoids by all the mutants, excepting mutant D10, which glucosylates naringenin more efficiently than luteolin (59% and 17%, respectively). In contrast to the wild-type enzyme, 9 mutants (B5, B9 / C12, E2, F4, G1, C8, C'5, E'10, and C1) converted luteolin better than quercetin. Among them, 6 mutants showed remarkable levels of conversion, above 90%. Only three mutants were less efficient than the wild type for luteolin glucosylation (F3, D10 and H'5).

Concerning morin, 11 mutants glucosylated this flavonol more efficiently than the wild-type enzyme, and up to 67% of conversion was reached for mutant H10 (versus 20% for the wild-type enzyme). Of note, all the previous studies related to flavonoid glucosylation reported that only molecules harboring vicinal hydroxyl groups on the B-ring (C3' and C4' positions) could be glucosylated by glucansucrases from GH family 70 (Bertrand et al., 2006; Meulenbeld and Hartmans, 2000). The results obtained herein with the wild-type  $\alpha$ -(1 $\rightarrow$ 2) branching sucrose and the mutants demonstrate that the conjugation of the hydroxyl groups at C3' and C4' positions on the flavonol B-ring may favor flavonoid reactivity, but is not a pre-requisite for glucosylation. The best mutants were B9 / C12 (same mutations W2135S-F2136L), and H10 (W2135G). The substitution of the tryptophan residue by a less bulky amino-acid like serine or glycine might liberate space and favor the accommodation of the acceptor.



**Figure 64. Conversion rates for the glucosylation of the six selected flavonoids with the 23 mutants.** Flavonoid conversion is represented as a colored spot for each mutant flavonoids: quercetin, red; luteolin, blue; morin, orange; naringenin, fuchsia; apigenin, green; chrysin, yellow. Inact refers to the inactive mutant (E2248Q) and to wild-type enzyme (WT). For WT, standard deviations are given. Relative activities correspond to the mutant activity divided by the wild type one. Q2 refers to the quercetin conversion observed in the present screening assay. Q1 refers to the first conversion determined in the preceding screening of the entire library.

Nine mutants were improved for naringenin glycosylation confirming the fact that vicinal groups on the B-ring are not necessary to catalyze flavonoid glucosylation with enzymes from GH70 family. The mutants exhibiting the best conversion rates for naringenin include mutants H5 (W2135F), H10 (W2135G), A10 (W2135V), which shows a 3-fold increase and mutant D10

(W2135I-F2136Y), that displays a 4-fold increase, compared to the wild-type. The highest conversion rate obtained with mutant D10 was around 60 %. When looking more closely at the mutation, it appears that replacing W2135 by a smaller apolar residue is sufficient to facilitate the accommodation of naringenin. For the best mutant (D10), the additional mutation of phenylalanine 2136 to tyrosine increases naringenin conversion.

Apigenin and chrysin were poorly recognized as acceptor by the wild-type enzyme and conversion rates only reached 2%. With the mutants, conversion never exceeds 5%. When looking at the structural differences between the six flavonoids (Figure 64), apigenin and chrysin both possessing an unbent structure, with a torsion angle below 30°, due to the presence of the C2,3-double bond (van Acker et al., 1996; Amat et al., 2008). In addition, they also harbor only one or no free hydroxyl group on the B-ring. In contrast, naringenin was glucosylated by the wild-type enzyme and also by several mutants in larger proportion. The sole difference between apigenin and naringenin is the presence of the C2,3-double bond of the C-ring for apigenin. Thus, at equivalent structure this double bond appears harmful for glucosylation. This is in accordance with the previous studies, which reported that C2,3-double favor the inhibition of many enzymes, as previously shown (Table 2, p 32) of this thesis. In addition, we may suggest that the glucosylation of the tested flavonoid probably targets hydroxyl groups of the B-ring. Of course, these conclusions will have to be confirmed in depth by a detailed characterization of the flavonoid glucosides.

## II.4.2. Sucrose consumption analysis

For each reaction, the consumption of sucrose was determined after 24 hour reaction, (Table 17). In the absence of flavonoids, the wild-type enzyme entirely consumed sucrose at 292 mM (100 g.L<sup>-1</sup>). Of the 23 mutants, half used more than 80 % of the sucrose and half consumed it from 32 to 80 % (Table 17). This is in agreement with the initial activity previously determined.

When the reactions were carried out in the presence of flavonoids, we observed a decrease of sucrose consumption in many cases. However, the inhibiting effect of most of the flavonoids was moderate for most of the mutant (below 30 %), except for naringenin that strongly reduced sucrose consumption by more than 50% for almost all the mutants.

The mutant C'5 (F2163H) is the more strongly inhibited by all the flavonoids, with a total sucrose consumption that fall below 39% in the presence of flavonoids. However, the inhibition level of sucrose consumption is not correlated to the glucosylation level. To illustrate, mutant



H10 is inhibited up to 83% for sucrose consumption but was one of the best mutant for naringenin glucosylation (38%).

**Table 17. Sucrose consumption rates of the 23 selected mutants in the presence or absence of flavonoids. Standard deviations are given for the wild-type (WT). Mutations are also given for each mutant.**

Name	Mutations	sucrose consumption (%)						
		No Flavonoid	Quercetin	Luteolin	Morin	Naringenin	Apigenin	Chrysin
B5	W2135C-F2136I	54	28	52	34	5	20	44
B9	W2135S-F2136L	61	48	64	55	31	35	54
C12	W2135S-F2136L	80	48	84	53	23	40	59
H10	W2135G	53	31	50	40	9	27	48
A10	W2135V	100	86	99	86	54	84	93
E2	W2135C	83	60	80	53	29	51	77
E12	W2135L	99	86	99	77	66	74	93
F4	W2135L-F2136L	72	40	78	42	38	29	51
H1	W2135C-F2136N	33	40	34	38	20	16	37
H5	W2135F	98	83	97	86	46	76	88
G1	W2135F-F2136I	58	24	53	17	1	19	41
C2	W2135N-F2136Y	58	31	55	61	33	25	41
F3	W2135N-F2136H	32	24	27	28	9	21	35
B'2	F2163G	69	62	70	48	48	34	66
D10	W2135I-F2136Y	100	94	100	99	80	95	99
G'5	F2163L	99	93	100	92	64	79	95
F'9	A2162E-F2163L	99	78	100	87	48	64	84
C8	W2135N	44	36	31	26	2	19	38
D'1	F2163?-L2166?	100	90	100	84	59	57	89
H'5	F2163I-D2164E-L264I	100	100	100	100	96	99	100
B'5	L2166I	66	39	64	15	1	33	47
C'5	F2163H	100	16	39	18	5	24	35
E'10	F2163G-L2166I	66	34	64	35	13	36	55
C1	W2135I-F2136C	38	46	34	29	11	16	38
Inactive	E2248Q	0	0	0	0	0	0	0
WT		99 ± 2	90 ± 11	100 ± 1	88 ± 11	73 ± 18	68 ± 19	95 ± 6

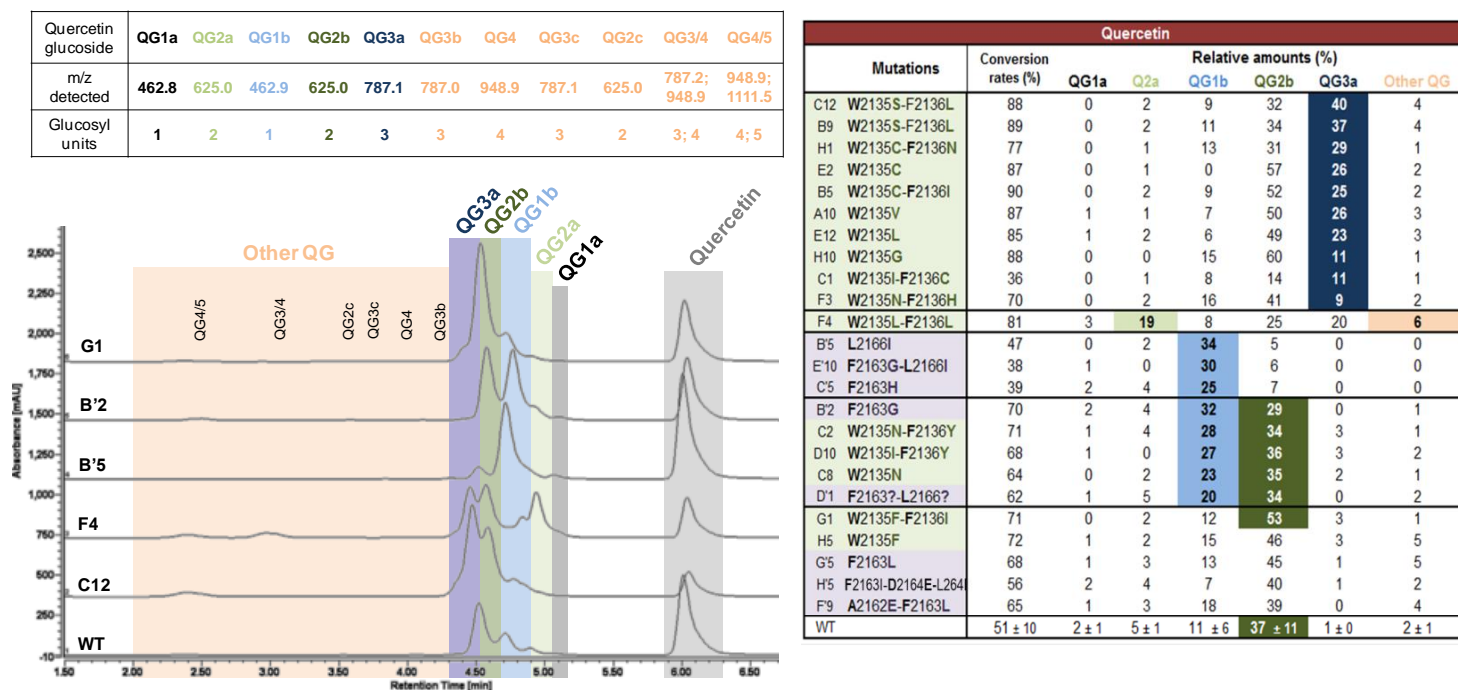
Equivalent levels of conversion were obtained with mutants A10 and H5 but these two enzymes were much less inhibited for sucrose consumption. This absence of correlation was also observed for most of the reactions and was previously reported for *NpAS* mutants (Malbert et al., 2014a; Thesis Chapter II). It could be explained by the large excess of sucrose used for the assays and the formation of different amounts of poly-glucosylated flavonoids.

### II.4.3. Glucosylation profile analysis

We analyzed the glucosylation profiles of each flavonoid by HPLC-C18-UV/MS separation and detection. Chrysin and apigenin were excluded of this analysis due to their very low conversion. The relative amounts of flavonoid glucosides were calculated by dividing the amount of the considered glucoside product by the amount of the total flavonoid content (glucosides and aglycon).



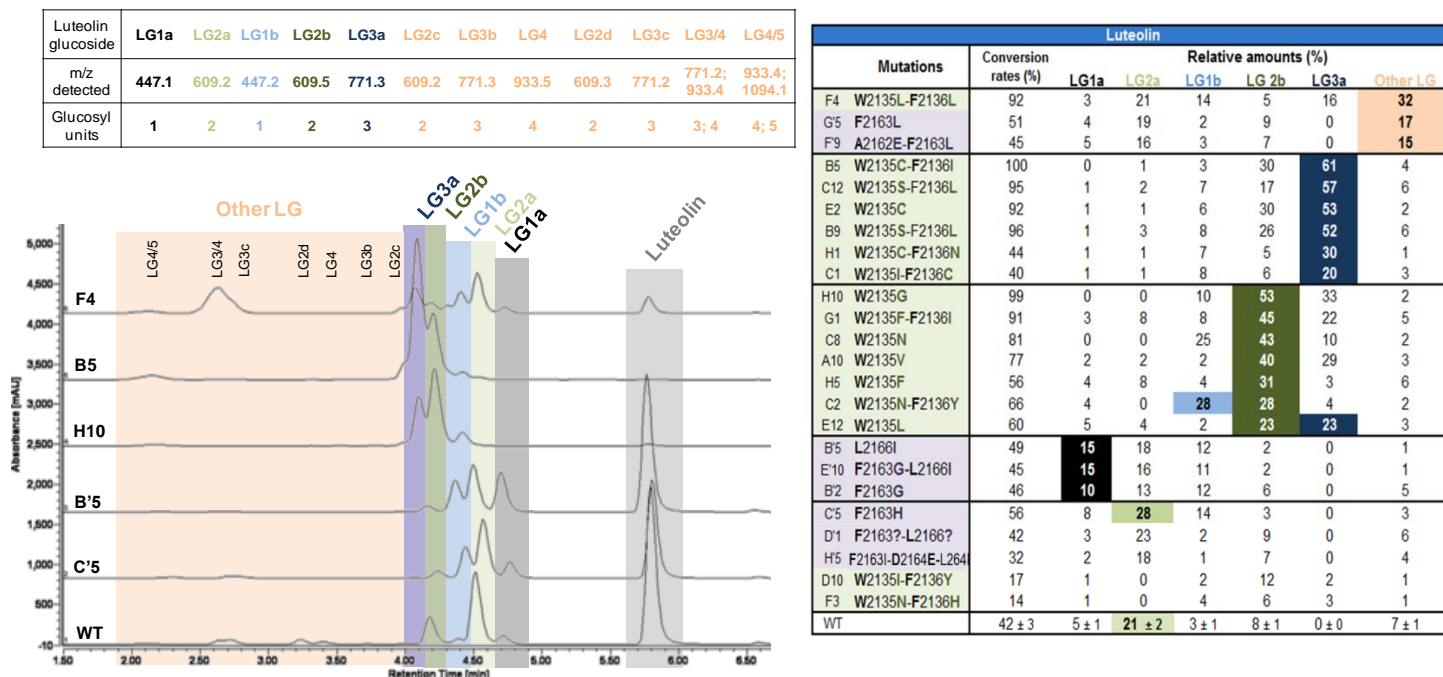
## II.4.3.a. Quercetin glucosylation profile (Figure 65)



**Figure 65. Quercetin glucosylation profiles obtained with the 23 selected mutants.** HPLC profiles of the best mutants for the production of each glucoside are reported. A table of MS data is supplied and explicit the number of glucosyl unit transferred for each product. The relative amount glucosylation products are reported in the table. Mutants are ranked relatively to the product synthesized in the highest amount (column highlighted by background color and bold numbers). The products of low retention times produced in lower amounts are grouped under the denomination "Other QG".

Of the 23 mutants, ten were shown to produced high amounts of quercetin triglucoside QG3a (up to 40%), which was not synthesized by the wild-type enzyme. These mutants all belong to the W2135-F2136 library. Interestingly, the 4 best mutants of this category possess a cysteine or a serine replacing the tryptophan at position 2135. Mutant F4 (W2135L-F2136L) significantly synthesized quercetin diglucoside QG2a, representing 19% of the converted quercetin, and equivalent amounts of QG2b and QG3a. Finally, some mutants produced QG1b monoglucoside in higher amount (up to 4-fold increased) than the wild-type.

