

Université de Lille, Faculté des sciences et technologies

# **THESE DE DOCTORAT**

**Ecole doctorale Biologie Santé**

Présenté par

**Stanislas HELLE**

En vue de l'obtention du titre de Docteur de l'Université de Lille

**Etude de la stœchiométrie des protéines liées à l'amidon : vers des programmes novateurs pour la sélection et la production de nouvelles variétés de pommes de terre**

Présenté le 23 Novembre 2020 devant la commission d'examen :

<b>Président :</b>	<b>Yann GUERARDEL</b>
<b>Rapporteurs :</b>	<b>Séverine LE GAC</b> <b>Michel ZIVY</b>
<b>Examineurs :</b>	<b>Theresa FITZPATRICK</b> <b>Marc VAN DER MAAREL</b> <b>Bruno DESPREZ</b>
<b>Co-Directeurs de thèse :</b>	<b>Christophe D'HULST</b> <b>Christian ROLANDO</b>
<b>Encadrant de thèse :</b>	<b>Nicolas SZYDŁOWSKI</b>
<b>Encadrant au sein de l'entreprise :</b>	<b>Pierre DEVAUX</b>

# Résumé

La pomme de terre représente, en termes de quantité produite, la 5<sup>ème</sup> culture agricole, et la 4<sup>ème</sup> source d'amidon dans le monde. De plus, son marché est en pleine expansion parce qu'elle possède le plus haut rendement d'amidon produit par parcelle de terre. L'amidon est aujourd'hui la source principale de carbohydrates dans la nourriture humaine et du bétail. Il est aussi utilisé comme additif alimentaire, colle, bioplastique et biocarburant. Il s'agit d'un polymère de glucose lié en  $\alpha$ -1,4 avec quelques points de branchements en  $\alpha$ -1,6. Les recherches en matière de compréhension des protéines responsables de la synthèse et de la dégradation de l'amidon ont permis de mieux appréhender son métabolisme. Toutefois, les mécanismes d'association de ces protéines à l'amidon sont encore aujourd'hui mal compris. D'un point de vue fondamental, les objectifs de ma thèse étaient de mieux comprendre cette association entre le grain d'amidon et ces protéines. À l'aide d'analyses par spectrométrie de masse, nous avons, dans un premier temps, identifié l'ensemble des protéines liées à l'amidon de pomme de terre. Nous avons ensuite établi leur stœchiométrie par protéomique quantitative. Puis, nous avons comparé cette stœchiométrie entre différentes formes de grains et mis en évidence que cette stœchiométrie fluctuait en fonction de la taille des grains. Par la suite, nous avons étudié l'incorporation d'azote lourd  $^{15}\text{N}$  dans les séquences des protéines liées à l'amidon à différents temps de la tubérisation pour suivre de manière dynamique l'association des protéines au grain d'amidon durant la synthèse du tubercule. Ces résultats ont montré une plus forte incorporation de l'azote durant les 3 premières semaines de la tubérisation et une différence marquée de ce taux d'incorporation entre les amidons synthases, responsables de l'élongation des chaînes de glucanes de l'amidon, et les enzymes de branchement, responsables de la formation des points de branchements. Nous avons aussi étudié l'impact que pouvait avoir la protéine majoritaire du grain, la Granule-Bound Starch Synthase (GBSS), sur l'association des autres protéines au grain par l'analyse protéomique comparative des protéines liées au grain d'amidon de pommes de terre mutantes pour cette protéine. Enfin, nous avons miniaturisé les protocoles jusqu'à l'analyse d'un grain unique, évitant ainsi les biais amenés par l'hétérogénéité des grains à l'intérieur d'un même échantillon. Les avancées techniques de ce projet permettront, au travers de la collaboration de mon laboratoire d'accueil et de l'entreprise Florimond Desprez Veuve et Fils, d'identifier de nouveaux marqueurs de sélection pour contribuer au développement de programmes novateurs pour la sélection et la production de nouvelles variétés de pommes de terre.

# Abstract

Potato represents, in term of quantity produced, the 5th crop culture, and the 4th starch source in the world. Moreover, its market is booming, as a result of its highest starch/land parcel yield. Starch is nowadays the main source of carbohydrate in human and cattle food. It is also used as food additive, glue, bioplastic and biofuel. Researches that have been led in order to understand the functions of proteins responsible of starch synthesis and degradation allowed to understand better its metabolism. Nevertheless, the mechanisms of association between the granule and the proteins are still misunderstood. In a fundamental point of view, the objective is to better understand these association mechanisms. With the help of mass spectrometry, we first identified the whole proteome of potato starch. Then, established stoichiometry of proteins bound to potato starch by quantitative proteomics. Afterward, we have compared this stoichiometry between different granules morphologies, and have highlighted that protein stoichiometry inside the granule fluctuates with its size. Thereafter, we have studied the incorporation rate of exogen heavy nitrogen  $^{15}\text{N}$  in the sequence of starch bound proteins at different time, in order to follow dynamically the association of proteins to the granule. These results demonstrated a higher incorporation rate during the first three weeks of the tuberization, and a difference of incorporation between the starch synthases, which elongate the glucan chains, and the starch branching enzymes, which form the branching points. Moreover, we have investigated the impact of the major protein inside the starch granule, the Granule Bound Starch Synthase (GBSS), on the association of the other proteins by comparative proteomics between starch granules from mutant potatoes for this protein. Finally, protocols will be miniaturized up to the analysis of a single starch granules, to avoid bias due to inter-granule heterogeneity. In a research and development point of view, the collaboration with the breeding company Florimond Desprez Veuve et Fils will allow the development of programs of selection and production of new potato varieties, particularly by the identification of protein markers for a given phenotype.

# Remerciements

La réalisation de ces travaux a été permise par la collaboration de l'équipe de Glycobiologie végétale de l'Unité de Glycobiologie Structurale et Fonctionnelle (UGSF) et de l'Unité de Miniaturisation pour la Synthèse l'Analyse et la Protéomique (MSAP) à l'Université de Lille, Faculté des Sciences et Technologies, ainsi que l'entreprise Florimond Desprez, Veuve-et-Fils. Ces travaux n'ont été possible que grâce à l'aide de nombreuses personnes.

Je remercie en premier lieu le directeur de l'Unité UGSF, Yann GUERARDEL, pour avoir accepté de présider la commission d'examen de ma thèse, j'en profite pour remercier l'ensemble de l'équipe de l'UGSF, qui m'a apporté son aide et son soutien pendant l'intégralité de cette thèse.

Je remercie Séverine LE GAC, pour avoir accepté de faire partie de la commission d'examen de thèse en qualité de rapportrice, mais aussi pour avoir suivi ma progression depuis le début de ma thèse par le biais des comités de suivis individuels, pour m'avoir permis d'intégrer son équipe pendant mon Master, et du soutien apporté durant 4 années. Je remercie Michel ZIVY pour avoir accepté d'être rapporteur durant la commission d'examen de thèse.

I would like to thank Theresa FITZPATRICK for her presence as an examiner to my thesis defence. Ik dank Marc J.E.C. VAN DER MAAREL voor zijn deelname aan mijn proefschrift tot rapporteur. Je remercie Bruno DESPREZ, le Directeur de Florimond DESPREZ, pour m'avoir permis de réaliser cette thèse CIFRE, pour son soutien, et pour sa participation en tant qu'examineur lors de mon jury de thèse.

Je remercie mes deux directeurs de thèses et anciens directeurs d'unités du MSAP et de l'UGSF, Christian ROLANDO et Christophe D'HULST, qui suivent mes travaux depuis mon arrivé au MSAP durant mon Master. Merci de m'avoir permis d'intégrer votre équipe durant 4 ans, voir 5 ans dans le cas du MSAP, pour votre aide apportée sur les différents projets de recherches, pour le temps que vous avez consacré à la correction des différents manuscrits et pour votre aide dans ma recherche de post-doc. Un grand merci pour m'avoir plusieurs fois donné l'occasion de présenter mes travaux lors de congrès internationaux. Enfin, merci pour votre sympathie et votre soutien aussi bien au travail qu'en dehors du labo.

Un grand, grand merci à Nicolas SZYDLOWSKI, mon superviseur, sans qui ces travaux n'auraient pas été possible. Merci pour toutes les formations que tu m'as apportées, aussi bien à la paillasse que dans la rédaction des manuscrits. Merci pour ta compréhension et ta compassion qui m'ont permis de garder le moral dans les coups dur. Pour ton humour, en partie hérité de Christophe, même si j'ai parfois du mal à le saisir. Merci pour tous les bons moments

passés aussi bien au labo qu'en dehors. Ce fut pour moi un plaisir de travailler avec toi durant ces 4 années.

Merci à Pierre DEVAUX, mon encadrant de thèse au sein de l'entreprise Florimond DESPREZ, pour son aide, son soutien, et pour les variétés de pommes de terre fournies. J'en profite pour remercier les personnels de Florimond DESPREZ qui m'ont aidé aussi bien sur le plan recherche qu'administratif durant mes 3 années de thèse.

Un gros merci à tous les membres, passés et présents, de la « Team Patate ». Merci à Jérémy VERBEKE qui m'a accompagné durant mes premiers mois de Master 2. Merci à Maud FACON et François SEVESTRE, pour leur aide et leur soutien depuis le début de ce projet ainsi que pour tous les bons moments passés en dehors du labo. Merci à Mara BERTI, Océane GEORGE, Claudine RAZANABOAHIRANA et Camille LOCQUET qui ont récemment rejoint l'équipe et qui m'ont aidé et soutenu durant ma dernière année de thèse.

Merci aux stagiaires que j'ai eu l'occasion d'encadrer, Walaa DERBASS, et Benoît POUILLAIN, qui m'ont beaucoup aidé à avancer sur les différents projets. Bon courage à tous les deux pour la suite !

Un remerciement tout particulier aux membres de la plateforme de protéomique, Stéphanie FLAMENT et Fabrice BRAY, qui m'ont formé à la spectrométrie de masse dès mon arrivé au MSAP en 2015, et qui n'ont cessé de m'épauler durant toutes ces années. Merci aussi à Amandine PRUVOST pour m'avoir aidé plusieurs fois aussi bien au niveau de la masse que de l'électrophorèse capillaire, et pour son soutien.

Merci aussi aux autres membres du MSAP, passés et présents, Laetitia CHAUSSEZ, Françoise PRETE, Karine WARME, Caroline TOKARSKI, Olga PLECHAKOVA, Ahmed MAZZAH, Maël PENHOAT, Christophe PENVERNE ; ainsi qu'aux doctorants et post-doc, Violaine PIOTROWSKI, Mélanie ROSEAU, Wissal ADHAMI, Marie YAMMINE, Ranin DABBOUSY, Katernia BRETOSH, Marc HAELGELIN, Ziad MAHMOUD, Sergi MANSOUR, Julien MOLINA, Wael BARAKAT et Mathieu BEAUCAMPS pour leur aide et leur soutien.

Je remercie les membres de l'équipe de Glycobiologie végétales, Coralie BOMPARD, Adeline COURSEAU, Fabrice WATTEBLED, David DAUVILLÉE et Xavier ROUSSEL qui m'ont tous apporté leur aide à un moment donné de ma thèse, que ce soit à la paillasse, dans l'élaboration d'une manip' ou dans ma recherche de post-doc. Un grand Merci aussi aux anciens

doctorants Adeline WYCHOWSKI, Camille VANDROMME, et Sylvain LAURENT, pour leur aide, leur soutien, et les bons moments passés ensemble au travail et en dehors du labo. Merci aussi à Corentin SPRIET pour son soutien et les bons moments passés ensemble, ainsi que pour son aide dans le projet d'art de mon cousin.

Merci aux membres de l'équipe de Génétique microbienne Steven BALL, Christophe COLLEONI, Ugo CENCI, Malika CHABI, Derifa KADOUCHE, ainsi que les doctorants Marie LELEU et Mathieu COLPAERT pour leur aide et leur soutien.

Merci aux membres du secrétariat et du service financier de l'UGSF pour leur aide apporté durant mes démarches administratives, et plus particulièrement à Nadège VEREECKE qui m'a été plusieurs fois d'un grand secours.

Merci à Jean-Luc PUTAUX et au CERMAV de Grenoble pour leur aide dans les différentes expériences de structures réalisées sur les grains d'amidons de pomme de terre, ainsi qu'à Florian VEILLET, Jean-Eric et Laura CHAUVIN et aux membres de l'INRA Agrocampus Ouest pour leur collaboration et pour nous avoir fourni les tubercules de pommes de terre des lignées mutantes.

I would like to thank Samuel ZEEMAN, who follow my work during my first CSI. Muchas Gracias también a Ángel MERIDA por participar en mi segundo CSI.

Merci à toute la bande de loustics de médecine pour leur soutien indéfectible et les bonnes soirées passés ensemble : Marie-Colombe SCHMITTBIEL, Marie-Alix, Léa VERRIER, Sandra LE LANN, Valentin PALIE, Marco RODRIGUEZ, Pierre SOULA, Corentin MEERSSEMAN, Jean-Paul NIGUET, Marc CHANCELADES et Phillippe HERPELINCK. Merci à mes amis de fac : Emilie COCARDE, Julie GUYONNAUD, Camille HACHE, Alice CAPUZ, Anne-Sophie ROY, Cyrille CAUDROY, Quentin MAYOLLE, Joris TANIÈRE, et tous ceux que j'ai oublié.

Merci à l'association Bioaddoct pour les soirées organisées durant ces 3 ans de thèse.

Merci à Éric SAGNIEZ pour la conception de mon costume de thèse.

Enfin, Un grand merci à ma famille, qui m'a soutenu durant toutes ces années, plus particulièrement à ma famille proche : ma mère Douceline HELLE, mon père, Bruno HELLE, ma petite sœur Bettina HELLE, et mon petit chat, Lollipop.

# Abréviations

AA : Acide aminé

ABC : Ammonium bicarbonate

ADN : acide désoxyribonucléique

ADNc : Acide désoxyribonucléique complémentaire

ADP : adénosine diphosphate

ADP-Glc : ADP-Glucose

AGC target : Automatic Gain Control target

AGPase : ADP-Glucose pyrophosphorylase

amol : attomole

AMP : adénosine monophosphate

AMPK : AMP-Kinase

AMY :  $\alpha$ -amylase

ANOVA : Analyse de variance

AP2 : APETALA2

APTS : acide 8- aminopyrene-1,3,6-trisulfonique

ARNm : acide ribonucléique messenger

*At* : *Arabidopsis thaliana*

ATP : adénosine triphosphate

BAM :  $\beta$ -amylase

BCA : Acide bicinchoninique

BE : Branching enzyme

BES1\_N: *Brassinosteroids insensitive 1*- Ethyl methanesulfonate -Suppressor 1, N-ter

BT-1 : protéine Brittle-1



CAZy : Carbohydrate-Active enZYmes

CBM : Carbohydrate binding module

CBQCA : 3-(4-carboxybenzoyl) quinoline-2-carboxaldehyde

CC : Coil-coiled

CERMAV : Centre de recherches sur les macromolécules végétales

CNRS : Centre National de la Recherche Scientifique

Col-0 : fond génétique Columbia chez la plante *Arabidopsis thaliana*

Con-A : Concanavaline A

C-ter : région C-terminale

cv. : cultivar

CYP-20-2 : cyclophiline 20-2 peptidyl-prolyl cis trans isomerase

DBE : Debranching enzyme

DCA : Acide deoxycholique

DMSO : Diméthylsulfoxyde

DP : degré de polymérisation

DSP : Dual specificity phosphatases

DTT : Dithiotréitol

DUF : domain uncharacterized and functional

eFASP: enhanced Filter Aided Sample Preparation

EMBL : European Molecular Biology Laboratory

ESV1 : Early StarVation 1

FDR : False discovery rate

FBN : fibrilline

g : gramme

GBSS : Granule Bound Starch Synthase

GFP: Green Fluorescent Protein

GH : Glycoside hydrolase

Glc : Glucose

Glc-1-P : Glucose-1-phosphate

Glc-3-P : Glucose-3-phosphate

Glc-6-P : Glucose-6-phosphate

GT : Glycosyltransférase

GWD : Glucan water dikinase

HCD : dissociation collisionnelle à plus forte énergie

HPLC : Chromatographie liquide haute performance

iBAQ score : intensity based absolute quantification score

ISA : Isoamylase

iST : in Stage Tip

kDa : kilo Dalton

kV : kilovolt

LC : Chromatographie liquide

LDA : Limit DextrinAse

LFQ score : Label free quantification score

LSF : Like Sex Four

LESV : Like-Early StarVation 1

Mal-6-P : Maltose-6-phosphate

MFP1 : MAR Binding Filament- Like Protein 1

Milieu MS : milieu Murashig & Skoog

mg : milligramme

MJ.L<sup>-1</sup> : Mégajoule par litre

mL : millilitre

MRC : myosin-resembling chloroplast protein

MS : spectrométrie de masse

MS/MS : spectrométrie de masse en tandem

MSAP : Miniaturisation pour la synthèse, l'analyse et la protéomique

<sup>14</sup>N : azote 14

<sup>15</sup>N : azote 15

NCBI : National Center for Biotechnology Information

nm : nanomètre

N-ter : région N-terminale

OGM : Organisme génétiquement modifié

PHS : Phosphorylase

PKPI : Potato Kunitz-type protease inhibitors

Pi : Phosphate inorganique

PII1 : Protein involved in starch initiation 1

PINII : proteinase inhibitor type II

PPDK\_N : Pyruvate phosphate dikinase, N-ter

PSM : Peptide spectrum matches

PTST : Protein Targeting to STarch

PU1 : Pullulanase

PWD : Phosphoglucan water dikinase

SDS : Sodium dodécyl sulfate

SDS-PAGE : Sodium Dodecyl-Sulfate-Polyacrylamide Gel Electrophoresis

SEM : Microscopie électronique à balayage

SEX : Starch EXcess

SMART : Simple Modular Architecture Research Tool

SS : Starch Synthase

*St* : *Solanum tuberosum*

TCEP : tris (2-carboxyethyl) phosphine

TFA : acide trifluoroacétique

THF : tétrahydrofurane

TP : Transfer Peptide

UGSF : Unité de glycobiologie structurale et fonctionnelle

UMR : Unité mixte de recherche

USR : Unité service et recherche

µg : microgramme

µL : microlitre

µm : micromètre

# Table des matières

Résumé .....	2
Abstract.....	3
Remerciements .....	4
Abréviations.....	8
Table des matières .....	13
Introduction .....	17
I. Utilisation de l'amidon .....	18
1. Utilité physiologique de l'amidon chez la plante.....	18
2. L'amidon dans l'industrie .....	18
II. Structure et composition de l'amidon .....	21
1. Structure des glucanes .....	21
2. Structure de l'amylose et de l'amylopectine .....	21
3. Cristallinité de l'amylopectine .....	25
4. Assemblage en anneaux de croissance.....	26
5. Le hile.....	28
6. Les pores et les canaux de l'amidon.....	28
7. Morphologie des grains d'amidon.....	29
8. Autres composants du grain d'amidon.....	31
III. Synthèse de l'amidon .....	31
1. Synthèse de l'ADP-Glucose.....	32
2. Élongation des glucanes de l'amidon par les amidons synthases .....	32
3. Formation des points de branchements par les enzymes de branchements.....	40
4. Rôle des enzymes de débranchement dans la stabilisation de la cristallinité.....	45
IV. Dégradation de l'amidon.....	48
1. Décompaction de l'amylopectine par les dikinases de l'amidon .....	48
2. Dégradation des chaînes linéaires par les $\beta$ -amylases.....	52

3.	La déphosphorylation par les phosphatases de l'amidon .....	56
4.	Les enzymes de débranchement dans la dégradation.....	59
5.	Les amidons phosphorylases .....	61
6.	Les enzymes de dismutations ou $\alpha$ -glucanotranférases (DPE) .....	63
7.	Les protéines sans activité catalytique connue impliquées dans le métabolisme de l'amidon <sup>64</sup>	
8.	Les protéines d'adressage « Protein Targeting to Starch » .....	65
9.	Early StarVation proteins .....	68
10.	Autres protéines récemment découvertes comme impliquées dans l'initiation : PI1 (MRC) et MFP1. ....	69
V.	L'analyse des protéines par spectrométrie de masse .....	73
VI.	La pomme de terre ( <i>Solanum tuberosum</i> ).....	76
VII.	Objectifs de la thèse .....	78
1.	Établissement du protéome complet de l'amidon de pomme de terre .....	78
2.	Fluctuation de la stœchiométrie basée sur la morphologie des grains. ....	79
3.	Dynamique d'association des protéines au grain d'amidon.....	80
4.	Impact d'une protéine sur l'ensemble du protéome, le cas des mutants <i>gbss</i> ....	81
5.	Protéomique à l'échelle du grain d'amidon .....	81
	Matériels et Méthodes .....	84
I.	Matériels .....	85
1.	Amidon utilisé pour l'étude du protéome complet, de la stœchiométrie des protéines liées au grain et pour les études grain unique.....	85
2.	Amidon utilisé pour les études de dynamiques d'association des protéines au grain	85
3.	Amidon utilisé pour l'étude de l'influence de GBSS sur l'ensemble du protéome	86
II.	Méthodes.....	87
1.	Traitements des grains d'amidon .....	87