## II.4.3.b. Luteolin glucosylation profile (Figure 66)



**Figure 66. Luteolin glucosylation profiles obtained with the 23 selected mutants.** HPLC profiles of the best mutants for the production of each glucoside are reported. A table of MS data is supplied and explicit the number of glucosyl unit transferred for each product. The relative amount glucosylation products are reported in the table. Mutants are ranked relatively to the product synthesized in the highest amount (column highlighted by background color and bold numbers). The products of low retention times produced in lower amounts are grouped under the denomination "Other LG".

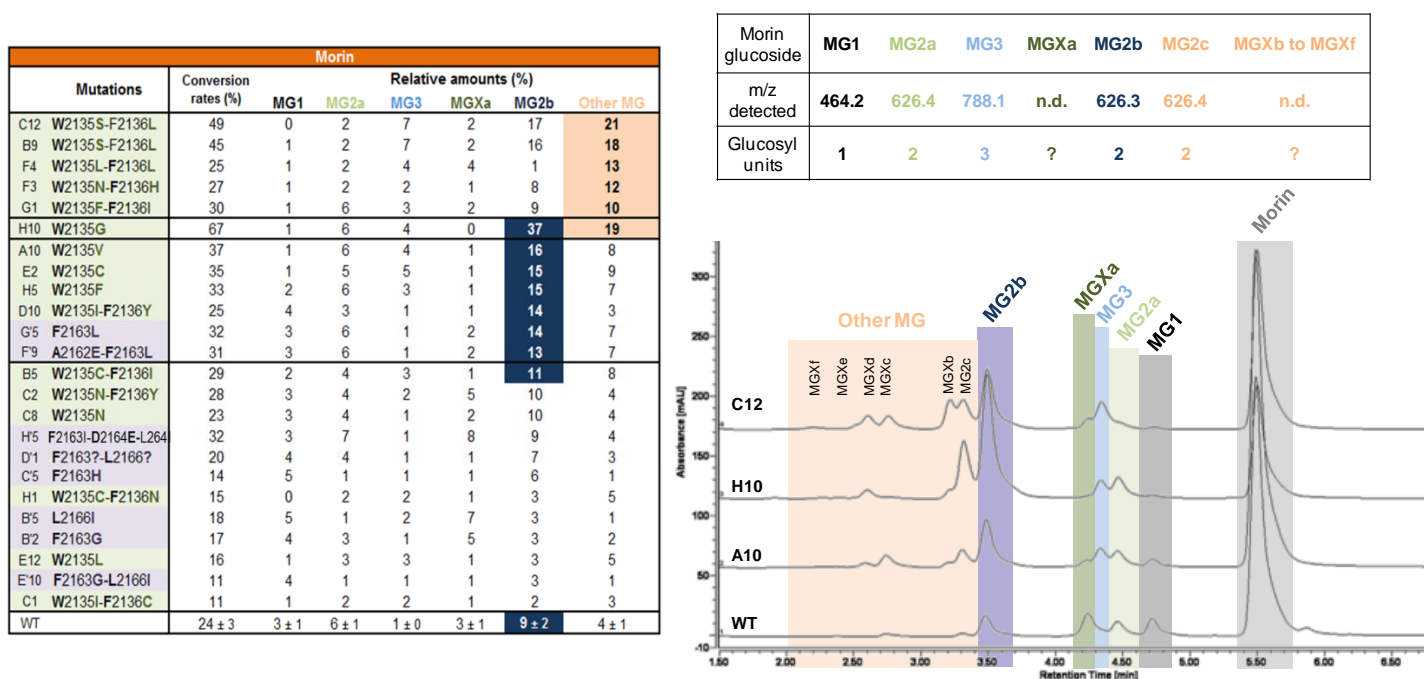
Two monoglucosides, four diglucosides, three triglucosides, two tetraglucosides and one pentaglucoside were detected. The presence of two monoglucosides with different retention time indicates that two different hydroxyl groups of luteolin can be glucosylated. By comparison with the results previously obtained with *NpAS* library (Malbert et al., 2014a; Thesis Chapter II), LG1b was eluted with the same retention time as that of luteolin-4'-O-monoglucoside suggesting that LG1b is glucosylated at 4'-hydroxyl position of cycle-B. The structure of LG1a is more elusive but due to the reactivity of the vicinal hydroxyl groups, we may suspect a glucosylation at the 3'-hydroxyl position of cycle-B. The presence of four diglucosides also shows that glucosylation may also occur on the A-ring and not only on the B-ring.

Three mutants (F4, G'5 and F'9) were improved up to 32% for the production of LG3c and LG3/4. A second series of five mutants (B5, C12 / B9, E2, H1, and C1) mainly synthesized triglucoside LG3a. All of them also produced mainly QG3a and exhibit a W2135C or S mutations, which clearly favor the triglucosylation of these flavonoids. LG3a and QG3a may

possess a similar glucosylation pattern. Again, this should be confirmed by structural characterization of the purified glucosides.

The third series of interesting mutants (H10, G1, C8, A10, H6, C2, and E12) mainly produced LG2b glucoside whereas the wild-type enzyme was more specific for the production LG2a. Three mutants (B5, E10, and B2) were more specific for monoglucoside synthesis and produced higher amounts of LG1a and LG1b than the wild-type enzyme. The mutations introduced in these variants did not favor of poly-glucosylation of luteolin.

### II.4.3.c. Morin glucosylation profile (Figure 67)

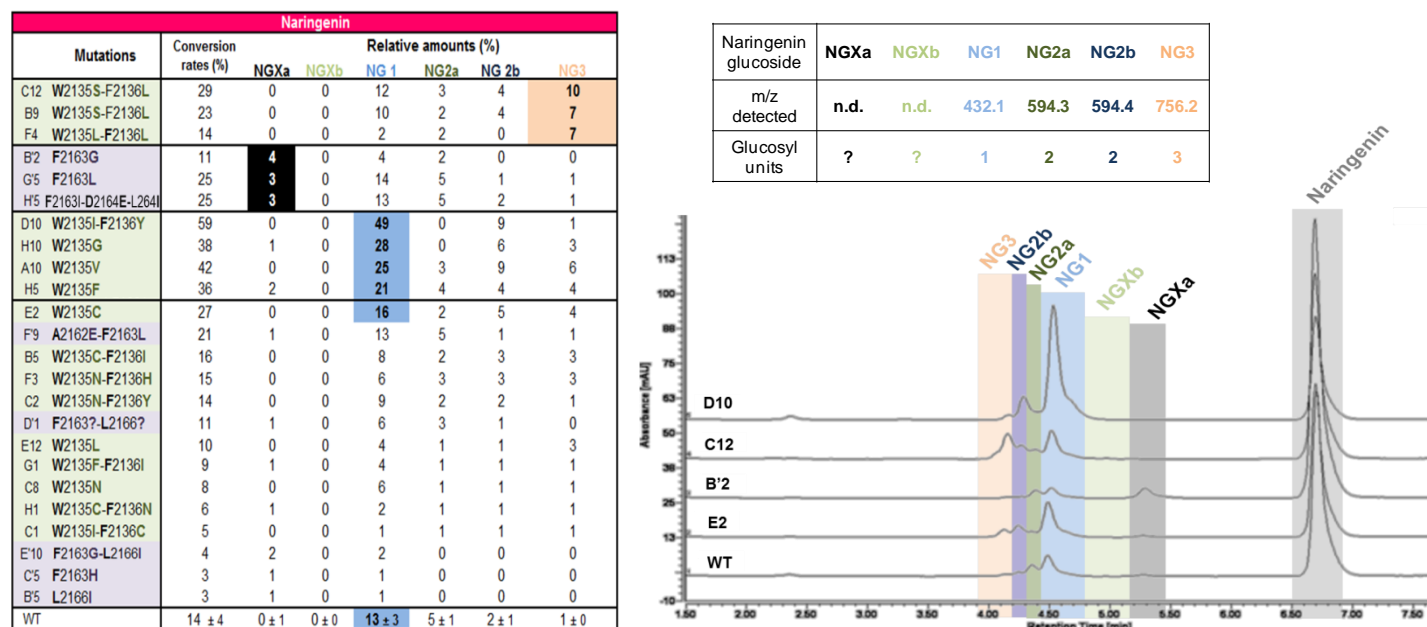


**Figure 67. Morin glucosylation profiles obtained with the 23 selected mutants.** HPLC profiles of the best mutants for the production of each glucoside are reported. A table of MS data is supplied and explicit the number of glucosyl unit transferred for each product. The relative amount glucosylation products are reported in the table. Mutants are ranked relatively to the product synthesized in the highest amount (column highlighted by background color and bold numbers). The products of low retention times produced in lower amounts are grouped under the denomination “Other MG”.

As previously seen, morin was converted less efficiently as luteolin or quercetin. Six of the eleven reaction products could not be identified by LC-MS. They were all referred as “MGX”, due to their undefined content of glucosyl units. Besides, one monoglucoside, two diglucosides and one triglucoside were identified. Five mutants (C12 / B9, F4, F3, G1, and H10) preferentially synthesized morin glucosides referred as MGX. Most of them are likely glucosides of higher

molecular mass showing low retention times. The best mutant for morin glucosylation is mutant H10 (W2135G) that synthesized four times more MG2b glucoside than the wild-type.

#### II.4.3.d. Naringenin glucosylation profile (Figure 68)



**Figure 68. Naringenin glucosylation profiles obtained with the 23 selected mutants.** HPLC profiles of the best mutants for the production of each glucoside are reported. A table of MS data is supplied and explicit the number of glucosyl unit transferred for each product. The relative amount glucosylation products are reported in the table. Mutants are ranked relatively to the product synthesized in the highest amount (column highlighted by background color and bold numbers).

Only seven reaction products were detected. As for morin, the mass of two minor products could not be determined. One naringenin monoglucoside, two diglucoside and one triglucoside were clearly detected. Naringenin was poorly glucosylated by the wild-type enzyme (14%) and most of the converted naringenin was monoglucosylated (NG1, 13%). Five mutants from library W2135-F2136 exhibited an improved production of NG1, up to the 49% achieved with D10 mutant (W2135I-F2136Y). Two different mutants (C12/B9 and F4) converted 10% of naringenin to a triglucoside form (versus 1% for the wild-type enzyme).

### III. Conclusion

In this work, enzyme engineering efforts were focused on the catalytic pocket of the  $\alpha$ -(1→2) branching sucrose. Four residues, not located in the catalytic amino acids sequence vicinity, were targeted and changed through saturation mutagenesis. From medium-size

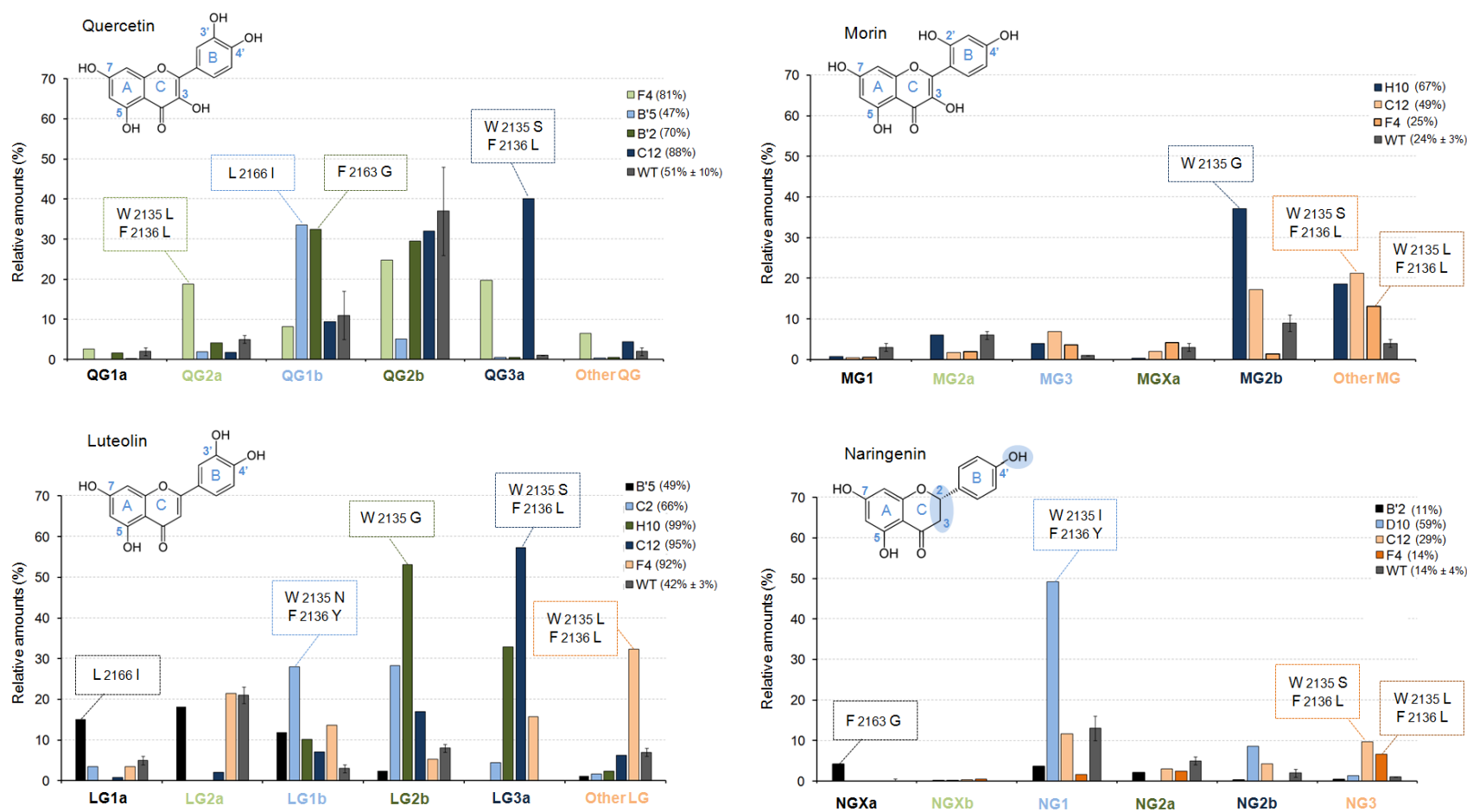
libraries, sucrose active mutants were rapidly isolated with a pH-based assay. Screening on quercetin glucosylation allowed further selection of a 23 variants set which were more efficient than the wild-type enzyme. This set of variants can be considered as neutral variants in regard to their activity on sucrose. From this very small set, a high ratio of enzymes showing promiscuous activities was sorted out.

The Figure 69 proposes a compilation of the diversity reached using this approach. A total of 10 variants were finally retained on the basis of their product profiles or catalytic performances. For each flavonoid, variants showing distinct specificities were identified. These results confirmed the interest of glucansucrase engineering to diversify flavonoid glucosides. This was previously exemplified with family 13 amylosucrase (paper 1 of this thesis) but is herein extended for the first time to GH70 enzymes.

Besides, the results also well illustrate the efficacy of the neutral drift approach to generate promiscuous activities. In particular, variants remarkably improved for morin and naringenin glucosylation were generated. First, this shows that contrary to what was previously reported, some enzymes from GH70 family are able to glucosylate flavonoids which do not possess vicinal hydroxyl groups on their B-ring. In addition, these enzymes can be significantly improved by protein engineering.

Finally, this small set of variants was tested with a larger panel of flavonoid acceptors. It provided some information on the probable effect of the mutations. However, to investigate these effects in more details, the isolation and structural characterization of the reaction products is now necessary. This will open the way to in depth structure-function relationship studies that should help to further improve the variants and adapt them to the glucosylation of recalcitrant flavonoids such as chrysin or apigenin.

The generated platform could be further extended by applying the same strategy, to target both other  $\alpha$ -(1→2) branching sucrose ( $\Delta N_{123}$ -GBD-CD2) regions and other glucansucrases. This should allow generating new flavonoids structures or achieving the glucosylation of recalcitrant flavonoids, such as chrysin, diosmetin or apigenin. In a same time, the mutants could be tested for the glucosylation of other flavonoids to enlarge the panel of acceptors and generated an extended library of glucosyl flavonoids. Thus, we have initiated the generation of a general flavonoid glucosylation platform that could become a powerful tool for valorization of flavonoids in therapeutics, cosmetic or food-processing industries.



**Figure 69. Compilation of the best hits obtained for the glucosylation of quercetin, luteolin, morin and naringenin.** The best relative amounts of each glucosides are presented together with the mutation harbored by the best producing mutant. Relative amounts of the glucosides produced by the wild-type  $\alpha$ -(1 $\rightarrow$ 2) branching sucrose are given (dark grey bars). The structure of each flavonoid structure is also presented.



## IV. Material and methods

### IV.1. Bacterial Strains, Plasmids, and Chemicals

Construction with the expression vector Gateway® Nova pET-53-DEST™ DNA (Novagen) containing *Δn123-gbd-cd2* gene (accession number: AJ430204) was used as DNA template for libraries construction (Malbert et al., 2014b; Thesis Chapter III). Fusion DNA-polymerase was purchased from Finnzymes (Espoo, Finland), and DpnI restriction enzyme from New England Biolabs (Beverly, MA, USA). Primers were synthesized by Eurogentec (Liège, Belgium). DNA extraction (QIAspin), gel extraction and purification (QIAquick) kits were purchased from Qiagen (Chatsworth, CA). *E. coli* TOP 10 electrocompetent cells (Invitrogen, Carlsbad, USA) were used as host for the plasmid library and *E. coli* BL21 Star DE3 (Life Technologies) for gene expression. DNA sequencing was performed by GATC Biotech (Mulhouse, France). All positive selected clones for quercetin glucosylation were sequenced using forward primer 5'-CCAACGAACACGAATGGGC-3' and the reverse one 5'-CTGTCATGATTGAATGCAAC-3'.

Ampicillin (Amp), lysozyme, and DNase I were purchased from Euromedex (Souffelweyersheim, France); Durapore® Membrane Filter 0.22μm GV from Merck Millipore (Darmstadt, Deutschland); bromocresol purple (BCP), sucrose, lactose, glucose, glycerol, dimethylsulfoxide, 2-(*N*-morpholino)ethanesulfonic acid (MES), sodium acetate and flavonoids from Sigma–Aldrich (Saint-Louis, MO, USA).

### IV.2. Production of recombinant $\alpha$ -(1→2) branching sucrase

Cultures in 5L-Erlenmeyer flask containing one liter of optimized auto-inducible ZYM5052 medium and recombinant  $\alpha$ -(1→2) branching sucrase ( $\Delta N_{123}$ -GBD-CD2) extraction were performed according to the previously described method in (Malbert et al., 2014b; Thesis Chapter III).

### IV.3. Sucrose enzymatic activity assay of the $\alpha$ -(1→2) branching sucrase

All assays were performed at 30 °C in 50 mM sodium acetate, pH 5.75. Enzyme activity was determined using the 3,5-dinitrosalicylic acid (DNS) as previously described in (Malbert et al., 2014b; Thesis Chapter III).

#### IV.4. Buffer and pH analyses

To analyze the  $\alpha$ -(1→2) branching sucrase stability at different pH, 1 mL of the soluble fraction of the enzymatic extract were added to 9 mL of MES or sodium acetate at 50 mM and ranging from pH 4 to 8. Solutions were stored for 24 h at 4°C. Thus final pH were checked and sucrose enzymatic activity DNS measurements carried out for each solution, in ultrapure water. Buffer capacity to neutralize  $\text{H}_3\text{O}^+$  analyzes were carried out using 10 mL of MES or sodium acetate buffer at 50 mM. Fixed volumes of a 1 M hydrochloric acid solution were gradually added to buffers and pH decreased followed through classic pH measurements.

#### IV.5. Molecular modeling and docking studies

The three-dimensional models of the  $\alpha$ -(1→2) branching sucrase in complex with sucrose or mono-glucosylated luteolin were generated using as template the coordinates of the crystallographic structures of  $\Delta\text{N}_{123}$ -GBD-CD2 (PDB accession code: 3TTQ; Brison et al., 2012) and *L. reuteri* GTF-180 in complex with sucrose (PDB accession code: 3HZ3; Vujičić-Žagar et al., 2010).

Molecules were initially manually docked in the active site of the enzymes. The four mono-glucosylated luteolin were constructed starting from the model of the  $\alpha$ -(1→2) branching sucrase in complex with sucrose, in which the glucosyl moiety of sucrose was attached to the luteolin moiety, generated using Corina 3D converter (Molecular Networks, Erlangen, Germany). They were positioned using as template the complex of *Streptococcus mutans* GTF-SI in complex with acarbose inhibitor (PDB accession code: 3AIC; Ito et al., 2011). Mutations were introduced using the Biopolymer module of Insight II software package (Accelrys, Saint Louis, USA).

The conformation of the mutated residue side chain was optimized by manually selecting a low-energy conformation from a side-chain rotamer library. Steric clashes (van der Waals overlap) and non-bonded interaction energies (Coulombic and Lennard–Jones) were evaluated for the different side-chain conformations. All energy minimization calculations were performed using the CFF91 force field and steepest descent algorithm (10,000 iterations) (Maple et al., 1988). Drawings were performed using PyMOL software (DeLano, 2002).



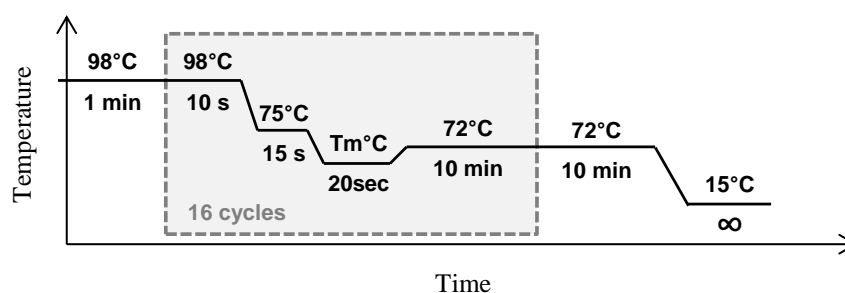
## IV.6. Molecular biology for $\alpha$ -(1→2) branching sucrase libraries creation

Mutagenesis experiments were carried out using pDest-53- *δn123-gbd-cd2* as a vector template. The primers used to achieve NDT mutations at each position are listed in Table 18.

**Table 18. Primers used to generate the monomutants for the libraries W2135-F2136 and F2163-L2166.** *NDT* and *NHA* codon indicates the bases which were used to obtain the replacement by the desired amino acids with N (A, T, C or G), D (not C) and H (not G).

Library	Primer direction	Nucleotide sequence	Tm°C
W2135-F2136	Forward	5'- GGT GGT GAT GCT <b>NDT NDT</b> CAA GGT GGT TAT CTG AAG -3'	60.5°C
	Reverse	5'- G ATA ACC ACC TTG <b>AHN AHN</b> AGC ATC ACC ACC ACC ATG -3'	61.6°C
F2163-L2166	Forward	5'- CCT GGT AAT GCA <b>NDT</b> GAT TTC <b>NDT</b> CTA GCC AAC GAC GTG G -3'	63.5°C
	Reverse	5'- GTC GTT GGC TAG <b>AHN</b> GAA ATC <b>AHN</b> TGC ATT ACC AGG TTG ACG -3'	63.4°C

Inverse PCR amplifications were carried out using the high fidelity Phusion DNA polymerase (Thermo Scientific) according to the following protocol: Template DNA, 0.12 ng.μL<sup>-1</sup>; forward and reverse primers, 200 nM; dNTP, 200μM each; Phusion DNA polymerase, 0.02 U/μL; in 50μL of HF Buffer 1X (thermo Scientific). The the PCR program diagram is provided below:



DNA was directly digested by DpnI endonuclease (20 U, 37°C, 2 h) to eliminate the methylated parental template and loaded on a 0,8% agarose, Tris, Acétate, EDTA (tris-base, 24,2 g.L<sup>-1</sup>; acetic acid, 5,7%; EDTA, 50 mM) gel for separation. The 8 kb bands, corresponding to pET53-*δn123-gbd-cd2* plasmids, were taken and purified using a Qiaquick Gel extraction kit (Qiagen®) and following the manufacturer recommendations. For each library, electrocompetent *E. coli* TOP10 were transformed with the amplified plasmids. The resulting clones were isolated by plating on LB agar (yeast extract, 5 g.L<sup>-1</sup>; Bacto Tryptone,

10 g.L<sup>-1</sup>; NaCl, 10 g.L<sup>-1</sup>; agar, 15 g.L<sup>-1</sup>) supplemented with 100 µg.mL<sup>-1</sup> ampicillin and contained in a 22 cm square sterile plates.

Clones were then scraped, using 5 mL of physiological saline water per plate and plasmids containing the mutated  $\delta n_{123}$ -*gbd-cd2* gene extracted. Purified DNA were quantified using a nanodrop (Thermo Scientific™) and freezed at -20°C.

## IV.7. $\alpha$ -(1→2) branching sucrase mutant libraries primary screening

On the first day, library plasmid stocks were transformed into electrocompetent *E. coli* BL21 Star. Recombinant clones were first plated for six hours on a Durapore® membrane, previously placed onto 22 cm square plates (Corning, USA) containing 200 mL of LB agar medium (200 mL) supplemented with ampicillin (100 mg.mL<sup>-1</sup>) to enable bacteria first growth.

The membranes were then transfered onto new 22 cm square plate containing the expression solid medium, adapted from (Malbert et al., 2014b; Thesis Chapter III) as follows: solid agar ZYM5052 optimized with lactose, 0.5% w/v; glucose, 0.01%, w/v; glycerol, 0.5%, w/v; ampicillin, 100 mg.mL<sup>-1</sup>; adjusted at pH 6.1.

The plates were then incubated at 23°C during 2 days. On the fourth day, membranes were transferred onto new 22 cm square plates containing the colorimetric screening medium: solid minimum M9 medium agar (Na<sub>2</sub>HPO<sub>4</sub>, 20 mM; KH<sub>2</sub>PO<sub>4</sub>, 40 mM; NH<sub>4</sub>Cl, 20 mM; NaCl 10 mM; CaCl<sub>2</sub>, 0.1 mM; pH 6.1) supplemented with lactose, 0.25% w/v; glycerol, 0.15%, w/v; thiamine, 5 µg.L<sup>-1</sup>; ampicillin, 100 mg.mL<sup>-1</sup> and sucrose (100 g.L<sup>-1</sup>). The medium was stained purple by addition of 50 mM 2-(*N*-morpholino)ethanesulfonic acid (MES) at pH 6.1, and bromocresol purple (BCP) indicator at 2.5 mg.L<sup>-1</sup>.

The plates were incubated at 30°C until clones expressing a sucrose active mutant turn yellow (12 to 24 h). Positive clones (yellow) were picked into two 96-wells microplates containing LB medium (200 mL) supplemented with glycerol (8%, w/v) and ampicillin (100 mg.mL<sup>-1</sup>). After 24 h of growth at 30°C, they were stored at -20°C and -80°C.

## IV.8. Semi-automated procedure for $\alpha$ -(1→2) branching sucrose mutant libraries secondary screening

The protocol used was previously established for *NpAS* library screening for rapid identification of variants improved for luteolin glucosylation, in a 96-wells microplates format (Malbert et al., 2014a; Thesis Chapter II). Here, flavonoid glucosylation reactions were adapted to the  $\alpha$ -(1→2) branching sucrose and carried out in the presence of sucrose (292 mM final), flavonoid (5 mM) and DMSO (10% final, v/v). The buffer for enzyme extractions and glucosylation reactions was sodium acetate at 20 mM and pH 5.75.

## IV.9. Analytical methods

Flavonoids and their glucosylated forms were separated and detected by LC-MS analysis with additional UV 340 nm detection, as described in Malbert et al. (2014a), Thesis Chapter II.

Sucrose consumption was monitored by high performance liquid chromatography with a refractometric detection as described in Malbert et al. (2014a), Thesis Chapter II.

## V. References

- Van Acker, S.A.B.E., de Groot, M.J., van den Berg, D.-J., Tromp, M.N.J.L., Donné-Op den Kelder, G., van der Vijgh, W.J.F., and Bast, A. (1996). A Quantum Chemical Explanation of the Antioxidant Activity of Flavonoids. *Chem. Res. Toxicol.* 9, 1305–1312.
- Amat, A., Sgamellotti, A., and Fantacci, S. (2008). Theoretical Study of the Structural and Electronic Properties of Luteolin and Apigenin Dyes. In *Computational Science and Its Applications – ICCSA 2008*, (Springer Berlin / Heidelberg), pp. 1141–1155.
- Bertrand, A., Morel, S., Lefoulon, F., Rolland, Y., Monsan, P., and Remaud-Simeon, M. (2006). *Leuconostoc mesenteroides* glucansucrase synthesis of flavonoid glucosides by acceptor reactions in aqueous-organic solvents. *Carbohydr. Res.* 341, 855–863.
- Brison, Y., Fabre, E., Moulis, C., Portais, J.-C., Monsan, P., and Remaud-Siméon, M. (2009). Synthesis of dextrans with controlled amounts of  $\alpha$ -1,2 linkages using the transglucosidase GBD-CD2. *Appl. Microbiol. Biotechnol.* 86, 545–554.
- Brison, Y., Pijning, T., Malbert, Y., Fabre, É., Mourey, L., Morel, S., Potocki-Véronèse, G., Monsan, P., Tranier, S., Remaud-Siméon, M., et al. (2012). Functional and structural characterization of  $\alpha$ -(1→2) branching sucrose derived from DSR-E glucansucrase. *J. Biol. Chem.* 287, 7915–7924.
- 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.
- Champion, E., Guérin, F., Moulis, C., Barbe, S., Tran, T.H., Morel, S., Descroix, K., Monsan, P., Mourey, L., and Mulard, L.A. (2012). Applying pairwise combinations of amino acid mutations for sorting out highly efficient glucosylation tools for chemo-enzymatic synthesis of bacterial oligosaccharides. *J. Am. Chem. Soc.* 134, 18677–18688.
- DeLano, W.L. (2002). The PyMOL Molecular Graphics System (San Carlos, CA: DeLano Scientific).
- Holms, H. (1996). Flux analysis and control of the central metabolic pathways in *Escherichia coli*. *FEMS Microbiol. Rev.* 19, 85–116.