2.	Morphologie des grains d'amidon.....	87
3.	Fractionnement des grains d'amidon en fonction de leur taille. ....	87
4.	Capture des grains d'amidon en capillaire pour les analyses avec 30 grains et 1 grain.	87
5.	Imagerie des grains d'amidon par microscopie électronique à balayage.....	88
6.	Etude de la cristallinité des grains d'amidons par diffraction aux rayons X.....	88
7.	Dosage de l'amylose. ....	89
8.	Distribution des longueurs de chaines d'amylopectine.....	89
9.	Contenus en phosphates. ....	90
10.	Extraction des protéines .....	90
11.	Electrophorèse en gel de polyacrylamide contenant du sodium dodécyl sulfate (SDS-PAGE).....	91
12.	Digestion à la trypsine en gel .....	91
13.	Digestion des protéines par « enhanced Filter Aided Sample Preparation » (eFASP) 92	
14.	Extraction et digestion des protéines de l'amidon par C18 in Stage Tip (iST)..	93
15.	Analyses en spectrométrie de masse par acquisition dépendante de données (DDA) 94	
16.	Identification des protéines associées au grain d'amidon avec PEAKS .....	95
17.	Quantification avec MaxQuant et Perseus .....	95
18.	Calcul du taux d'incorporation de <sup>15</sup> N.....	96
	Résultats.....	97
I.	Etablissement du protéome complet du grain d'amidon de pomme de terre .....	98
1.	Protéines du métabolisme de l'amidon. ....	98
2.	Les autres protéines identifiées au sein du grain d'amidon .....	103
3.	Quantification des protéines et traitement des grains d'amidon au SDS et à la thermolysine.....	106

II.	Fluctuations de la stœchiométrie des protéines en fonction de la morphologie du grain.	108
1.	Analyse morphologique des grains d'amidon de pomme de terre (cv. Monalisa) et fractionnement .....	108
2.	Structure et composition de l'amidon dans les différentes fractions.....	112
3.	Protéomique quantitative comparative des différentes fractions .....	116
III.	Dynamique d'association des protéines avec le grain d'amidon .....	124
1.	Marquage et sélection des peptides .....	125
2.	Identification des massifs de pics correspondants aux peptides d'intérêt .....	126
3.	Taux d'incorporation pour chacune des protéines. ....	129
IV.	Impact d'une protéine sur l'ensemble du protéome, étude des mutants pour la protéine GBSS. ....	130
1.	Dosage des protéines dans l'amidon de pomme de terre des mutants <i>waxy</i> ...	131
2.	Protéomique quantitative comparative des mutants <i>waxy</i> .....	132
V.	Protéomique du grain unique d'amidon.....	134
1.	Analyse protéomique à partir de 30 ou d'un unique grain d'amidon.....	135
2.	Identification d'espèces à partir de la séquence de GBSS. ....	137
	Discussions et Perspectives .....	140
I.	Établissement du protéome complet de l'amidon de pomme de terre.....	141
II.	Fluctuation de la stœchiométrie basée sur la morphologie des grains. ....	145
III.	Dynamique d'association des protéines au grain d'amidon.....	148
IV.	Impact d'une protéine sur l'ensemble du protéome, le cas des mutants <i>gbss</i> .....	150
V.	Protéomique du grain unique .....	152
	Bibliographie .....	155
	Annexes .....	181



# Introduction

## **I. Utilisation de l'amidon**

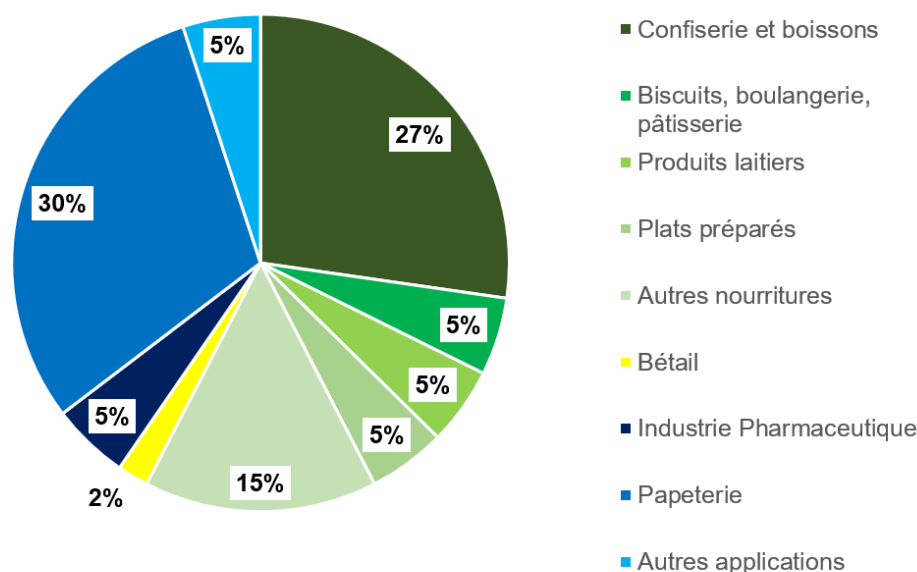
### **1. Utilité physiologique de l'amidon chez la plante.**

L'amidon permet le stockage du glucose dans les cellules de plantes terrestres et les algues vertes (*Viridiplantae* ou *Chlorobiontes*), ainsi que celles d'autres organismes photoautotrophiques telles que les algues rouges (*Rhodophytes*) et les cyanobactéries diazotrophes (*Cyanophyceae*) (Goren et al., 2018). Les plantes dites C3 et C4, qui convertissent le CO<sub>2</sub> en un composé à 3 ou 4 carbones via la photosynthèse, vont synthétiser de l'amidon durant la journée dans les chloroplastes, au niveau des tissus verts (Geiger and Servaites, 1994). Cet amidon sera ensuite consommé la nuit lors de la respiration, fournissant ainsi à la plante l'énergie nécessaire pour sa survie hors photosynthèse. Cet amidon est appelé amidon transitoire. L'amidon peut aussi être stocké de manière plus durable dans les amyloplastes, au niveau des tissus de réserve (Tyson and ap Rees, 1988). La conservation d'énergie dans les tissus de réserve va permettre à la plante de réutiliser cette énergie au moment opportun, par exemple lors de la germination dans le cas de l'amidon de l'albumen des grains de céréales, ou du tubercule de pomme de terre. Cet amidon est appelé amidon de réserve. Des grains d'amidon sont aussi retrouvés au niveau des racines dans les statolithes, des amyloplastes spécialisés dans la perception de la gravité chez la plante (Vandenbrink and Kiss, 2019).

### **2. L'amidon dans l'industrie**

L'amidon est aussi la principale source de glucides dans l'alimentation humaine ainsi que celle du bétail. Il représente 50 à 55% de la masse de notre alimentation quotidienne actuelle (Szabo de Edelenyi et al., 2014). Il est consommé sous différentes formes, du produit cru tel que dans les fruits, aux produits transformés tel que la confiserie ou les produits laitiers (Figure 1). La forte présence de l'amidon dans l'alimentation humaine commence avec la sédentarisation des populations et le développement de l'agriculture il y a 9 000 à 6000 ans (Piperno, 2011). Des grains d'amidons retrouvés sur des dents et des outils rudimentaires datant de cette époque démontrent l'existence d'un régime riche en glucide dès cette période (Piperno and Dillehay, 2008). La grande qualité de l'amidon, qui en fait l'un des plus importants nutriments dans notre alimentation actuelle, est sa versatilité. L'amidon, aussi bien natif que modifié chimiquement, est utilisé dans l'industrie agroalimentaire comme agent épaississant et comme agent gélifiant (Rapaille and Vanhemelrijck, 1997). Les propriétés de gélification et de résistance à la digestion de l'amidon sont aussi étudiées pour sélectionner de nouvelles variétés

de riz et de blé dont l'amidon possède des propriétés intéressantes lors de la cuisson (Lin et al., 2019; Hong et al., 2020). De même, des recherches sont menées sur les propriétés de l'amidon chez la pomme de terre pour pouvoir limiter les effets délétères à sa consommation, comme l'accumulation de sucres solubles dans les tubercules lors de leur conservation au froid (Tai et al., 2020). Le secteur industriel de l'amidon connaît actuellement un renouveau. La majeure partie du secteur de l'amidon était consacrée aux plantes céréalières, et plus particulièrement au maïs et au blé. L'intérêt grandissant pour une alimentation sans gluten, végétarienne et sans Organismes Génétiquement Modifiés (OGMs) de certaines populations a toutefois amené une diversification des sources d'amidon, diminuant la demande pour l'amidon issu des plantes céréalières et plus particulièrement du blé, au profit de celui d'autres plantes déjà dominantes sur le marché de l'amidon comme la pomme de terre ou le manioc; mais aussi vers des plantes qui ne représentaient avant qu'une économie de niche, comme la banane verte. Bien que l'agro-alimentaire représente la majeure partie du secteur industriel de l'amidon en Europe (60% de l'amidon purifié), celui-ci est aussi utilisé dans d'autres domaines (Figure 1.).



**Figure 1. Utilisation de l'amidon en Europe.** Graphique extrait et modifié à partir du site starch Europe (<https://starch.eu/the-european-starch-industry/#overview>). Données collectées en 2018.

L'amidon est aussi de plus en plus utilisé pour la synthèse de bioplastiques, du fait de l'intérêt grandissant pour l'utilisation de matériaux renouvelables et biodégradables. Les bioplastiques issus d'amidons natifs sauvages montrent toutefois une trop faible résistance à la tension et à l'humidité pour remplacer le plastique issu du pétrole (Zhang et al., 2014). Les tests

effectués ont montré qu'un amidon riche en amylose est plus résistant à la tension et à l'exposition à l'humidité (Phan et al., 2005; Muscat et al., 2012). L'agent plastifiant utilisé et les modifications chimiques jouent aussi un grand rôle dans la résistance du plastique. Les principaux agents plastifiants utilisés sont l'eau, le glycérol et le sorbitol. Des amides tels que l'urée, le formamide et certains acides aminés ont montré de meilleures propriétés plastifiantes, mais la présence d'amide dans les agents plastifiants empêche leur utilisation pour l'emballage de la nourriture (Zhang et al., 2014). La modification chimique de l'amidon permet aussi d'améliorer les qualités du bioplastique. L'amidon dialdéhyde montre ainsi une meilleure résistance que l'amidon natif (Yu et al., 2010). L'amidon est aussi utilisé pour la synthèse d'éthanol pour son utilisation en tant que biocarburant. L'éthanol est toutefois peu efficace en tant que carburant du fait de sa faible densité énergétique ( $\approx 25$  Méga joules par litre ( $\text{MJ.L}^{-1}$ )). Une récente étude a montré la possibilité d'obtenir du butanol à partir d'amidon de blé, dont la densité énergétique est meilleure que celle de l'éthanol ( $\approx 30 \text{ MJ.L}^{-1}$ ), mais reste plus faible que celle de l'essence et du diesel ( $\approx 35 \text{ MJ.L}^{-1}$ ) (Luo et al., 2018).

L'amidon peut aussi servir à renforcer les fibres des vêtements, rendant celles-ci plus résistantes à la traction et au déchirement. Cela permet aussi de rigidifier un vêtement de manière à lui donner une meilleure tenue. Cette meilleure tenue de la fibre a facilité l'automatisation de la découpe et de l'assemblage des vêtements. Il est aussi utilisé pour la fixation des couleurs lors de la teinture.

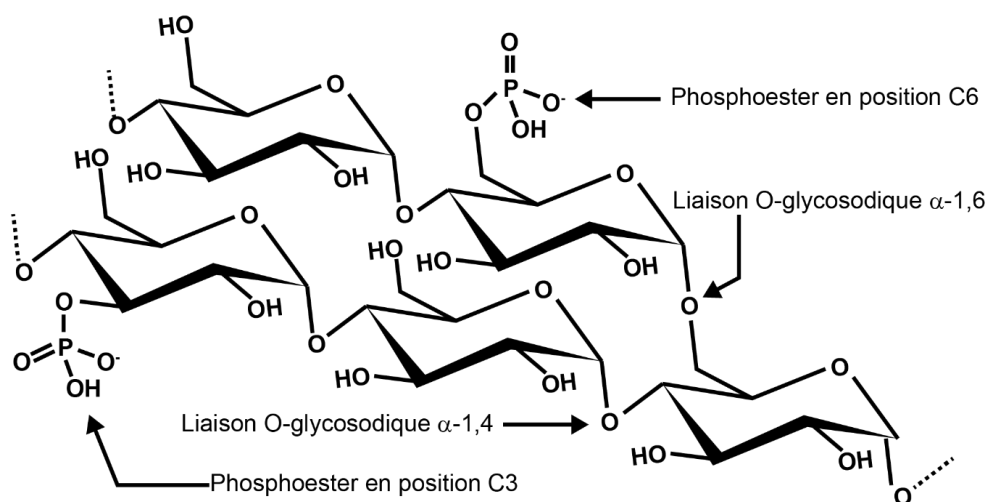
L'amidon sert à la confection du papier dans les 3 principales étapes de sa fabrication. Pour assurer la cohésion interne de la feuille et augmenter la résistance du papier, on ajoute des liaisons cellulose-amidon-cellulose aux liaisons cellulose-cellulose. De par l'augmentation de l'utilisation de papier recyclé, l'amidon est de plus en plus utilisé pour assurer cette résistance mécanique malgré la dégradation de la qualité des fibres lors du recyclage. L'amidon est utilisé lors de l'étape de presse pour éviter l'effet de peluchage dû au manque de cohésion des fibres superficielles de la feuille. L'amidon est aussi utilisé dans les produits de couchages, ces produits permettent de réduire les irrégularités de la surface du papier, remplaçant les macroporosités du papier par des microporosités propices à l'écriture et à l'impression et évitant l'effet buvard.

Enfin, il est aussi utilisé dans l'encapsulation ou comme excipient de médicaments du fait de sa facilité de dégradation par l'organisme et parce qu'il est généralement inerte et ne déclenche pas de réaction inflammatoire.

## II. Structure et composition de l'amidon

### 1. Structure des glucanes

L'amidon est composé de deux polymères de glucose, l'amylose et l'amylopectine. Ces polymères sont composés de résidus de glucose liés ensemble par des liaisons O-glycosidiques  $\alpha$ -1,4 et  $\alpha$ -1,6, ces dernières formant ce qu'on appelle des points de branchement (Figure 2). (Figure 2.). L'amylose est majoritairement linéaire avec très peu de points de branchements (moins de 1% de l'ensemble des liaisons). L'amylopectine est modérément branchée (4% à 6% de l'ensemble des liaisons). On retrouve aussi au niveau des chaînes de l'amylopectine quelques résidus de phosphate, liés de manière covalente au glucose par des liaisons phosphoesters sur les positions C3 et C6 (Figure 2.). Les grains d'amidons les plus chargés en phosphate, tels que ceux de la pomme de terre et du curcuma, possèdent à peine plus de 1 ‰ résidus de glucose-phosphate (Verbeke et al., 2016). L'assemblage de l'amylose et l'amylopectine donne un grain semi-cristallin, insoluble dans l'eau froide.



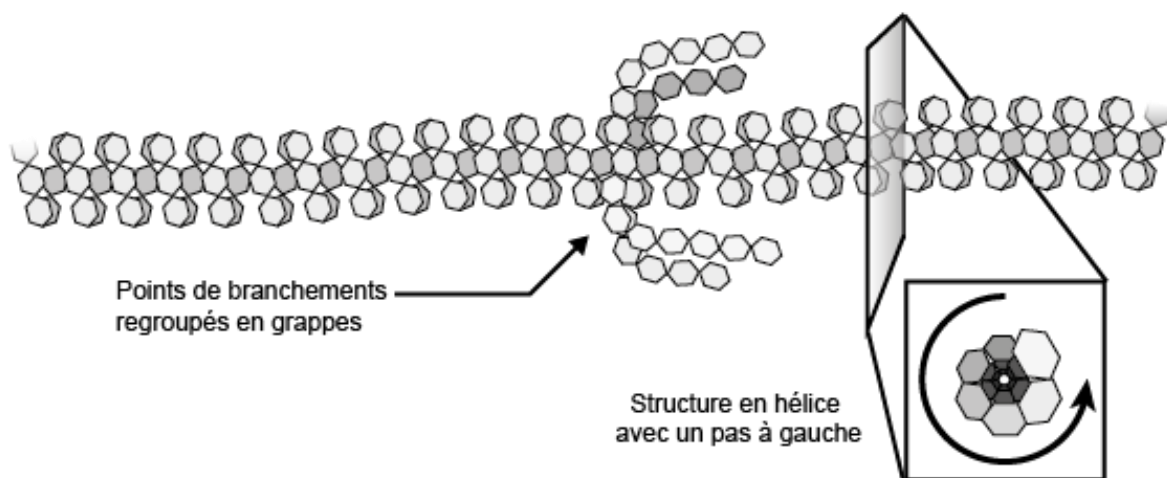
**Figure 2. Représentation schématique de la structure primaire des liaisons O-glycosidiques entre les résidus de glucose de l'amidon ainsi que des résidus de phosphates.**

### 2. Structure de l'amylose et de l'amylopectine

#### a. L'amylose

L'amylose est, comme décrit précédemment, principalement linéaire. Elle représente 15 à 30 % des polymères de l'amidon des tissus/organes de réserve chez les espèces de grande culture. Son degré de polymérisation varie de 500 à 6000 résidus de glucoses pour un poids moléculaire de plusieurs centaines de kilo Daltons (kDa). ». Deux populations de chaînes

d'amylose sont retrouvées dans le grain d'amidon. La première population est strictement linéaire, alors que la deuxième population, qui représente 25 à 55 % de l'amylose selon les espèces végétales, possède 2 à 8 points de branchements par chaîne d'amylose (Takeda et al., 1987). Les points de branchement sont regroupés en grappes de chaînes allant de 4 à une centaine de résidus de glucose (Hizukuri et al., 1997). L'amylose peut adopter une configuration cristalline *in vitro*. Elle va s'agencer sous forme d'hélices simples avec un pas à gauche allant de 6 à 8 unités de glucoses et interagir avec des lipides présents au sein du grain (Figure 3.) (Yamashita and Monobe, 1971; Godet et al., 1995).



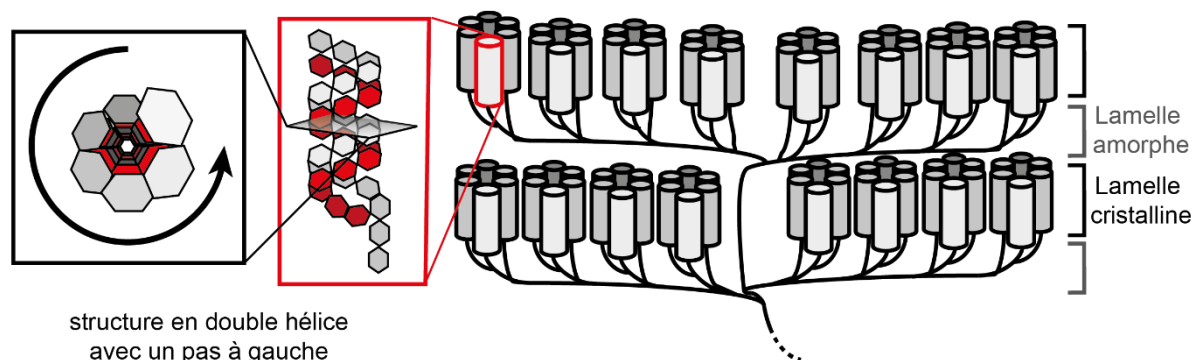
**Figure 3. Représentation schématique de l'amylose.** Chaque hexagone de la figure représente une unité de glucose.