- 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.
- Malbert, Y., Pizzut-Serin, S., Massou, S., Cambon, E., Laguerre, S., Monsan, P., Lefoulon, F., Morel, S., André, I., and Remaud-Simeon, M. (2014a). Extending the structural diversity of alpha-flavonoid glycosides with engineered glucansucrases. *ChemCatChem Accepted*.
- Malbert, Y., Vuillemin, M., Laguerre, S., Remaud-Siméon, M., and Moulis, C. (2014b). Optimizing the production of an  $\alpha$ -(1→ 2) branching sucrose in *Escherichia coli* using statistical design. *Appl. Microbiol. Biotechnol.* 1–12.
- Maple, J.R., Dinur, U., and Hagler, A.T. (1988). Derivation of force fields for molecular mechanics and dynamics from ab initio energy surfaces. *Proc. Natl. Acad. Sci.* **85**, 5350–5354.
- Meulenbeld, G.H., and Hartmans, S. (2000). Transglycosylation by *Streptococcus mutans* GS-5 glucosyltransferase-D: acceptor specificity and engineering of reaction conditions. *Biotechnol. Bioeng.* **70**, 363–369.
- Moon, Y.-H., Lee, J.-H., Jhon, D.-Y., Jun, W.-J., Kang, S.-S., Sim, J., Choi, H., Moon, J.-H., and Kim, D. (2007). Synthesis and characterization of novel quercetin-[alpha]-d-glucopyranosides using glucansucrase from *Leuconostoc mesenteroides*. *Enzyme Microb. Technol.* **40**, 1124–1129.
- Presser, K.A., Ratkowsky, D.A., and Ross, T. (1997). Modelling the growth rate of *Escherichia coli* as a function of pH and lactic acid concentration. *Appl. Environ. Microbiol.* **63**, 2355–2360.
- Reetz, M.T., Kahakeaw, D., and Lohmer, R. (2008). Addressing the Numbers Problem in Directed Evolution. *ChemBioChem* **9**, 1797–1804.
- Tucker, D.L., Tucker, N., and Conway, T. (2002). Gene Expression Profiling of the pH Response in *Escherichia coli*. *J. Bacteriol.* **184**, 6551–6558.
- Vujičić-Žagar, 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.* **107**, 21406–21411.



# **Conclusion & Prospects**



Along this thesis, we have seen that flavonoids and their glycosides are natural plant secondary metabolites exhibiting physico-chemical and biological properties of prime interest for food, therapeutic and cosmetic applications. Glycosylation reactions naturally occur in plants where they often have major impacts on the solubility, stability, as well as biological activities of the aglycon molecules. However, access to flavonoid glycosides from plant extracts remains limited. One way to facilitate the production of these compounds and open the route to a larger diversity is the development of *in vitro* enzyme-based processes of glycosylation.

This thesis work fits into this context and pursued several objectives including:

- Enzymatic glycodiversification of flavonoids.
- Generation of enzymatic tools dedicated to this function, using protein engineering.
- Demonstration of the benefits related to glycosylation.

To address these questions, we selected two sucrose-active enzymes. One amylosucrase from GH family 13 and one  $\alpha$ -(1 $\rightarrow$ 2) branching enzyme from GH family 70. Before concluding on the major results achieved, we briefly summarized hereafter the main motivation for the selection of these enzymes.

### **Motivation for selecting sucrose-active enzymes from GH 13 and 70 families**

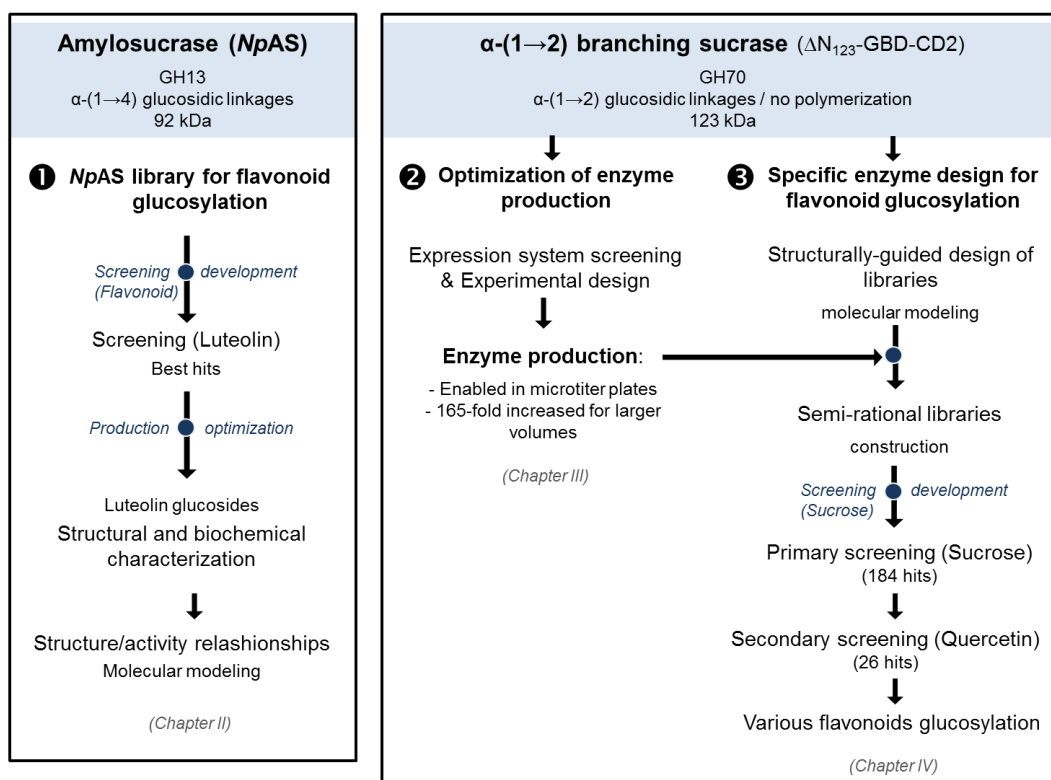
- These two enzymes use sucrose as glucosyl donor, a cheap and readily available substrate. In comparison, glycosyltransferases, which catalyze the formation of flavonoid glycosides in nature, use nucleotide-activated sugars as glucosyl donor substrates. If practical access to these molecules has been improved, their availability is yet not comparable to that of sucrose. This remains a major drawback for their *in vitro* use and such enzymes are more appropriate for *in vivo* applications.
- They were shown to be very efficient transglucosylases, exhibiting high specific activities and promiscuity toward various hydroxylated acceptors, including flavonoids.
- They specifically catalyze the formation of  $\alpha$ -glucosidic linkages and could potentially lead to original molecular structures differing from the natural ones which usually display  $\beta$ -linkages. This is appealing when diversity and innovation are searched.



- Enzyme engineering was successfully applied to tailor *NpAS* for the glucosylation of non-natural acceptors with high catalytic efficiency and controlled regiospecificity. Notably, none of these enzymes has previously been engineered to specifically improve flavonoid glucosylation.
- The 3D-structures of the two enzymes are available what enables their engineering through structurally-guided approaches, usually preferred for enzyme specificity modifications.

Taking into account these considerations and within the frame of a collaboration with “Les Laboratoires Servier”, we first assessed the potential of an amylosucrase (*NpAS*) library containing variants modified at their acceptor binding site for the glucosylation of luteolin (our model molecule). This library and the  $\alpha$ -(1→2) branching sucrase were then tested with other flavonoid acceptors. No hits were identified from the *NpAS* library, however, the preliminary tests confirmed the potential of the  $\alpha$ -(1→2) branching sucrase. They motivated the construction, from this template and with a structurally-guided strategy, of a platform of mutants improved for flavonoid glucosylation. The overall approach is depicted in the following workflow. We will conclude on the main achievements, discuss the advancements and provide the perspectives of this approach in the following parts.

### Engineering sucrose-active enzymes for flavonoid glucodiversification

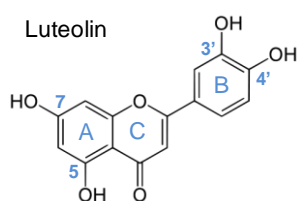


## Establishing efficient production systems for the selected glucansucrases

At the beginning of this thesis, *NpAS* was already produced in a soluble form in both microtiter plates and larger volumes. But the  $\alpha$ -(1→2) branching sucrase was mainly produced in inclusion bodies by *E. coli*, at the level of soluble enzyme reaching only 76 U.L<sup>-1</sup> of culture. Time-consuming and tedious steps of denaturation/renaturation of the aggregated enzyme were required to finally recover 2300 U.L<sup>-1</sup> of culture. Moreover, no activity in microplates could be recovered in the soluble fraction, preventing any attempt of enzyme evolution. To face this problem, implementing an efficient production process was thus a prerequisite.

As there is no “universal rule” to avoid the formation of inclusion bodies, experimental design and Response Surface Methodology (RSM) were used to optimize the production of the  $\alpha$ -(1→2) branching sucrase. The  $\Delta N_{123}$ -GBD-CD2 encoding gene was cloned into two expression vectors in fusion with 6xHis tag or Strep tag II encoding sequences at 5' or 3' ends of the gene. The plasmidic constructions and expressed in four *E. coli* strains. Response analyses allowed finding optimized conditions and producing up to 1260 U.L<sup>-1</sup> of culture of soluble enzyme with *E. coli* BL21 Star DE3 cells transformed with the pET-53 vector (B53).

RSM was applied to Box–Behnken designs to further optimize the expression conditions of the enzyme in an auto-inducible medium. Five factors were considered, *i.e.* culture duration, temperature and the concentrations of glycerol, lactose inducer and glucose repressor. The productions with the three tested expression systems were improved. The best response (production level of soluble enzyme) was obtained with *E. coli* BL21 Star DE3 cells transformed with the pET-55 vector (B55). However, production with B53 expression system led to better purification yields (48%) than that with B55 (11%). Using the predicted optimal conditions for B53, 3060 U. L<sup>-1</sup> of culture of soluble enzyme were produced in microtiter plates, enabling the implementation of screening procedures and the development of engineering approaches. Thus, we selected this B53 expression system for further engineering of the  $\alpha$ -(1→2) branching sucrase. Using the same conditions, more than 9,500 U.L<sup>-1</sup> of culture were produced in Erlenmeyer flask, which represents a 125-fold increase compared to the production levels previously reported, simplifies the production and purification process of the protein and opens new perspectives for its industrial applications.



Notably, a similar approach was conducted to optimize luteolin glucoside production using *NpAS* I228A mutant. A 4.5-fold increase in the glucosides yield (up to 13.7 mM) was obtained, enabling the easy production of several hundred mgs of products and allowing their structural characterization.

These results show that multi-factorial experimental design was reliable to define the best combination between various factors, reduce the number of experiments required to scan a large statistical space of conditions, and predict an optimal combination, of course dependant on the systems.

Another critical step in protein engineering is the development of efficient screening assays allowing accurate isolation of positive hits from variants libraries. Before our work, no miniaturized assays had been proposed to screen variants improved for flavonoid glucoside synthesis.

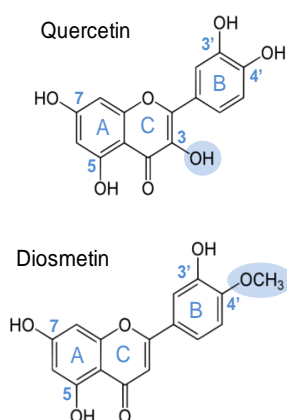
### Development of a screening procedure

A semi-automated and standardized screening protocol including cell growth in deep-well, cell lysis, activity assay (sucrose consumption) and acceptor reaction assay in the presence of luteolin was developed. The main difficulty we had to face was linked to the very poor solubility of luteolin in water (37  $\mu$ M). This clearly hampered the design of miniaturized, accurate and rapid screening protocols. This problem was partially circumvented by performing the acceptor reaction in heterogeneous medium containing 5 mM of luteolin and 3% (v/v) of DMSO. In these conditions, most of the luteolin was still insoluble at the beginning of the reaction but a phase transfer occurs from solid to liquid phase during the glycosylation reaction. Tested on the *NpAS* wild-type enzyme, the miniaturized assay was confirmed reproducible and yielded sufficient amounts of glucosides for LC-MS/UV<sub>340nm</sub> detection. Notably, this method was established with luteolin and easily adapted to screen seven other flavonoid acceptors. Thus, we developed a screening method which could be applied generally to evaluate the capacity of glucansucrases to glucosylate flavonoids.

### Assessing the engineering approach with *NpAS* mutant library promiscuity

Docking experiments of two luteolin monoglucosides (substituted at C3' or C4' position) in the *NpAS* acceptor binding site allowed to propose mutation targets in subsite +1 and +2 of the enzyme. A small library of *NpAS* single mutants targeting these positions (171 mutants), already available in the laboratory (LISBP) collection, was screened for luteolin glucosylation according to the established protocol. The results showed that an interesting

diversity was present in this small library. Indeed, 50 mutants were remarkably improved for luteolin conversion. 40 variants converted more than 20% of luteolin (1 mM) with conversions ranging from 20% to 66% depending on the mutant (7% for WT). Four positions were found critical for luteolin glucosylation (226, 228, 290 and 331). Among them, position 226 appeared clearly strategic as almost all mutants converted more than 25% luteolin. In addition, the library contained mutants with different regiospecificity, as reflected by the four different glucosylation profiles obtained.



These results confirmed the power of enzyme engineering. An interesting promiscuity was decrypted, which shows the potential of structurally-guided approaches to reshape the acceptor binding sites of enzymes, so that the designed libraries can efficiently glucosylate different interesting acceptors. However, this acceptor promiscuity remains limited. Indeed, even if the *NpAS* library was efficient for luteolin glucosylation, but only poor conversion rates, below 20%, were obtained for quercetin, and none of the mutants was able to glucosylate diosmetin.