La synthèse de l'amylose, quelle que soit l'espèce végétale ayant été étudiée, est strictement dépendante d'une amidon synthase particulière dénommée Granule Bound Starch Synthase (GBSS) (Delrue et al., 1992; Ball, 1998; van de Wal et al., 1998). En effet, toutes les plantes déficientes pour GBSS ne sont plus capables de produire de l'amylose. Cependant, cela n'entrave pas la formation d'un grain d'amidon d'apparence normale constitué uniquement d'amylopectine dont la structure cristalline reste la même que celle de son équivalent sauvage.

#### *b. L'amylopectine*

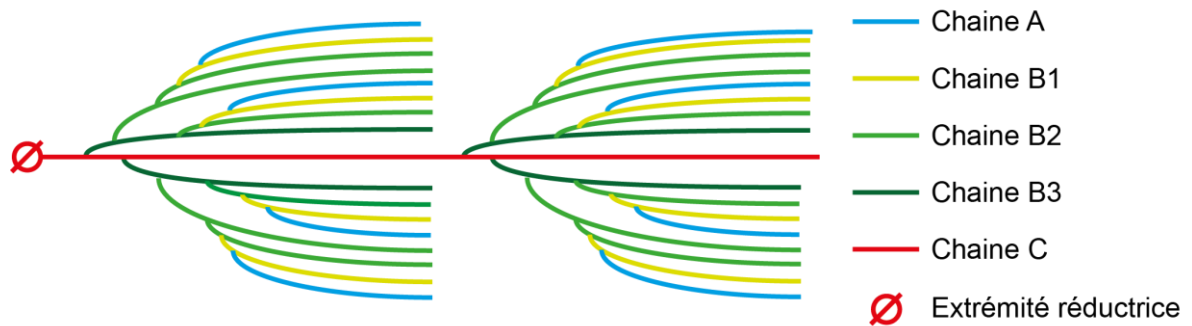
L'amylopectine est le polymère majoritaire de l'amidon sauvage, représentant 70 à 85% des polymères de l'amidon. La molécule d'amylopectine possède un DP compris entre  $10^5$  et  $10^6$  unités de glucose, pour un poids moléculaire de  $10^4$  à  $10^5$  kDas. De par son nombre plus conséquent de points de branchement que l'amylose, elle possède une structure très différente de ce dernier. C'est l'amylopectine qui fournit pour l'essentiel sa cristallinité au grain d'amidon. Les points de branchement de l'amylopectine ne sont pas distribués de manière homogène au

sein de la molécule, comme cela peut l'être pour le glycogène, mais regroupés dans certaines zones (French, 1972; Robin, 1974). Cela conduit à la formation de grappes d'amylopectine constituées à la base d'une lamelle amorphe où se trouve l'essentiel des points de branchement, prolongée d'une lamelle cristalline dans laquelle les glucanes linéaires forment des structures en double hélices parallèles avec un pas à gauche de 6 molécules de glucose (Figure 4)(Hizukuri, 1986). La taille d'une grappe composée d'une lamelle amorphe et d'une lamelle cristalline, a été déterminée à 9-10 nm quelle que soit l'origine botanique de l'amidon (Jenkins and Donald, 1995). Les lamelles amorphes mesurent environ 3 nm et les lamelles cristallines 6 nm.



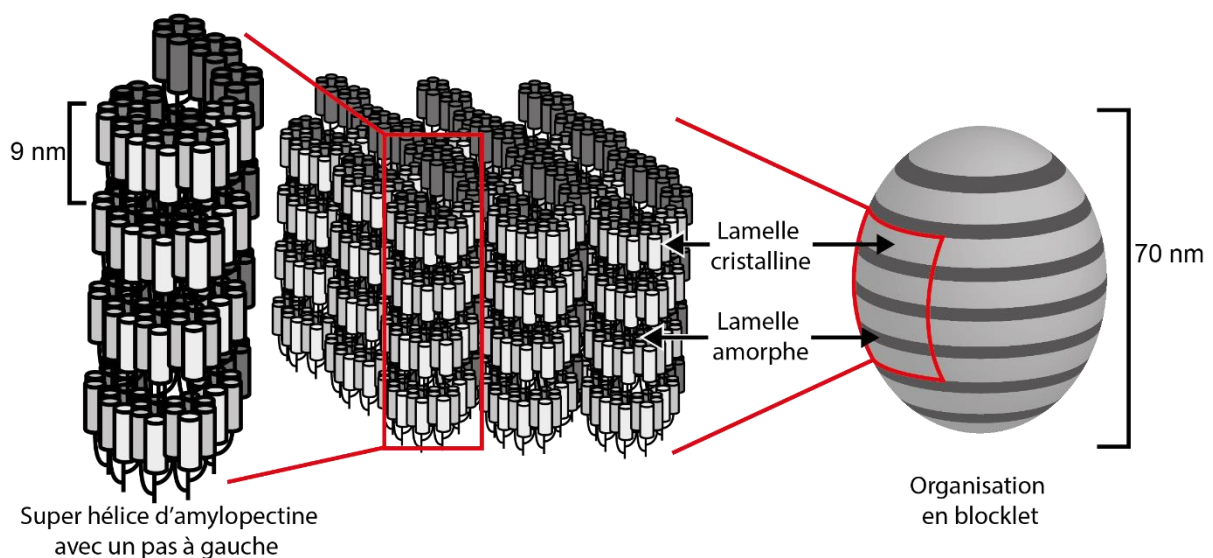
**Figure 4. Représentation schématique des grappes de l'amylopectine.** Les cylindres représentent des doubles hélices de deux glucanes, les traits noirs représentent des glucanes linéaires. Chaque hexagone représente une unité de glucose. Le glucane représenté en rouge correspond à une chaîne de type A, branchée sur un glucane représenté en gris qui forme une chaîne de type B (voir le paragraphe ci-après pour la définition des types de chaîne). Les deux chaînes s'enroulent parallèlement sous forme de double hélice avec un pas gauche de 6 unités de glucose.

On retrouve 3 types de chaînes dans l'amylopectine, classées selon leur longueur et leur degré de branchement (Figure 5)(Peat et al., 1956). Les chaînes de type A sont les chaînes les plus courtes et les plus externes de la grappe d'amylopectine. Elles ont un DP maximum de 15 et ne possèdent pas de points de branchement supplémentaires à celui qui les lie au reste de la grappe. Les chaînes de type B sont les chaînes de taille intermédiaire variant d'un DP 10 à des DP >40. Elles portent d'autres chaînes de type B ou des chaînes de type A. Elles se subdivisent en 3 sous-classes : B1, B2 et B3. Les chaînes B1 sont les plus courtes, se retrouvent en périphérie de la grappe et ne portent que des chaînes de type A. Les chaînes B2 et B3 peuvent quant à elles porter des chaînes de type A et B et sont plus longues que les chaînes de type B1. Enfin, la chaîne de type C contient la seule extrémité réductrice de l'amylopectine. Il n'y a qu'une seule chaîne C par molécule d'amylopectine.



**Figure 5. Schéma des différentes chaînes de l'amylopectines.** Chaque ligne représente une chaîne de glucane linéaire. La ligne rouge représente la chaîne C, les lignes vertes les chaînes de type B, branchées sur d'autres chaînes de type B ou sur la chaîne C. Les lignes bleues représentent les chaînes de types A, branchées sur des chaînes de type B et ne portant pas d'autres chaînes.

Les grappes d'amylopectine formées par la succession des lamelles amorphes et cristallines, s'organiseraient ensuite en superstructures dénommées « blocklets », d'une taille variant de 20 à 500 nm (Gallant et al., 1997). Ces superstructures ont été observées par microscopie électronique sur de l'amidon de pomme de terre et seraient formées par des superhélices de lamelles cristallines avec un pas à gauche de 5 nm (Oostergetel and van Bruggen, 1993; Gallant et al., 1997) (Figure 6). Si la structuration de l'amylopectine en lamelles amorphes et cristallines est aujourd'hui acceptée par l'ensemble de la communauté scientifique, la formation de superstructures fait encore débat. Ces superstructures pourraient également être dues à des artefacts provoqués lors de la préparation des échantillons, notamment du fait de la déshydratation de l'amidon. L'amylopectine déshydratée s'agrègerait pour former les blocklets, qui ne seraient pas présent à l'état naturel dans le plaste.

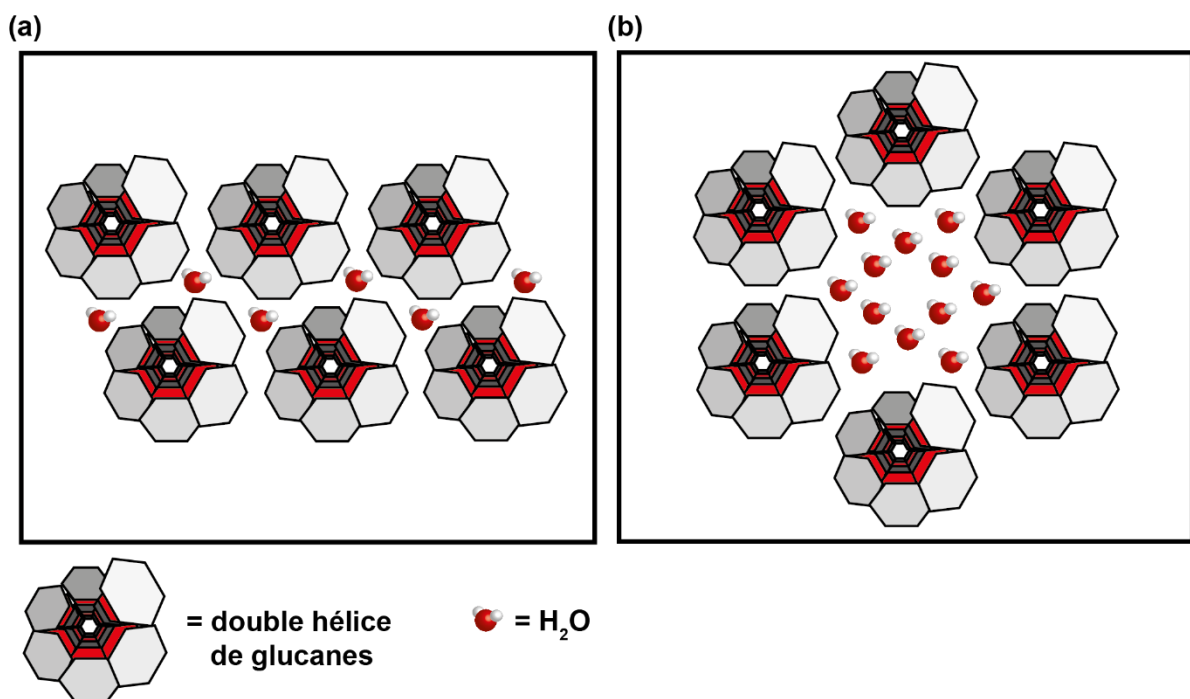




**Figure 6. Représentation des superstructures de l'amylopectine.** Cette représentation est basée sur le modèle proposé chez la pomme de terre par Gallant et coll. (1997) (Gallant et al., 1997). Les cylindres de la figure représentent les doubles hélices des grappes de l'amylopectine.