## Engineering the $\alpha$ -(1 $\rightarrow$ 2) branching sucrose for flavonoid glucosylation

Engineering of the  $\alpha$ -(1 $\rightarrow$ 2) branching sucrose was based on molecular modeling and docking. Four luteolin glucosides, substituted at C3', C4', C5 or C7 position, were docked in the enzyme acceptor binding site. Four amino acids potentially involved in the recognition of flavonoids were selected, *i.e.* W2135, F2136, F2163 and L2166. We opted for a pairwise combinations strategy involving the construction of two medium size sub-libraries of several thousand clones. The first library was targeting on residues W2135 and F2136 while the second one focused on residues F2163 and L2166. The screening method conducted for the small *NpAS* library could not be used to directly screen these larger libraries. Indeed, it requires enzyme production and extraction steps, which are not compatible with a high throughput screening (HTS). Moreover, flavonoids are expensive and difficult to handle. Thus, the lack of HTS method, such as FACS, strongly limited the exploration of the diversity proposed by computational design analyzes and development of such HTS screening assays still remain incredibly challenging.

## Development of a rapid primary screen for activity on sucrose donor

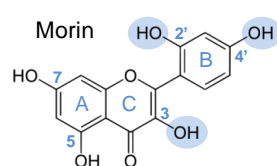
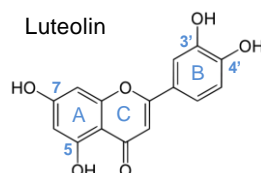
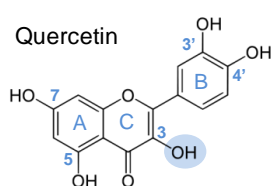
To overcome this limitation, a two stage screening procedure was developed. We decided to reduce the size of libraries by first isolating only the variants that had kept their ability to use sucrose as glucosyl donor. To achieve this, a rapid and efficient pH-based

colorimetric screening assay, adapted from that previously described by Champion et al. (2010), was implemented. The recombinant enzyme production on solid medium was optimized and we defined conditions for which acid production could be related to sucrose consumption by active  $\alpha$ -(1 $\rightarrow$ 2) branching sucrase. An appropriate pH indicator (bromocresol purple) allowing detection of active clones at pH values that were not detrimental for enzyme activity (optimal pH of 5.7) was selected. A total of 3,000 recombinant clones, ensuring a high coverage of the variant space, were screened. Less than 20% of the clones were active on sucrose, highlighting the relevance of this primary screening. 184 of the active clones were selected and confirmed to be active on sucrose. Library F2163-L2166 appeared negatively affected by mutations for sucrose activity, although, most of the library W2135-F2136 mutants showed a sucrase activity close to that of the wild-type.

### Secondary screening for flavonoid glucosylation

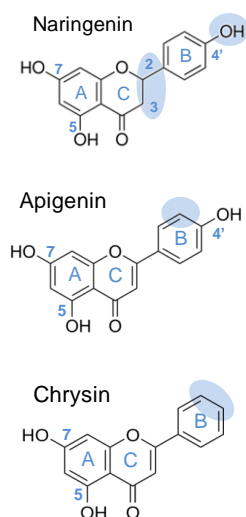
The screening procedure previously developed for NpAS was applied to screen the  $\alpha$ -(1 $\rightarrow$ 2) branching sucrase libraries for diosmetin and quercetin glucosylation. Of the 184 mutants, some mutants were found slightly improved for diosmetin glucosylation but the best conversion rates only reached 5%. Nevertheless, 80% and 41% of the mutants from W2135-F2136 and F2163-L2166 library, respectively, showed an improved conversion rate of quercetin, above 70%. The library that was the less affected for sucrase activity also provided the highest number of improved mutants for quercetin glucosylation.

From this secondary screening a set of 23 interesting mutants was sub-selected to further explore their promiscuity toward six structurally different flavonoids: quercetin, luteolin, morin, naringenin apigenin and chrysin.



Quercetin and luteolin were shown to be the best flavonoid acceptors, with 6 mutants exhibiting remarkable levels of conversion, above 90%. Some mutants also showed a modified regiospecificity toward luteolin, as revealed by the glycosylation profile diverging from that of the wild-type.

Morin and naringenin were successfully glucosylated up to 67% and 60%, respectively (versus 20% and 14% for the wild-type enzyme). The results obtained herein with the wild-type  $\alpha$ -(1 $\rightarrow$ 2) branching sucrase and the mutants demonstrated for the first time that the conjugation of the hydroxyl groups at C3' and C4' positions on the flavonol B-ring may favor flavonoid reactivity, but is not a pre-requisite for glucosylation with enzymes from GH70 family. These innovatives



results led to a patent application (Morel et al., 2014).

Diosmetin, apigenin and chrysin were poorly recognized by the mutants with conversion rates never exceeding 5%. When comparing apigenin and naringenin, the rigidity conferred by the C2,3-double bond was suspected to prevent a proper accommodation of the flavonoid in the active site. In addition, we may suggest that the glucosylation of the tested flavonoid probably targets preferentially the hydroxyl groups of the B-ring. Of course, these conclusions will have to be confirmed in depth by a detailed characterization of the flavonoid glucosides.

Thus, from this very small set of mutants, a high ratio of modified enzymes showing promiscuous activities was sorted out. A total of 10 variants was finally retained on the basis of their product profiles or catalytic performance. The generated platform could be further engineered and mutations recombined to generate catalysts able to achieve the glucosylation of the recalcitrant flavonoids. In a same time, these mutants could be further tested with other flavonoids to enlarge the panel of acceptors and generate a more diversified library of flavonoid glucosides. Thus, we have initiated the generation of a promiscuous platform for flavonoid glucosylation that clearly demonstrates the key contribution of enzyme engineering to this purpose.

The contribution of engineered enzymes for the glucosylation of non-natural exogenous acceptors was previously exemplified with family 13 amylosucrase (Champion et al., 2010; Daudé et al., 2014). This is confirmed here, and it was shown for the first time been for a GH70 family enzyme.

## Perspectives regarding the library of flavonoid glucosides

### *Structural characterization for structure-function relationships*

The 10 best mutants of the  $\alpha$ -(1 $\rightarrow$ 2) branching sucrose showed distinct specificities toward the tested flavonoids and many glucoside derivatives, never obtained before, were synthesized. Indeed, when considering quercetin, luteolin, morin and naringenin, at least 33 flavonoid glucosides were produced as revealed by LC-MS analyzes, some of them harboring up to 5 glucosyl units. For each flavonoid, several mono, di, or triglucosylated flavonoids were produced, raising the question of the glucosylation regioselectivity. For example, the identification of four luteolin diglucosides shows that glucosylation should also occur on the A-ring and not only on the B-ring (2 hydroxyl groups). Moreover, position

W2135 and F2136 appeared clearly well-targeted and highly strategic for flavonoid glucosylation.

But, to investigate the effect of the mutations in more details, the isolation and structural characterization of the reaction products should be investigated. This would open the way to in depth studies of structure-function relationships, comparable to those performed for the luteolin glucosides produced by *NpAS* mutants.

The three main luteolin glucosides produced using with I228A and R226Y *NpAS* mutants were characterized by NMR as luteolin-4'-O- $\alpha$ -D-glucopyranoside (L1), luteolin-4'-O- $\alpha$ -D-glucopyranosyl- $\alpha$ -(1 $\rightarrow$ 4)-D-glucopyranoside (L2) and luteolin-4'-O- $\alpha$ -D-glucopyranosyl- $\alpha$ -(1 $\rightarrow$ 4)-D-glucopyranosyl- $\alpha$ -(1 $\rightarrow$ 6)-D-glucopyranoside (L3). Thus, glucosyl units were preferentially and successively transferred to only one position of the luteolin with the unexpected formation of an  $\alpha$ -(1 $\rightarrow$ 6) linkage for the triglucoside. Molecular modeling studies showed that mutations at positions 226 and 331 facilitated the accommodation of luteolin monoglucoside in productive conformation. Elongation of luteolin diglucoside appeared to be achievable only through a more flexible  $\alpha$ -(1 $\rightarrow$ 6) glucosidic linkage. Indeed, it provides enough conformational flexibility to the nearly planar luteolin for its binding at subsites +3 and +4 through the establishment of beneficial van der Waals interactions with hydrophobic residues at position 226, such as for R226Y mutant. The characterization of these structures was essential to elucidate the mode of binding and get insight in the mutant structural determinants involved in their specificity. This is clearly essential to pursue engineering by taking advantage of the acquired knowledge and develop improved methods of computational design.

#### *Increasing the yield of productions*

We have seen that flavonoid lack of solubility limits the yield of production. DMSO was a good solubilizing agent but it may be detrimental for enzyme activities. Thus, different solubilizing agents could be tested. For example, addition of bovine serum albumin (BSA) to the reaction could be attempted to complex flavonoids (Fang et al., 2011; Xiao et al., 2011), enhance their solubility and possibly decrease their inhibitory effect, which was also observed in this work. The use of more stable immobilized enzymes or the removal of fructose that accumulates during the reaction and competes with the flavonoid could also help to increase production yields.

To improve the production of flavonoid glucosides, the transglucosylases could also be evolved to generate variants exhibiting an enhanced stability in organic solvents or at higher temperature. Getting access to such improved enzymes by directed molecular



evolution would be of great interest to serve as scaffold and facilitate both the access to the less produced flavonoid glucosides and the development of more efficient screening assays. To this purpose and in parallel to the work presented in the previous chapters, we generated a library of  $\alpha$ -(1 $\rightarrow$ 2) branching sucrase variants by error-prone PCR. This library was screened for sucrase activity using the pH-based colorimetric screening. The activity of the active mutants was further assayed in the presence of 30% DMSO or 37°C. Unfortunately, none of the variants showed significant improvements. In the frame of our work, we could not perform iterative cycles of evolution. Nevertheless, additional rounds of mutations and screening should be performed to really assess the relevance of such a strategy.

### The benefits of glucosylation

The water-solubility of the three characterized luteolin glucosides (L1, L2 and L3) was clearly promoted by glucosylation and increased gradually with the number of glucosyl units. L3 was shown to be soluble at concentrations of more than 647 mM (500 g kg<sup>-1</sup>), which represent a 17,700-fold increase, compared to luteolin aglycon (37  $\mu$ M). Such a level of solubility has never been reported for functionalized forms of flavones. The UV spectra of L1, L2 and L3 were also recorded and showed only little affected compared to that of luteolin. In contrast, the *in vitro* antioxidant properties (DPPH assays) of the three luteolin glucosides were totally lost showing that glycosylation can be seen as a protection by increasing their stability. We performed a limited characterization of the novel compounds. According to the numerous physicochemical properties of the flavonoids, other investigations should be conducted to explore the impact of the glucosylation on their, stability toward various environmental parameters, chemical reactivity, inhibitory effects, protein or DNA binding, browning resistance, etc. Increasing the number of novel structurally and biochemically characterized flavonoid glucosides could also be very useful to develop QSAR models and better predict the influence of a specific glycosylation pattern on the properties of the flavonoid glucosides. These data could also be very useful for enzyme engineering to predict the reactivity of a given flavonoid, and provide more reliable computational design approaches.

The *in vitro* characterization of flavonoids, is required to consider their application as active compounds in industrial formulations, but is not sufficient. Indeed, we suggested that L1, L2 and L3 *in vitro* loss of anti-antioxidant activity could help to preserve the native structure of the active molecule from rapid oxidation, delayed their metabolization and facilitate their absorption. Now, *in vivo* studies of L1, L2 and L3 properties should be carry out to further validate this hypothesis.



In conclusion, our work has demonstrated that a wide range of novel flavonoid  $\alpha$ -glucosides can be produced by engineered sucrose-active enzymes. Now, these compounds should be tested through pharmacological and formulation assays. They might be very attractive for many industrial purposes and especially in the pharmaceutical field. According to the promiscuity of glucansucrases and their mutants, the enzymatic biocatalysts developed here are not restricted to the glucosylation of flavonoids and could also be assayed for the glucosylation of other hydroxylated acceptors of interest, in order to further expand their applications.

## References

- 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.
- Daudé, D., André, I., Monsan, P., and Remaud-Siméon, M. (2014). Chapter 28. Successes in engineering glucansucrases to enhance glycodiversification. In *Carbohydrate Chemistry*, A. Pilar Rauter, T. Lindhorst, and Y. Queneau, eds. (Cambridge: Royal Society of Chemistry), 624–645.
- Morel, S., André, I., Brison, Y., Cambon, E., Malbert, Y., Pompon, D., Remaud-Siméon, M., and Urban, P. (2014). Projet GLYCOFLAV - Modification à façon de flavonoïdes pour créer des composés à forte valeur ajoutée. Patent in registration, DI 13-50, Toulouse White Biotechnology, INRA UMS 1337.
- Fang, R., Jing, H., Chai, Z., Zhao, G., Stoll, S., Ren, F., Liu, F., and Leng, X. (2011). Study of the physicochemical properties of the BSA: flavonoid nanoparticle. *Eur. Food Res. Technol.* 233, 275–283.
- Xiao, J.B., Huo, J.L., Yang, F., and Chen, X.Q. (2011). Noncovalent Interaction of Dietary Polyphenols with Bovine Hemoglobin in Vitro: Molecular Structure/Property–Affinity Relationship Aspects. *J. Agric. Food Chem.* 59, 8484–8490.