### 3. Cristallinité de l'amylopectine

L'assemblage en grappe des chaînes d'amylopectine va procurer à l'amidon son caractère partiellement cristallin. L'analyse par diffraction des rayons X de l'amidon a permis de mettre en évidence l'existence de deux allomorphes cristallins, A et B (Figure 7). Ces allomorphes se distinguent par l'organisation des doubles hélices au sein de la grappe d'amylopectine et notamment par la position des molécules d'eau au sein du cristal. Dans les cristaux de type A, retrouvés principalement dans l'amidon de réserve des céréales, l'empaquetage des doubles hélices d'amylopectine forme un maillage dense sous forme de feuillets, laissant peu de place aux molécules d'eau pour s'agencer entre les doubles hélices. Les cristaux de type B, retrouvés principalement dans l'amidon des tissus hydratés tels que les feuilles ou les tubercules, ont une structure plus lâche en forme d'hexagone, laissant plus de place aux molécules d'eau pour s'accumuler entre les doubles hélices (Imberty et al., 1988; Pérez and Bertoft, 2010).



**Figure 7. Structures cristallines de l'amidon.** (a) Structure des cristaux de type A. (b) Structure des cristaux de type B. Adaptation de Imberty et coll. (1988) (Imberty et al., 1988).

#### 4. Assemblage en anneaux de croissance.

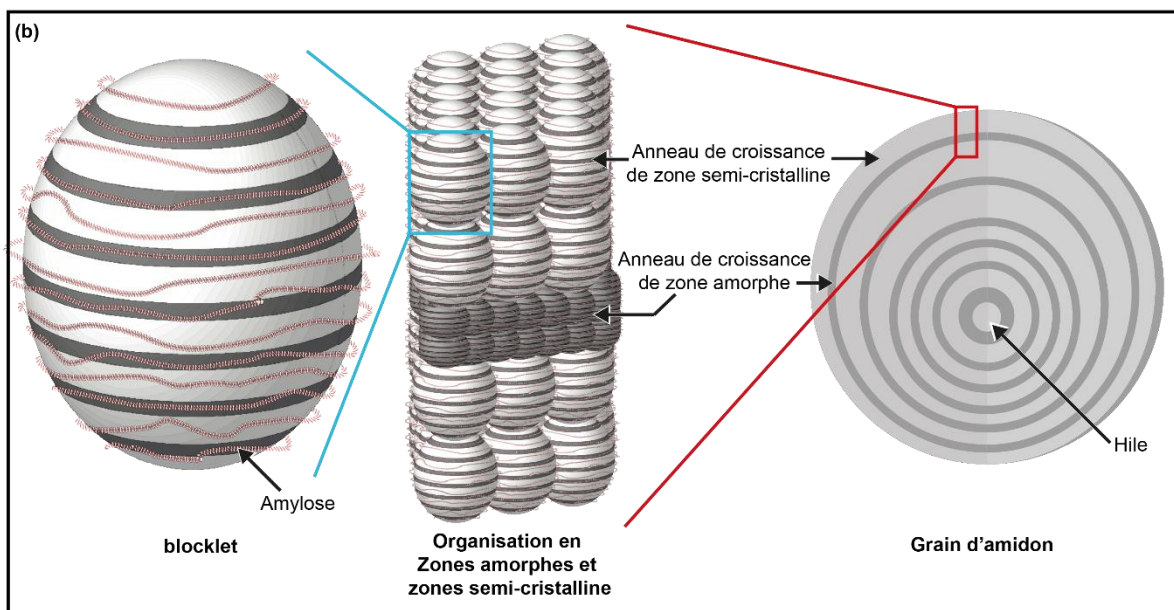
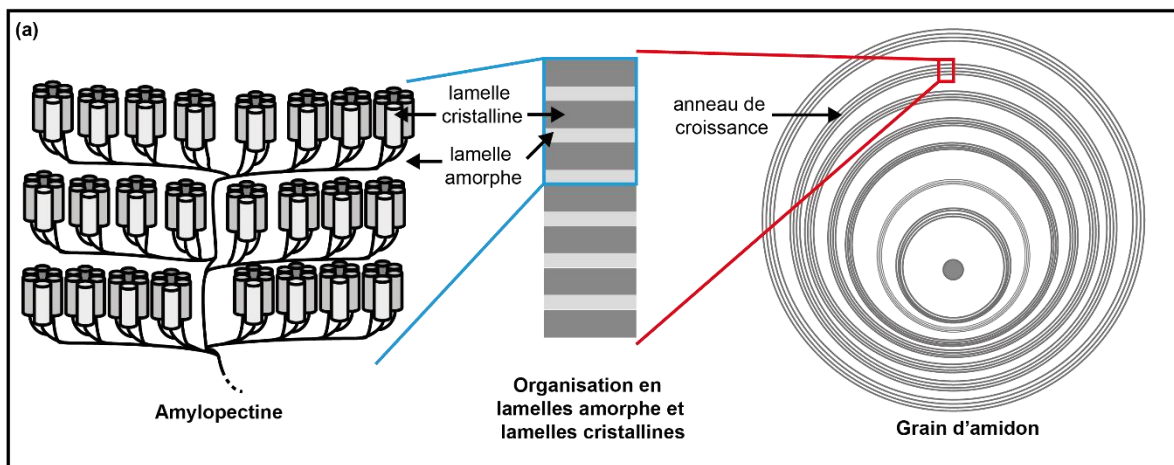
Les molécules d'amylopectine, structurées elles-mêmes en lamelles amorphes et cristallines, puis organisées possiblement en superstructures de type superhélices et blocklets, vont ensuite s'agencer sous formes de couches concentriques de manière à former le grain d'amidon. Ces couches sont organisées en symétrie radiaire, sous la forme de sphérulites, les chaînes de l'amylopectines étant organisées de manière perpendiculaire à la surface. Cela a été démontré par l'apparition d'une « croix de malte » lors de l'observation des grains à la lumière polarisée, et ce quelle que soit l'espèce végétale d'origine de l'amidon (Figure 8) (Ziegler et al., 2005).



**Figure 8. Imagerie de grains d'amidon de pomme de terre sous lumière polarisée.** Dr. Robert Berdan, *Potato starch grains in polarized light 100X*. Image numérique. Canadian nature photographer (<https://www.canadiannaturephotographer.com/>), édité par Dr. Robert Berdan. Image Numérique publiée le 22 Juin 2017. Les grains d'amidon apparaissent comme des structures ovoïdes de couleur claire avec une croix de Malte (traits sombres dans les grains) qui indiquent une organisation radiale des molécules.

Ces couches partent du centre du grain, appelé hile, vers l'extérieur et sont qualifiés d'anneaux de croissance (Figure 9a)(Yamaguchi et al., 1979; Pérez and Bertoft, 2010). L'épaisseur de ces anneaux de croissance varie de 100 à 400 nm d'épaisseur. Les anneaux de croissance se succèdent en anneaux amorphes, peu dense aux électrons lors de l'observation par microscopie électronique, et en anneaux semi-cristallins, plus denses aux électrons (Yamaguchi et al., 1979). Les anneaux de croissance semi-cristallins sont plus résistants à l'hydrolyse que les anneaux de croissance amorphe. La composition des zones amorphes de l'amidon reste mal connue à ce jour. Elles seraient composées de matériels intermédiaires peu structurés, d'amylose, ou d'amylopectine plus lâche que celle des anneaux de croissances semi-

cristallins. La localisation précise de l'amylose au sein des anneaux de croissance reste à ce jour inconnue. L'hypothèse la plus probable actuellement serait que l'amylose est entremêlée sous forme de chaînes individuelles autour des grappes d'amylopectine (Pérez and Bertoft, 2010). Elle pourrait ainsi servir au maintien de la structure des chaînes de l'amylopectine. Dans l'amidon *waxy*, c'est-à-dire dépourvu d'amylose après mutation de la plante, on constate lors de la cuisson que les grains d'amidon gélatinise plus vite et plus facilement, ce qui peut laisser penser que sans amylose, l'amylopectine se décompacte plus facilement sous l'effet de la chaleur (Hsieh et al., 2019). Des expériences de cross-linking sur l'amylose et l'amylopectine de maïs ont en effet démontré que les chaînes d'amylose n'interagissaient pas entre elles, mais qu'elles interagissaient avec les chaînes de l'amylopectine (Kasemsuwan and Jane, 1994). Si l'on considère l'hypothèse des blocklets comme valide, l'amylose pourrait former un maillage autour de l'amylopectine pour stabiliser la structure des blocklets (Figure 9b) (Ridout et al., 2003).



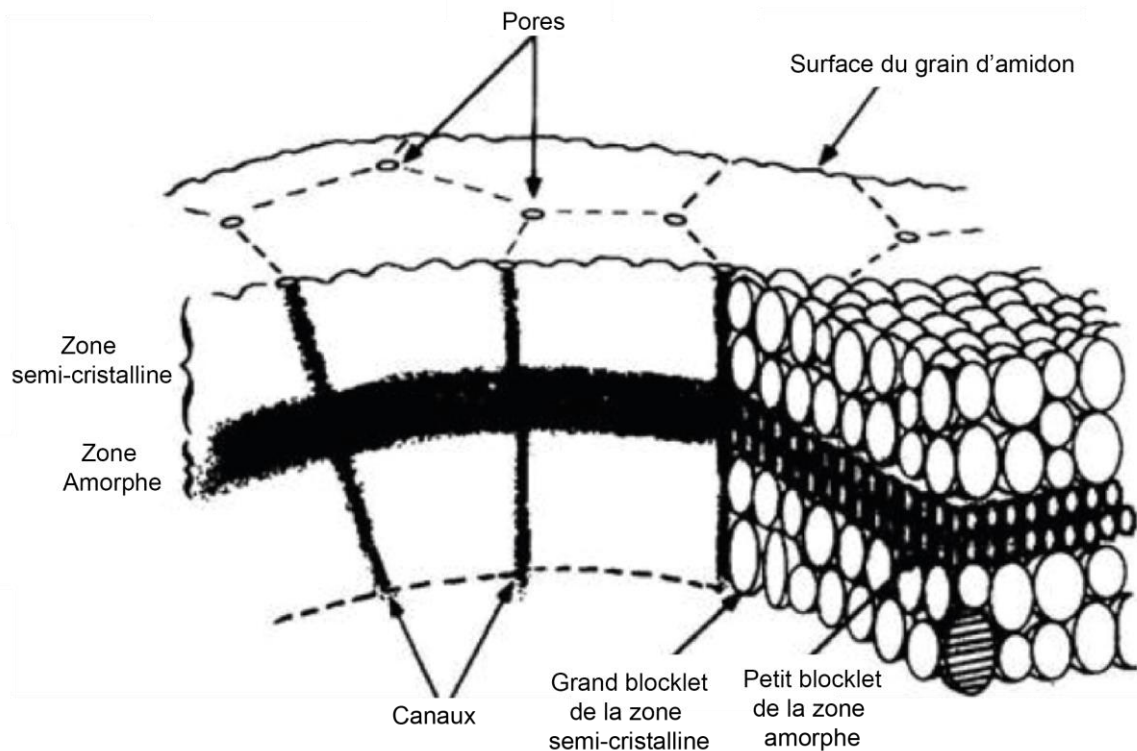
**Figure 9. Représentations des anneaux de croissances d'un grain d'amidon.** (a) représentation schématique des anneaux de croissance montrant l'alternance des zones amorphes et des zones semi-cristallines, elles-mêmes composées des lamelles amorphes et cristallines de l'amylopectine. (b) représentation schématique de l'organisation structurale des anneaux de croissance tenant compte de l'hypothèse des superstructures en blocklets proposées par Gallant et coll. (1997), ainsi que de l'organisation de l'amylose par Ridout et coll. (2003) (Gallant et al., 1997; Ridout et al., 2003).

## 5. Le hile.

Le « foyer probable de l'initiation de la synthèse » du grain, autour duquel les anneaux de croissances sont compilés, est appelé hile. Tout comme les zones amorphes, le hile est peu dense aux électrons (Yamaguchi et al., 1979). Encore aujourd'hui, la composition du hile reste mal connue, notamment du fait de la complexité à étudier sa structure. Il semble toutefois que le hile soit composé de glucanes s'apparentant à de l'amylose ou à du matériel intermédiaire peu branché (Ziegler et al., 2005). L'hypothèse actuelle de la mécanique de synthèse du grain voudrait que le hile soit le siège de l'initiation de la synthèse, et que les anneaux de croissance viendraient s'apposer couche par couche autour de celui-ci. Mais cela reste à démontrer formellement.

## 6. Les pores et les canaux de l'amidon

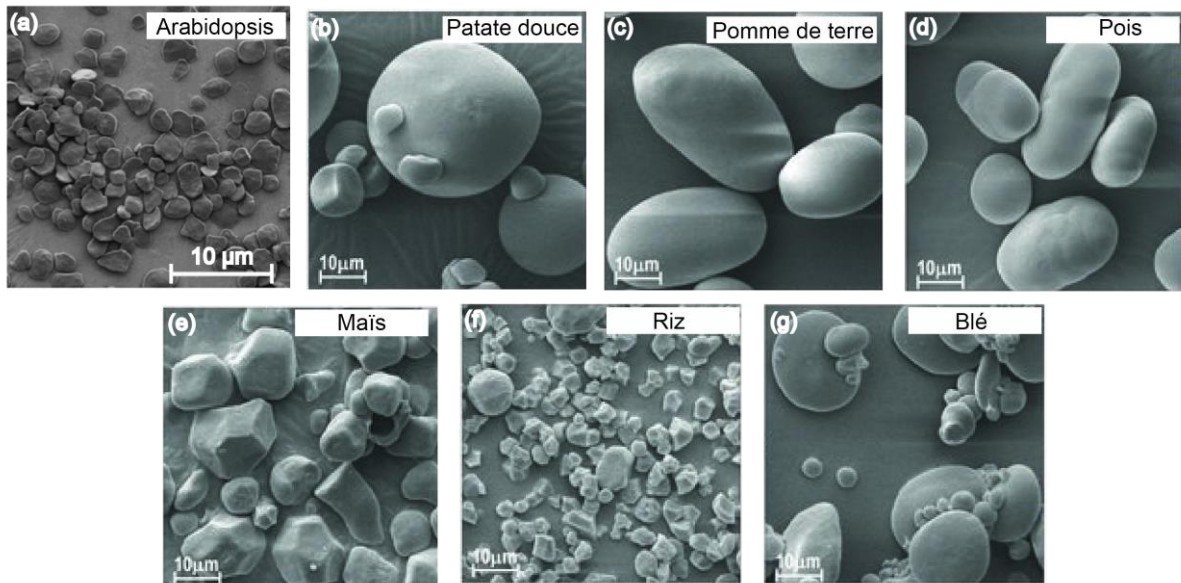
Chez certaines espèces (maïs, blé et sorgo, riz et manioc) de petits pores et canaux, reliant la surface du grain aux parties internes, ont été observés (Figure 10) (Fannon et al., 1992; Fannon et al., 1993). Chez le sorgo, ces canaux ont un diamètre variant de 100 à 300 nm en surface et de 70 à 100 nm à l'intérieur du grain. Les études qui ont été menées par la suite sur les canaux du sorgo, du maïs et du blé ont mis en évidence la présence de protéines spécifiques (Han et al., 2005). La coloration au 3-(4-carboxybenzoyl) quinoline-2-carboxaldehyde (CBQCA), un colorant de grande sensibilité pour les amines permettant de détecter des protéines et des peptides encapsulés et ne réagissant pas avec les lipides, a montré la présence de protéines au niveau des canaux. L'analyse des bandes de polypeptides trouvées après extraction des protéines de surfaces et des canaux, par traitement à la thermolysine, a permis d'identifier, chez le maïs, la protéine Brittle-1 (BT-1), un transporteur de nucléotides-sucres de type ADP-glucose. Les travaux sur ces structures font toutefois débat au sein de la communauté scientifique. La présence de pores et de canaux peut aussi s'expliquer par des artefacts dû à la formation de fissures par dégradation partielle, mécanique ou enzymatique, des grains d'amidon lors de l'extraction du tissu végétal.



**Figure 10. Représentation des pores et des canaux à la surface d'un grain d'amidon.** Les zones sombres représentent les zones amorphes du grain, composé des anneaux de croissances amorphes et des pores. Chaque sphère représente un blocklet. Adapté de Galland et coll. (1997) (Gallant et al., 1997).

## 7. Morphologie des grains d'amidon

Les grains d'amidon ont des tailles et formes diverses selon leur origine botanique. La taille d'un grain d'amidon peut varier d'une centaine de nm pour les plus petits jusqu'à 100  $\mu\text{m}$  pour les plus gros (Figure 11)(Buléon et al., 1998). Certains ont la forme de galets ronds et aplatis, c'est notamment le cas de l'amidon transitoire tel que retrouvé dans des chloroplastes d'*Arabidopsis thaliana*. D'autres ont une forme plus sphérique tel que l'amidon de réserve de patate douce (*Ipomoea batatas*). D'autres ont une forme ovoïde tel que l'amidon de tubercule de pomme de terre (*Solanum tuberosum*) ou de pois (*Pisum sativum*). Certaines espèces accumulent des grains de forme polygonale, comme le riz (*oryza sativa*) et le maïs (*Zea mays*). Chez le blé (*Triticum aestivum*), on retrouve deux populations de grains : les grains A dont la taille varie de 15 à 35  $\mu\text{m}$  et de forme aplatie et les grains de type B plus petits de 2 à 10  $\mu\text{m}$  et de forme sphérique (Tanaka et al., 2017).



**Figure 11. Images obtenues par microscopie électronique à balayage (SEM) de grains d'amidons de différentes espèces.** (a) Imagerie par SEM de grains d'amidon transitoire d'*Arabidopsis thaliana*, fond génétique Col-0. Adapté de Vandromme et coll. (2019) (Vandromme et al., 2019). (b)-(g) Images de SEM de grains d'amidon de réserve provenant de différentes espèces végétales : (b) patate douce (*Ipomoea batatas*), (c) pomme de terre (*Solanum tuberosum*), (d) pois (*Pisum sativum*), (e) maïs (*Zea mays*), (f) riz (*Oryza sativa*) et (g) blé (*Triticum aestivum*). Adapté de Khalid et coll. (2017) (Khalid et al., 2017).

Ces différences morphologiques pourraient, entre autres, s'expliquer par les contraintes d'encombrement au niveau du plaste, le lien de synthèse des grains d'amidon dans la cellule végétale. L'amidon transitoire d'*Arabidopsis* étant synthétisé entre les thylakoïdes, sa petite taille et sa forme aplatie pourraient s'expliquer par le faible espace disponible entre deux thylakoïdes (Gámez-Arjona et al., 2014). Au contraire, la grande taille des cellules et des amyloplastes du tubercule de pomme de terre permettrait la synthèse de grains atteignant jusqu'à 100 µm. Chez le blé, la distribution bimodale des grains d'amidon pourrait s'expliquer par une localisation différentielle des grains de type A et de type B au niveau de l'amyloplaste de l'albumen de blé. Les grains de type A sont synthétisés au centre de l'amyloplaste, alors que les grains de type B sont synthétisés au niveau d'extensions de la membrane de l'amyloplaste, appelés stromules, avec un espace plus limité (Tanaka et al., 2017). D'autres paramètres entrent en jeu dans la morphologie du grain, puisque les mutants pour certaines protéines responsables de la synthèse de l'amidon peuvent tout de même permettre la synthèse d'un grain, mais avec une morphologie différente sans que cela n'impacte la morphologie du plaste (Wattebled et al., 2005; Grzes et al., 2011; Szydłowski et al., 2011; Vandromme et al., 2019).