# **Artworks & Table Contents**



## I. Artwork content

Figure 1. Basic C6-C3-C6 skeleton of flavonoids.....	26
Figure 2. Flavonoid biosynthetic pathway from Lepiniec et al., 2006.....	27
Figure 3. Classes of flavonoids: chemical structures and examples of compounds for each group. ...	29
Figure 4. Scavenging of reactive oxygen species (R●) by flavonoid. ....	31
Figure 5. Main structural features for efficient radical scavenging, based on the quercetin structure. .	32
Figure 6. Metal chelation sites of the flavonol core. ....	32
Figure 7. Hypothesis of the links between the working mechanisms of flavonoids and their effects on disease. ....	37
Figure 8. Schematic representation of the pharmacokinetics of quercetin .....	41
Figure 9. Distribution of flavonoids in the splanchnic area .....	42
Figure 10. Examples of plant pigment formation through anthocyanin glycosylation .....	46
Figure 11. GT reaction mechanisms that lead to the inversion or retention of the anomeric stereocenter upon glycosyl transfer .....	47
Figure 12. Nomenclature system of the UGT superfamily .....	49
Figure 13. Diversity of CAZymes and auxiliary enzymes, according to the CAZy database .....	50
Figure 14. Structures of plant UDP glycosyltransferases.....	52
Figure 15. Flavonol glycosylation pathway in <i>Arabidopsis thaliana</i> .....	54
Figure 16. General concept of the one-pot system involving two enzymes for efficient synthesis of glucosides from flavonoids .....	56
Figure 17. Re-use of the nano-immobilized OleD GT for the synthesis of the glycosylated naringenin. ....	57
Figure 18. A glycosyltransferase gene toolbox and a nucleotide-sugar biosynthesis gene toolbox provide the basis for a general in vivo pathway engineering strategy to generate novel glycosylated natural products .....	59
Figure 19. Nucleotide sugar biosynthesis pathway and production of UDP-rhamnose in engineered <i>E. coli</i> .....	60
Figure 20. Specific activity of UGT85H2 and its mutants toward kaempferol and biochanin A .....	63
Figure 21. Schematic representation of parent UGT74F1*, UGT74F2*, and the generated chimeras .....	65
Figure 22. General Glycoside Hydrolases mechanisms .....	69
Figure 23. Synthetic and hydrolytic reactions catalyzed by Glycoside Hydrolases .....	70
Figure 24. Biocatalysis by Cel7B–E197S glycosynthase .....	75
Figure 25. GS categorization according to the structure of the polymer produced .....	77
Figure 26. Main reactions catalyzed by sucrose-utilizing glucansucrases.....	79
Figure 27. Schematic structure of DSR-E from <i>L. citeum</i> B1299 and GBD-CD2 truncated variant. ....	81
Figure 28. Structures of the $\alpha$ -(1→2) GOS produced by GBD-CD2 using sucrose as donor and pentasaccharide as acceptor.....	82
Figure 29. The two-step $\alpha$ -retaining-type mechanism of glucansucrases.....	84
Figure 30. Structural/sequence alignment of proposed GBD domains of GTF180- $\Delta$ N, GTFA- $\Delta$ N and DSR-E $\Delta$ N123-GBD-CD2 and GTFI-SI .....	86
Figure 31. Topology diagrams of GH13 and GH70 sucrose-utilizing transglucosylases .....	87

Figure 32. Schematic representation of the NpAS structure with labeling and color-coding of the five domains. ....	88
Figure 33. Three-dimensional structures of glucansucrases .....	90
Figure 34. View of the $\alpha$ -(1→2) branching sucrase domain organization .....	91
Figure 35. Domain V organization observed in the structures of GTF180- $\Delta$ N, GTFA- $\Delta$ N and DSR-E and $\Delta$ N123-GBD-CD2 .....	92
Figure 36. Schematic representation of the acceptor binding subsites .....	93
Figure 37. Schematic planar view of NpAS catalytic pocket, in complex with maltoheptaose .....	94
Figure 38. (A) Active site of the $\alpha$ -(1→2) branching sucrase superimposed with sucrose from GTF180-DN:sucrose complex structure (Brison, 2010). (B) View of subsites -1 and +1 of $\Delta$ N123-GBD-CD2, with sucrose from the GTF180- $\Delta$ N-sucrose complex superimposed.....	95
Figure 39. Kinetic parameters computed for the $\alpha$ -(1→2) transglucosylation of $\alpha$ -(1→6) dextrans catalyzed by the $\alpha$ -(1→2) branching sucrase.....	98
Figure 40. View of Amylosucrase active site in complex with maltoheptaose (PDB: 1MW0).....	101
Figure 41. Screening of the library of AS monomutants for their ability to synthesize the desired disaccharide.....	102
Figure 42. Catechin glucoside formation by <i>S. mutans</i> GS-5 GTF-D in the absence (♦) or presence of <i>P. pastoris</i> (●) or <i>S. cerevisiae</i> T2-3D (▲).....	103
Figure 43. Structures of the EGCG glucosides reported in <sup>a</sup> Moon et al., 2006b and <sup>b</sup> Moon et al., 2006a. Water solubility, conversion rates and purified amounts of each glucosides were also given.	105
Figure 44. Structures of the: (A) astragalin glucosides reported in Kim et al., 2012; (B) ampelopsin glucoside AMPLS-G1 reported in Woo et al., 2012.....	106
Figure 45. (A) Chemical structure and numbering of luteolin. (B) Side view of <i>N. polysaccharea</i> amylosucrase active site pocket.....	135
Figure 46. Results of NpAS single mutant library screening.....	138
Figure 47. Comparison between luteolin glucosides productions with wild-type NpAS and I228A mutant in screening and optimized reaction conditions. ....	139
Figure 48. Physico-chemical properties of the luteolin glucosides L1, L2 and L3. ....	141
Figure 49. View of amylosucrase active site in complex with maltoheptaose (cyan) (PDB: 1MW0) (A). Superimposition of docked L1 (green), L2 (magenta, and L3 (purple) onto maltoheptaose (cyan) (B). Docked L1 into I228A mutant (C), L2 into V331G mutant (D) and L3 into R226Y mutant (E). ....	144
Figure 50. Selection of the best expression systems.....	166
Figure 51. Surface plots (3D) of the activity recovered in the soluble fraction for A53 productions ...	169
Figure 52. SDS-PAGE and western blot analyses of $\Delta$ N <sub>123</sub> -GBD-CD2 enzyme expression in A53, B53 and B55. ....	171
Figure 53. SDS-PAGE of $\Delta$ N <sub>123</sub> -GBD-CD2 expressed in B53 expression system in soluble extract (S) and purified by 6xHis tag affinity chromatography .....	172
Figure 54. Activity recovered in the soluble fraction for production with <i>E. coli</i> BL21 Star DE3 cells transformed with the pET-53 and pET-55 vector (B53 and B55, respectively) and <i>E. coli</i> ArticExpress cells transformed with the pET-53 vector (A53), in microplates and Erlenmeyer. ....	193
Figure 55. View of $\alpha$ -(1→2) branching sucrase active site pocket showing the 2 amino acid mutagenesis targets of each library (light green and purple) and the adjacent ones (green).....	196
Figure 56. Solid-state pH-based colorimetric assay for screening of sucrose active mutants.....	197
Figure 57. (A) Residual activity of the $\alpha$ -(1→2) branching sucrase as a function of pH. Residual activities were calculated by comparison to activities determined before enzyme storage. (B) pH and BCP color variation versus H <sub>3</sub> O <sup>+</sup> addition. ....	199

Figure 58. Solid-state pH-based colorimetric screening methods validation with a mix of clones expressing wild-type or inactive $\alpha$ -(1→2) branching sucrase (mutant E2248Q).....	200
Figure 59. Distribution of the mutant activity (relative to wild-type) according to their activity on sucrose .....	201
Figure 60. Molecular structure of luteolin, diosmetin and quercetin.....	202
Figure 61. Results of the secondary screening of $\alpha$ -(1→2) branching sucrose mutant libraries, for their capacity to glucosylate quercetin. ....	204
Figure 62. $\alpha$ -(1→2) branching sucrase mutant sequences. ....	206
Figure 63. Molecular structure of the flavonoids tested as acceptor for the 23 selected mutants isolated from the preliminary screening assays. ....	207
Figure 64. Conversion rates for the glucosylation of the six selected flavonoids with the 23 mutants. ....	209
Figure 65. Quercetin glucosylation profiles obtained with the 23 selected mutants. ....	212
Figure 66. Luteolin glucosylation profiles obtained with the 23 selected mutants. ....	213
Figure 67. Morin glucosylation profiles obtained with the 23 selected mutants. ....	214
Figure 68. Naringenin glucosylation profiles obtained with the 23 selected mutants.....	215
Figure 69. Compilation of the best hits obtained for the glucosylation of quercetin, luteolin, morin and naringenin. ....	217
Figure 70. Structures des 3 produits de glucosylation majoritaires de la lutéoline par les variants de l'ASNp.....	248
Figure 71. Optimisation de l'activité de la saccharase de branchement $\alpha$ -(1→2) en fraction soluble pour les productions avec les tandems vecteur-hôte A53; B53, et B55 .....	250
Figure 72. Structure moléculaire des 6 flavonoïdes testés comme accepteurs.....	252
Figure 73. Compilation des meilleurs mutants de la .....	254
Figure S1. Method developed for screening <i>NpAS</i> single mutants for luteolin glucosylation. ....	154
Figure S2. Inhibition of <i>NpAS</i> activity on sucrose in the presence of DMSO. ....	154
Figure S3. Response Surface Methodology 2D-plots showing the effect of different reaction parameters .....	155
Figure S4. NMR spectra of luteolin glucosides L2 and L3. ....	156

## II. Table content

Table 1. UV-visible absorption of flavonoid subclasses, in methanol .....	34
Table 2. Examples of enzymes inhibited by flavonoids. ....	36
Table 3. Potential benefic effects of flavonoids in animal and human health.....	38
Table 4. Glucosyltransferases tested as whole-cell biocatalysts for quercetin glucosylation.....	58
Table 5. The established Glycoside Hydrolases clans of related families in the CAZy database .....	68
Table 6. Overview of flavonoids glycosylations catalyzed by Glycoside Hydrolases (except glucansucrases) .....	73
Table 7. Apparent kinetic values for sucrose consumption (ViS), sucrose hydrolysis(ViG) and polymerization (ViGx) .....	97
Table 8. Comparison of the apparent kinetic parameters determined for sucrose hydrolysis and $\alpha$ -(1→2) dextran branching activity of GBD-CD2 and $\Delta$ N123-GBD-CD2.....	98
Table 9. GH70 glucansucrases engineering studies for polymerization efficiency and/or specificity changes (reported in literature at present). ....	99

Table 10. Transglucosylation of diverse flavonoids using either dextranucrase B-512F or alternansucrase B-23192 and molar ratio of the various flavonoid glucosides .....	104
Table 11. Amounts and relative distribution of luteolin glucoside products synthesized with the best representative mutant of each product profiles shown in Figure 46. ....	136
Table 12. Response surface quadratic model analysis of $\Delta N_{123}$ -GBD-CD2 soluble activity level obtained with A53, B53 and B55 productions, after model simplification. ....	167
Table 13 Predicted and measured $\Delta N_{123}$ -GBD-CD2 soluble activity level with optimal conditions from RSM analyses, for A53, B53 and B55 expression systems. ....	170
Table 14 Purification of $\Delta N_{123}$ -GBD-CD2 expressed in pET-53 and pET-55 constructions from BL21 Star DE3 overexpression scale-up experiments .....	172
Table 15. Experimental level of factors affecting the enzyme $\Delta N_{123}$ -GBD-CD2 production for each host / vector combination. ....	177
Table 16. Comparison of the activity values and the quercetin conversion rates of the two screening assays. ..	208
Table 17. Sucrose consumption rates of the 23 selected mutants in the presence or absence of flavonoids. Standard deviations are given for the wild-type (WT). Mutations are also given for each mutant.....	211
Table 18. Primers used to generate the monomutants for the libraries W2135-F2136 and F2163-L2166. ....	220
Table S1. 171 <i>NpAS</i> single mutant library bacteria growth and activities on sucrose only.....	157
Table S2. $^{13}\text{C}$ and $^1\text{H}$ NMR data for luteolin, L1, L2 and L3.....	159
Table S3. RSM and results of the Box-Behnken design with the three variables studied.....	160
Table S4. <i>E. coli</i> strain genotypes used in this study for recombinant protein expression and for cloning experiments.....	183
Table S5. Experimental design (Box-bekhen design) for each expression system (host / vector).....	184
Table S6. Response surface quadratic model analysis of soluble enzyme production with ArcticExpress DE3 <i>E. coli</i> strain / pET-53 expression vector (A53), BL21 star DE3 <i>E. coli</i> strain / pET-53 expression vector (B53) and BL21 star DE3 <i>E. coli</i> strain / pET-53 expression vector (B55) productions .....	190

# Résumé en Français





Depuis quelques années, le développement durable est devenu une préoccupation majeure de notre société qui se répercute sur le secteur industriel et entraîne une mutation de ses procédés de production. Dans ce contexte socio-économique, la demande en nouvelles structures moléculaires biosourcées produites par des procédés industriels écoresponsables s'est donc fortement accrue. Cette tendance stimule les efforts de recherche universitaire oeuvrant pour le développement de voies alternatives de biotransformation d'extraits et de principes actifs d'origine végétale, tels que les flavonoïdes ou les sucres.

Les flavonoïdes sont des pigments naturels largement répandus dans le règne végétal où ils agissent, entre autres, en tant que filtres UV, antioxydants et/ou agents protecteurs contre les pathogènes. Leurs propriétés physico-chimiques et biologiques, de plus en plus étudiées, en font de bons candidats pour le développement de nouveaux principes actifs d'intérêts pharmaceutiques, cosmétiques ou agro-alimentaires. Cependant, leur très faible solubilité en milieux aqueux ou hydrophobes complique leur formulation et impacte leur biodisponibilité ce qui entrave l'accessibilité aux effets biologiques recherchés.

Dans les plantes, la glycosylation est un moyen efficace de générer des dérivés de flavonoïdes qui présentent, généralement, une solubilité dans l'eau et une stabilité supérieure à leurs aglycons référents. Ces flavonoïdes glycosylés sont fréquemment retrouvés dans la nature mais sont généralement produits en faibles quantités et leur extraction reste fastidieuse et peu efficace. La glycosylation chimique des flavonoïdes permet de générer de telles structures. Cependant, la complexité de la chimie des sucres, nécessitant de nombreuses étapes de protection / déprotection moléculaire, limite le développement de ce type de procédés.

La glycosylation enzymatique des flavonoïdes apparaît donc comme une approche très prometteuse qui suscite, depuis quelques décennies, un engouement considérable de la communauté scientifique. En effet, elle met à profit la stéréo et régiosélectivité des enzymes et permet la synthèse de nouveaux composés dans des conditions plus douces et plus respectueuses de l'environnement. Parmi les glyco-enzymes envisageables pour la biosynthèse de flavonoïdes glycosylés, les glucane-saccharases se présentent comme des biocatalyseurs attractifs. Ces enzymes particulières sont des transglucosylases très efficaces, qui utilisent le saccharose, une agro-ressource peu chère et largement disponible, en tant que donneur d'unités glucosyle. En outre, ces enzymes peuvent naturellement transférer l'unité glucosyle sur une large variété d'accepteurs hydroxylés, et notamment des flavonoïdes. Néanmoins, leur efficacité catalytique pour la glycosylation de nombreux flavonoïdes reste fortement limitée par leur faible promiscuité envers les molécules acceptrices d'unités glucosyles.

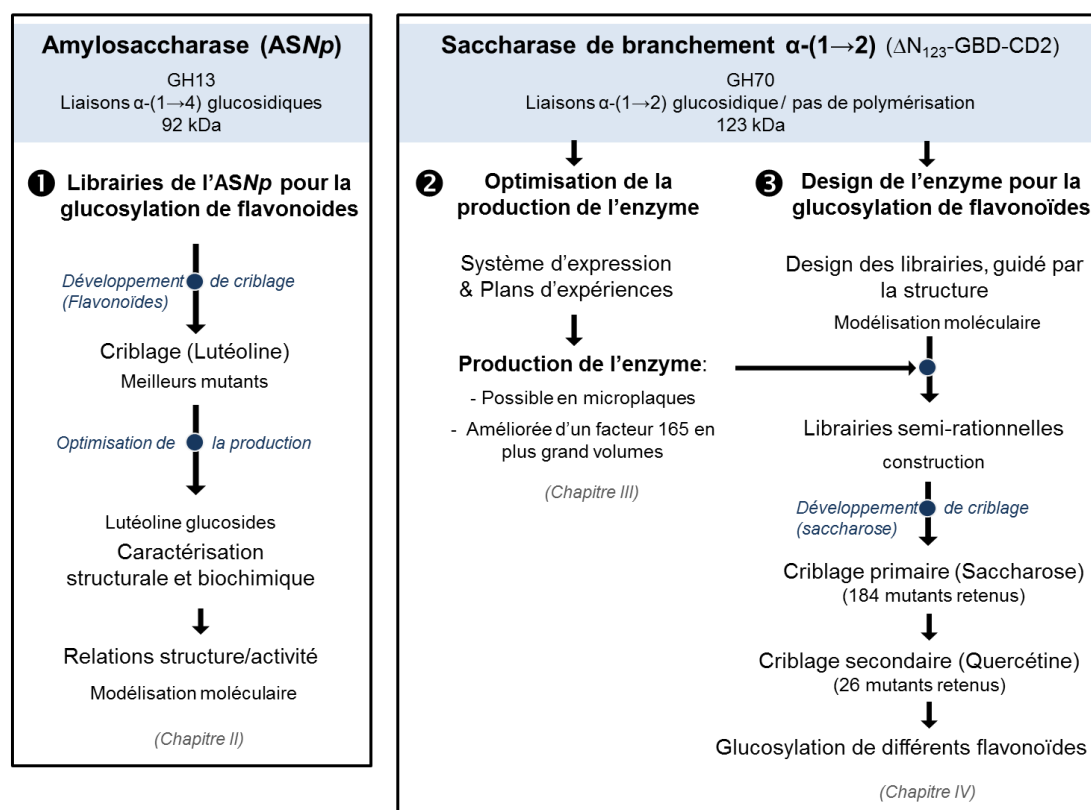
L'ingénierie des protéines, actuellement en plein essor, permet de remodeler les sites actifs des enzymes afin d'améliorer leur efficacité catalytique et/ou de changer leur spécificité réactionnelle. L'ingénierie enzymatique appliquée aux glucane-saccharases, pourrait donc permettre de surmonter leurs limitations, par remodelage de leur site actif, en vue de favoriser l'accommodation des flavonoïdes et d'augmenter leur réactivité catalytique. L'efficacité de telles stratégies a déjà été démontrée par des études visant à modifier la spécificité de ces enzymes et à améliorer la glycosylation d'accepteurs non-naturels. Cependant, aucun travail d'ingénierie des glucane-saccharases portant sur la glycosylation des flavonoïdes n'a été décrit dans la littérature.

Dans ce contexte, nos travaux visent à tirer parti des progrès de l'ingénierie des protéines afin de fournir des enzymes adaptées à la synthèse de nouveaux flavonoïdes-O- $\alpha$ -glucosylés, plus solubles en milieu aqueux, plus résistants à l'oxydation et présentant des biodisponibilités et bioactivités modifiées.

Ainsi, l'objectif principal de cette thèse s'attache à la construction, par ingénierie des protéines, d'outils enzymatiques efficaces, dédiés à la glucosylation des flavonoïdes ainsi qu'à la pré-caractérisation des produits de glucosylation obtenus. Pour mener à bien notre étude, deux transglucosylases recombinantes, biochimiquement et structuralement caractérisées, ont été sélectionnées :

- L'amylosaccharase de *Neisseria polysaccharea* (NpAS) qui appartient à la famille 13 des glycoside-hydrolases. Cette enzyme a été la première glucane-saccharase structurellement caractérisée (2001) et a déjà été remodelée par ingénierie pour différentes applications.
- La saccharase de branchement  $\alpha$ -(1 $\rightarrow$ 2), plus récemment caractérisée (2012), qui correspond à une forme tronquée de la dextrane-saccharase DSR- E de *Leuconostoc citreum* B1299, et appartient à la famille GH 70.

Le diagramme ci-après présente la stratégie et l'organisation générale des travaux menés au cours de cette thèse sur l'ingénierie enzymatique de ces deux enzymes actives sur saccharose, en vue de la glucodiversification des flavonoïdes.



## Crible et mise en oeuvre d'amylosaccharases (ASNp) modifiées pour étendre la diversité structurale des flavonoïdes α-glycosylés.

Dans la première partie de ces travaux de thèse, une librairie de mutants de l'ASNp a été sélectionnée, à partir de travaux de docking, et évaluée sur sa capacité à glucosyler la lutéoline, utilisée comme flavonoïde modèle. Cette petite librairie de l'ASNp, précédemment constituée par mutagenèse à saturation, est composée de 171 monomutants ciblés sur neuf positions du sous-site accepteur +1 de l'enzyme.

Dans un premier temps, nous avons développé un protocole de criblage standardisé afin de cribler cette librairie de mutants de l'ASNp sur sa capacité à glucosyler la lutéoline. Cette méthode de crible semi-automatisé comprend la croissance des cellules en microplaques, l'extraction de l'enzyme, l'analyse d'activité sur saccharose (donneur d'unité glucosyle) et la réaction d'accepteur.

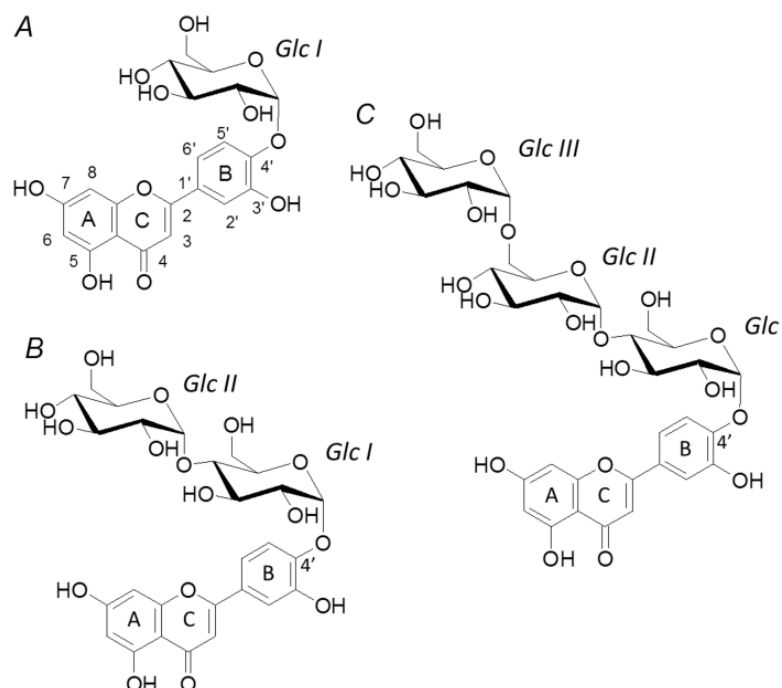
La très faible solubilité de la lutéoline dans l'eau (37 μM), comme pour de nombreux flavonoïdes, entravait fortement la conception d'un protocole de crible précis, rapide et miniaturisé. Nous avons démontré que ce problème pouvait être en grande partie résolu en effectuant la réaction d'accepteur en présence d'une petite quantité de solvant organique (3% de DMSO, v/v) et d'une concentration théorique de lutéoline fixée à 5 mM. Dans ces conditions, le milieu réactionnel reste hétérogène. Cependant, les essais se sont révélés

reproductibles, tout en permettant d'atteindre des concentrations de produits compatibles avec une analyse LC-MS/UV<sub>340nm</sub>.

Le crible de la banque d'ASNp a révélé que 29 % des mutants convertissaient significativement mieux la lutéoline que l'enzyme sauvage. Notamment, trois mutants ont permis la conversion de plus de 60% des 5mM de lutéoline, contre seulement 7% pour l'ASNp sauvage. En outre, pour l'ensemble de la librairie, quatre profils types de glucosylation ont pu être distingués.

Afin d'améliorer les rendements des trois lutéoline glucosylés majoritairement synthétisés (L1, L2 et L3) par les monomutants de la librairie d'ASNp, une Méthodologie en Surface de Réponse (MSR), appliquée à des plans expérimentaux de type Box-Benken, a été appliquée aux réactions de glucosylation catalysées par le monomutant I228A. Trois principaux paramètres réactionnels ont été analysés: la concentration en DMSO (v/v), la concentration en lutéoline, et le ratio molaire saccharose (donneur) / lutéoline (accepteur). Dans les conditions optimales prédites par l'analyse statistique, un gain de 10 % a été obtenu pour l'efficacité de conversion et 4.5 fois plus de lutéoline glucosylées ont pu être produites permettant d'atteindre 13,7 mM de produits.

Les dérivés glucosylés L1, L2 et L3 ont ensuite été purifiés afin de caractériser leur structure par RMN (Figure 70 A, B et C, respectivement).



**Figure 70. Structures des 3 produits de glucosylation majoritaires de la lutéoline par les variants de l'ASNp.** *lutéoline-4'-O-α-D-glucopyranoside* (A, L1), *lutéoline-4'-O-α-D-glucopyranosyl-α-(1→4)-D-glucopyranoside* (B, L2) et *lutéoline-4'-O-α-D-glucopyranosyl-α-(1→4)-D-glucopyranosyl-α-(1→6)-D-glucopyranoside* (C, L3).

Les formes di et tri-glucosylées de la lutéoline, respectivement L2 et L3, n'avaient jamais été décrites auparavant. Leurs structures ont révélé un transfert successif des unités glucosyle, préférentiellement sur la position C4' de la lutéoline, ainsi que la formation inattendue d'une liaison glucosidique de type  $\alpha$ -(1→6) pour la forme triglucosylée (L3).

Ces trois produits de glucosylation de la lutéoline ont montré une augmentation de leur solubilité en milieu aqueux, corrélée à leur degré de glucosylation et atteignant un facteur d'amélioration de 17 000 pour la forme triglucosylée. Nous avons également montré que la glucosylation protège ces produits de l'oxydation par des radicaux libres, ce qui pourrait s'avérer intéressant pour préserver leur intégrité fonctionnelle lors du transport dans l'organisme, jusqu'à ce qu'il soit déglucosylés.

Des analyses par modélisation moléculaire ont finalement été menées pour élucider le mode de liaison de la lutéoline et de ses dérivés glucosylés au sein du site actif de l'ASNp, et identifier les déterminants structuraux impliqués dans la spécificité des mutants d'intérêt. Elles ont mis en évidence que les mutations des résidus R226 et I331 et l'ASNp facilitent l'accommodation de la lutéoline monoglucosylée en conformation productive. L'élongation de la lutéoline diglucosylée n'apparaît possible que par la formation de la liaison  $\alpha$ -(1→6) qui confère suffisamment de flexibilité conformationnelle à la molécule pour son arrimage au niveau des sous-sites +3 et +4 de l'ASNp, via l'établissement d'interactions de type van der Waals avec les résidus hydrophobe de la position 226 (mutant R226Y).

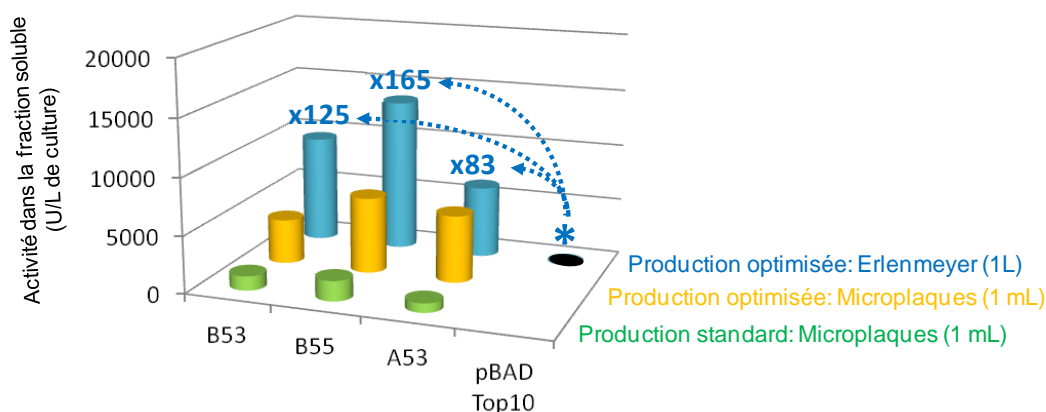
La promiscuité de cette banque de monomutants de l'ASNp a finalement été testée pour la glycosylation de deux autres flavones: la diosmétine et la quercétine. Toutefois, aucun des mutants ne s'est montré capable de glucosyler la diosmétine, et le meilleur mutant identifié pour la glucosylation de la quercétine ne présentait qu'un taux de conversion de 16%. Une autre transglucosylase, la saccharase de branchement  $\alpha$ -(1→2), capable de glucosyler la diosmétine (2%) et la quercétine (60%), a alors été sélectionnée en vue de travaux d'ingénierie enzymatique dédiés spécifiquement à la glucosylation des flavonoïdes.

### **Optimisation de l'expression soluble de la saccharase de branchement $\alpha$ -(1→2)**

L'une des premières exigences de l'ingénierie enzymatique est la production efficace de l'enzyme candidate sous forme soluble et active. Au début de cette thèse, la saccharase de branchement  $\alpha$ -(1→2) était principalement produite, chez *E. coli*, sous forme de corps d'inclusion avec seulement 76 U.L<sup>-1</sup><sub>de culture</sub> dans la fraction soluble. De plus, des étapes fastidieuses de dénaturation et renaturation de l'enzyme agrégée étaient nécessaires pour récupérer 2300 U.L<sup>-1</sup><sub>de culture</sub> d'enzyme soluble. Cette méthode de production s'avérait également incompatible avec une expression enzymatique en microplaques.

Nous avons donc établi une stratégie basée sur des plans d'expérience et une Méthodologie en Surface de Réponse (MSR) afin d'optimiser les conditions d'expression de l'enzyme en milieu auto-inductible. La première étape fut de tester plusieurs combinaisons plasmide-hôte, potentiellement plus favorables à une production soluble de l'enzyme. Le gène codant pour la saccharase de branchement  $\alpha$ -(1→2) a donc été cloné dans deux vecteurs d'expression, en fusion avec des étiquettes de purification 6xHis et Strep II, aux extrémités N- et C-terminales de l'enzyme. Ces constructions ont ensuite été transférées dans quatre souches différentes d'*E.coli*. Trois combinaisons hôte-plasmide ont été sélectionnées sur la base de la quantité d'enzyme soluble produite en microplaques. Les meilleurs résultats, en termes de niveau de production d'enzyme, ont atteint  $1260 \text{ U.L}^{-1}$  de culture pour le couple pET-53 et la souche *E. coli* BL21 DE3 Star (B53).

Trois plans expérimentaux de type Box-Behnken ont alors été établis pour analyser cinq facteurs importants de la production d'enzymes : la durée de la culture, la température et les concentrations en glycérol, lactose (inducteur) et glucose (répresseur). Trois blocs de quarante-cinq essais ont été réalisés en microplaques (1 mL de culture). Les modèles de régression ont été construits et une corrélation correcte a été obtenue avec les données expérimentales ( $R^2 > 94 \%$ ). D'importantes améliorations de la production ont ainsi été obtenues pour les 3 combinaisons construction-hôte testées, en microplaques comme en plus grands volumes (Figure 71). Les meilleures réponses (niveau de production de l'enzyme soluble) ont été obtenues avec le couple pET-55 et *E. coli* BL21 DE3 Star (B55, Figure 71).



**Figure 71. Optimisation de l'activité de la saccharase de branchement  $\alpha$ -(1→2) en fraction soluble pour les productions avec les tandems vecteur-hôte A53; B53, et B55. Les facteurs d'amélioration sont donnés pour les productions en erlenmeyer.**

En utilisant les conditions optimales prédites,  $5740 \text{ U.L}^{-1}$  de culture d'enzyme soluble ont été efficacement produites en microplaques, permettant ainsi le développement de procédures de crible et par conséquent d'approches d'ingénierie enzymatique. De plus, l'application de ces conditions sur de plus grands volumes de culture (1 L) a conduit à des

niveaux de production de plus de 12 000 U.L<sup>-1</sup><sub>de culture</sub> d'enzyme soluble. De tels niveaux de production représentent un gain spectaculaire, de l'ordre de 165 fois ceux préalablement décrit et inférieur à 80 U.L<sup>-1</sup><sub>de culture</sub>.

Afin de purifier l'enzyme et d'éliminer les formes dégradées co-produites (majoritairement dégradées en C-terminal de l'enzyme), la purification par chromatographie d'affinité avec l'étiquette Strep II (N-terminal) était requise. Malheureusement, la protéine pure n'a pu être isolée qu'avec un rendement de purification inférieur à 15%. C'est pourquoi, nous avons finalement opté pour le système de production B53 (pET-53 / *E. coli* BL21 Star) qui permettait d'atteindre des niveaux d'activité également très élevés, de l'ordre de 3060 U.L<sup>-1</sup><sub>de culture</sub> en microplaques et 9530 U.L<sup>-1</sup><sub>de culture</sub> en erlenmeyer (1L). En effet, la purification de l'enzyme produite est dans ce cas réalisée par chromatographie d'affinité avec l'étiquette poly-histidine (N-terminal) et conduit à des rendements de purification supérieurs à 40 %.

### **Design semi-rationnel de la saccharase de branchement $\alpha$ -(1→2) pour la diversification des flavonoïdes glucosylés**

Afin d'identifier les déterminants structuraux de la saccharase de branchement  $\alpha$ -(1→2), potentiellement impliqués dans la reconnaissance des flavonoïdes, quatre formes monoglucosylées de la lutéoline, substituées au niveau des hydroxyles en position 3', 4', 5 ou 7, ont été arrimées dans le site de liaison à l'accepteur par modélisation moléculaire. Quatre acides aminés potentiellement impliqués dans la reconnaissance des flavonoïdes ont ainsi été sélectionnés, soit W2135, F2136, F2163 et L2166. Nous avons opté pour une stratégie de mutagénèse en combinaisons par paires impliquant la construction de deux sous-librairies, W2135-F2136 et F2163-L2166, de tailles moyennes, contenant plusieurs milliers de clones. La méthode de criblage précédemment développée pour le tri des variants de l'ASNp n'étant pas adapté au criblage de telles librairies, un système de criblage en deux étapes a été développé.

### **Développement d'un criblage primaire sur saccharose**

Afin de réduire considérablement la taille des librairies à cribler avec les flavonoïdes, une procédure de criblage incluant une étape de pré-sélection des variants ayant gardé leur capacité à utiliser le saccharose comme donneur d'unités glucosyle a été établie. Elle consiste en un test de criblage rapide et efficace basé sur la détection colorimétrique d'un changement de pH, adapté de celui précédemment décrit par Champion *et al.* (2010).

La production de la saccharase de branchement  $\alpha$ -(1→2) recombinante, sur milieu solide, a été optimisée en ajustant les conditions optimales pour que la production d'acide soit spécifique de l'activité de l'enzyme sur saccharose. Un indicateur de pH approprié

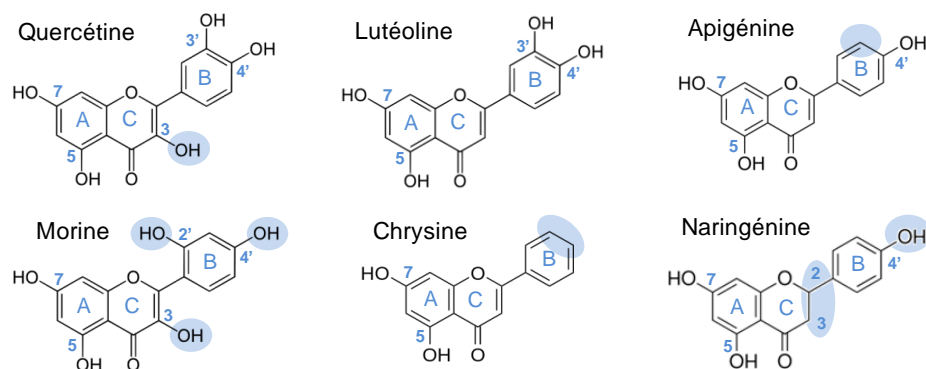


(pourpre de bromocrésol) permettant la détection des clones actifs à des valeurs de pH ne nuisant pas à l'activité enzymatique (pH optimal de 5,7) a été sélectionnée. Au total, 3000 clones recombinants ont été criblés pour assurer une couverture de l'espace de séquences, de l'ordre de 99%. Moins de 20% des clones se sont révélés actifs sur saccharose, soulignant la pertinence de ce criblage primaire. 92 clones actifs de chaque librairie ont été sélectionnés et leur activité de clivage du saccharose validée. Les mutations portées par la librairie F2163-L2166 ont affecté négativement l'activité de l'enzyme sur saccharose, tandis que la plupart des mutants de la librairie W2135-F2136 avaient gardé une activité proche de l'enzyme sauvage.

### Criblage secondaire pour la glucosylation des flavonoïdes

La méthode de criblage développée pour l'ASNp a été transposée à la saccharase de branchement  $\alpha$ -(1 $\rightarrow$ 2) pour permettre le crible des variants sur leur capacité de glucosylation de la diosmétine et de la quercétine. Parmi les 184 mutants testés, quelques uns présentaient une légère amélioration de leur capacité à glucosyler la diosmétine mais les taux de conversion obtenus ne dépassaient pas 5 % (WT, 2%). A contrario, une très forte proportion de variants améliorés et efficaces pour la glucosylation de la quercétine a été observée. En effet, 80 % et 41% des mutants provenant respectivement des librairies W2135-F2136 et F2163-L2166 ont converti à plus de 70% la quercétine (WT, 60%). La librairie W2135-F2136, moins affectée par les mutations au niveau de son activité sur saccharose, était également celle qui présentait le plus grand nombre de mutants améliorés pour la glucosylation de la quercétine.

A partir de ce crible secondaire, une série de 23 mutants intéressants a finalement été sélectionnés afin d'analyser leur promiscuité de glucosylation envers six flavonoïdes de structure différente: la quercétine, la lutéoline, la morine, la naringénine, l'apigénine et la chrysrine (Figure 72).



**Figure 72. Structure moléculaire des 6 flavonoïdes testés comme accepteurs.** Les différences structurales sont mises en évidence par des cercles bleus, par rapport à la lutéoline.

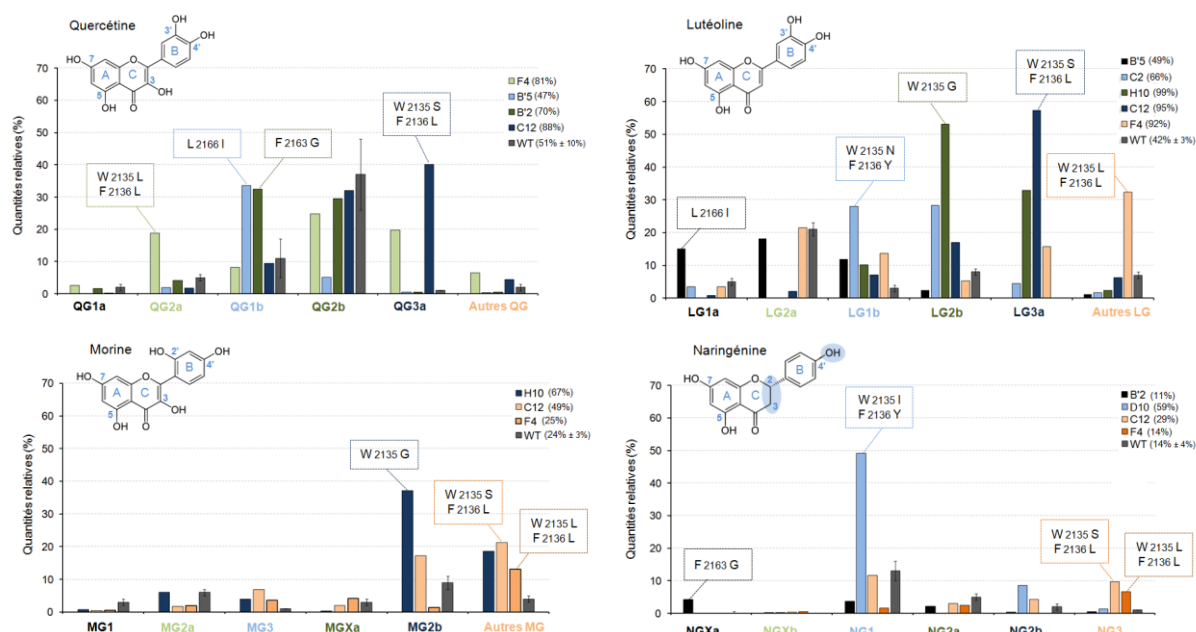
Globalement, la quercétine et la lutéoline se sont avérés être les meilleurs accepteurs, avec 6 mutants présentant des niveaux remarquables de conversion, de plus de 90 %. Certains mutants ont également montré une régiospécificité modifiée, tel que révélée par le profil de glycosylation divergeant de celui du obtenu avec l'enzyme sauvage.

La morine et la narigénine ont pu être glucosylées avec des taux de conversions atteignant 67 % et 60 %, respectivement (contre 20% et 14 % pour l'enzyme sauvage). Ces résultats démontrent, pour la première fois, que la conjugaison des groupements hydroxyle aux positions C3 'et C4' du cycle B (catechol) des flavonoïdes favorise leur réactivité, mais n'est pas un prérequis pour leur glucosylation avec des enzymes de la famille GH 70. Ces résultats très innovants ont conduit au dépôt d'un brevet (Morel *et al.*, 2014).

La diosmétine, l'apigénine et la chrysine se sont avérées être de mauvais accepteurs, avec des taux de conversion ne dépassant pas 5%. En comparant les résultats obtenus avec l'apigénine et la naringénine, il apparaît que la rigidité conférée au flavonoïde par la double liaison C2-C3 gêne très probablement son arrimage en conformation productive dans le site actif de l'enzyme. En outre, nous pouvons suspecter que la glycosylation des flavonoïdes testés vise préférentiellement les groupements hydroxyle du cycle B. Bien entendu, ces conclusions devront être validées par une caractérisation détaillée des structures des flavonoïdes glucosylés produits.

Au sein de cette plateforme de variants, nous avons obtenu un taux élevé d'enzymes présentant une promiscuité de glucosylation des flavonoïdes. Au terme du temps imparti pour ces travaux de thèse, un total de 10 variants a été retenu sur la base des profils de produits synthétisés et/ou de leurs performances catalytiques (Figure 73).

Ainsi, grâce à une approche d'ingénierie semi-rationnelle guidée par la structure de l'enzyme, nous avons développé une plate-forme d'enzymes de grand intérêt, actives sur saccharose et promiscuitaires pour la glucosylation des flavonoïdes. De plus il est important de noter que ces travaux correspondent au premier travail d'ingénierie enzymatique développé dans cet objectif et appliqué à la famille 70 des glycoside-hydrolases.



**Figure 73. Compilation des meilleurs mutants de la saccharase de branchement  $\alpha$ -(1→2) pour la glucosylation de la quercétine, de la lutéoline, de la morine et de la naringénine.** Les plus grandes quantités relatives de chaque glucoside sont présentées avec les mutations portées par le meilleur mutant producteur du dit glycoside. Celles de l'enzyme sauvage sont également données (barres grises). La structure de chaque flavonoïde est fournie avec l'histogramme correspondant. Les flavonoïdes glycosylés sont identifiés avec la première lettre correspondant à l'initiale du flavonoïde et le nombre d'unités glucosyle transférées.

En conclusion, nos travaux ont démontré qu'une large gamme de nouveaux dérivés  $\alpha$ -glucosylés de flavonoïdes pouvait-être produite à l'aide d'enzymes actives sur saccharose et modifiées à façon par ingénierie des protéines. Les nouvelles molécules produites doivent maintenant être testées par des essais pharmacologiques et de formulation afin d'évaluer leur propriétés et leur intérêt applicatif, en particulier dans les domaines pharmaceutique et cosmétique. Etant donnée la promiscuité, souvent observée, des glucansucrases et de leurs mutants pour la glycosylation d'accepteurs différents, les bibliothèques de variants construites dans le cadre de notre projet renferment probablement des enzymes performantes pour la glucosylation d'autres types d'accepteurs hydroxylés. Cette boîte à outils enzymatique pourrait donc s'avérer très utile pour étendre le champ d'applications des glucane-saccharases à de nombreux domaines, encore inexploités.

# Abbreviations



**ANOVA:** Analysis of Variance

**BET:** Ethidium Bromide

**BCP:** Bromocresol Purple

**CAZy:** Carbohydrate-Active enZYmes

**CD:** Catalytic Domain

**CGTase:** Cyclodextrine GlycosylTransferase

**DNS:** Dinitrosalicylic acid

**DP:** Degree of Polymerisation

**DSR:** Dextransucrase

**E.C.:** Enzyme Commission number

**EDTA:** Acide Ethylene Diamine Tetra Acétique

**GBD:** Glucan Binding Domain

**GBD-CD2:** Glucan Binding Domain – Catalytic Domain 2 ( $\alpha$ -(1→2) branching sucrase)

**GH:** Glycoside-Hydrolase

**GST:** Glutathion S Transferase

**GT:** GlycosylTransférase

**6xHis:** Six histidines tag

**HPLC:** High Performance Liquid Chromatography

**kDa :** kiloDalton

**LC-MS:** Liquid Chromatography – Mass Spectroscopy

**LB:** Lysogeny Broth

**MES:** 2-(*N*-morpholino)ethanesulfonic acid

**NMR:** Nuclear Magnetic Resonance

**NpAS:** Amylosucrase from *Neisseria polysacharea*

**OB:** Oligosaccharide Binding site

**PCR:** Polymerase Chain Reaction

**QSAR:** Quantitative Structure-Activity Relationships

**PDB:** Protein Data Bank

**RSM:** Response Surface Methodology

**SDS-PAGE:** Sodium Dodecyl Sulfate PolyAcrylamide Gel Electrophoresis

**Tris:** 2-amino-2-hydroxymethyl-1,3-propanediol

**U:** Unit of enzyme activity (correspond to  $\mu$ mol of fructose liberate per minute)

**UV:** Ultra Violet

**v/v:** volume per volume

**wt/v:** weight per volume

**WT:** Wild Type