## **8. Autres composants du grain d'amidon**

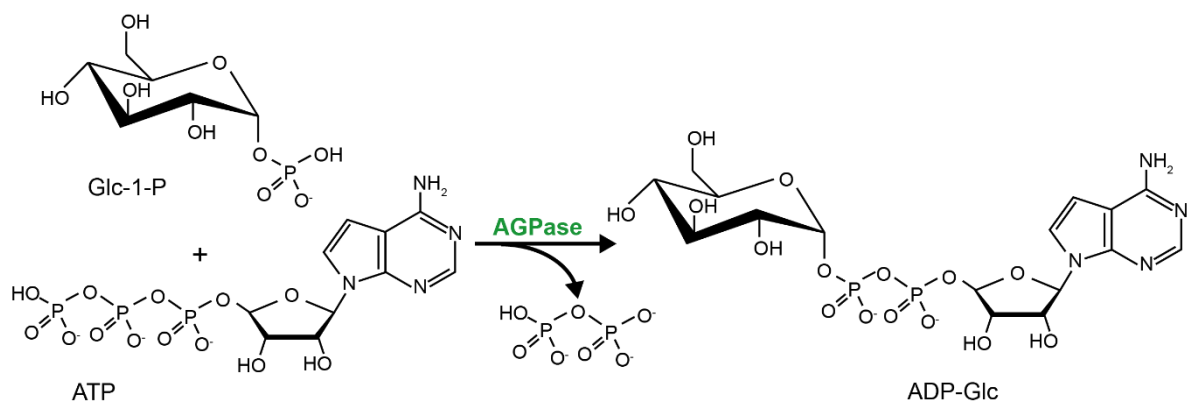
Au sein des glucanes composant l'amidon, on retrouve aussi des phosphoesters. Certaines chaînes sont phosphorylées, en position C3 ou en position C6 du glucose (Ritte et al., 2006). Ces phosphorylations permettraient la décompaction des chaînes linéaires de l'amylopectine, et ainsi d'initier le phénomène de dégradation de l'amidon (Blennow and Engelsen, 2010). La quantité de phosphoesters au sein du grain d'amidon reste très faible (moins de 1%) (Verbeke et al., 2016). Les tests de digestion *in vitro* menés sur de l'amidon de pomme de terre indiquent que les résidus de phosphates seraient principalement situés au centre des chaînes B de l'amylopectine, 9 résidus avant un point de branchement (Takeda, 1982). En plus de l'amylose et de l'amylopectine, d'autres composants sont retrouvés dans le grain d'amidon. On retrouve des lipides (jusqu'à 1.5% du poids sec), principalement dans l'amidon retrouvé au niveau de l'albumen des plantes céréalières (Morrison et al., 1984). L'amidon de tubercule de pomme de terre ne possède en revanche qu'une très faible quantité de lipides. Ces lipides se retrouvent sous forme d'acide gras ou de lysophospholipides. Du « matériel intermédiaire », nom donné à des dextrans ne pouvant ni être affiliés à de l'amylose, ni à de l'amylopectine, est aussi retrouvé au sein des grains (Yoon and Lim, 2003). Son poids moléculaire est intermédiaire aux deux polymères. Il possède trop de points de branchement pour être considéré comme de l'amylose et n'a pas les caractéristiques structurales de l'amylopectine. Sa quantité reste faible dans les grains d'amidons issus de plantes sauvages, mais semble être augmentée chez certains mutants. Des protéines, là aussi en très faible quantité (de 0,1% à 0,7% du poids du grain d'amidon selon les espèces), sont aussi retrouvées dans les grains d'amidon (Biliaderis et al., 1979; Yusuph et al., 2003). La majeure partie des protéines ayant été retrouvée au sein du grain est liée à son métabolisme (Grimaud et al., 2008).

## **III. Synthèse de l'amidon**

Bien que l'amidon soit une molécule chimiquement simple, sa structure est complexe, et implique donc un mécanisme de synthèse complexe nécessitant des dizaines d'activités enzymatiques ainsi que d'autres protéines non catalytiques jouant des rôles de régulation, d'adressage ou de maintien de la structure. Dans cette partie seront exposées les différentes étapes de la synthèse de l'amidon et le rôle précis que joue chaque enzyme impliquée.

## 1. Synthèse de l'ADP-Glucose

L'élongation des polymères de l'amidon est dépendante de l'ADP-Glucose (ADP-Glc). Celui-ci est synthétisé chez la plante à partir d'adénosine triphosphate (ATP) et de Glucose-1-Phosphate (Glc-1-P) par une transférase, l'ADP-Glucose pyrophosphorylase (AGPase) (Figure 12). Le Glc-1-P est lui-même obtenu par néoglucogénèse à partir du Glycéraldéhyde 3 phosphate (G3P) issu du cycle de Calvin. La synthèse d'ADP-Glc est conditionnée par la quantité de 3-phosphoglycérate (3-PGA) et de phosphate inorganique (Pi) dans le milieu (Mugford et al., 2014). En effet, plus le Pi sera abondant, plus l'activité de l'AGPase sera inhibée, au profit de la synthèse de saccharose.

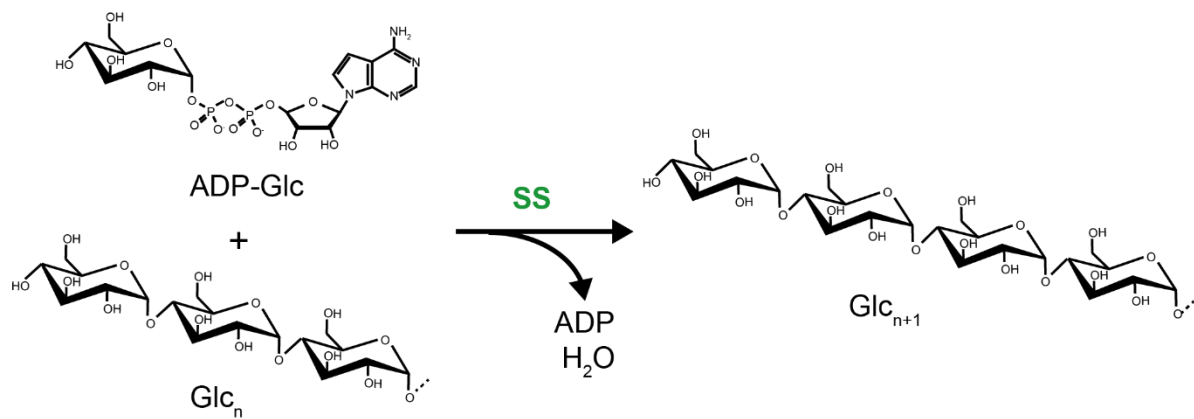


**Figure 12. Formation de l'ADP-Glucose.** Glc-1-P : Glucose 1 phosphate, ATP : Adénosine triphosphate, ADP-Glc : ADP-glucose. AGPase : ADP-glucose pyrophosphorylase.

## 2. Élongation des glucanes de l'amidon par les amidons synthases

Les glucanes de l'amidon sont allongés à partir d'ADP-Glc par les amidon-synthases (SSs). Celles-ci ajoutent une unité de glucose à l'extrémité non-réductrice du glucane en créant une nouvelle liaison  $\alpha$ -1,4 (Figure 13). 6 classes d'amidon-synthases sont aujourd'hui connues et caractérisées : la GBSS, et 5 amidon-synthases solubles : SS1, SS2, SS3 et SS4, et SS5 (Pfister and Zeeman, 2016; Abt et al., 2020). Une 7<sup>ème</sup> classe, SS6, a été retrouvée dans le génome de la pomme de terre (Deschamps et al., 2008; Liu et al., 2015; Van Harsselaar et al., 2017). Un preprint sur la caractérisation de cette enzyme chez le manioc est actuellement disponible sur BioRxiv, il n'a toutefois pas encore été revu par les pairs et publié dans un journal scientifique à ce jour (He et al., 2020).



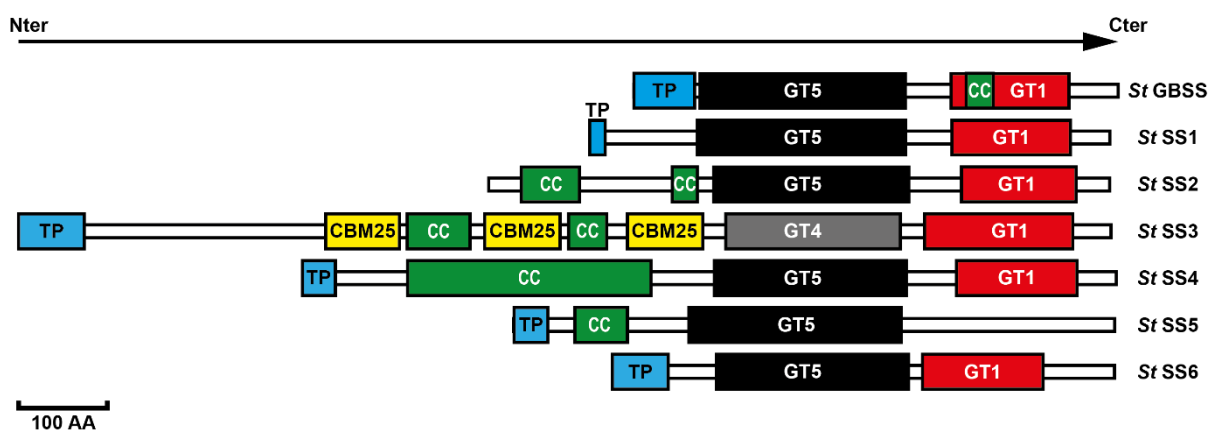


**Figure 13. Elongation d'un de glucane par les amidons synthases.** ADP-Glc : ADP-Glucose,  $\text{Glc}_n$  : Glucane linéaire à  $n$  unités de glucose.  $\text{Glc}_{n+1}$  : Glucane linéaire à  $n+1$  unités de glucose. SS : amidon synthase.

#### a. GBSS

La GBSS est l'amidon synthase responsable de la biosynthèse de l'amylose (Denyer et al., 1995b). Elle est aussi capable d'allonger les chaînes d'amylopectine (Pfister and Zeeman, 2016). Il semble toutefois que cette synthèse au niveau de l'amylopectine ne soit pas un produit final mais un intermédiaire de la synthèse de l'amylose (van de Wal et al., 1998). Cet intermédiaire serait ensuite clivé par des enzymes de branchement pour former l'amylose, selon le modèle de synthèse proposé par Ball et coll. (1998) (Ball, 1998). Les mutants dépourvus de GBSS active, appelés « *waxy* » en raison de l'aspect cireux des grains de maïs mutés pour ce gène, n'accumulent pas d'amylose (Tsai, 1974; Okagaki and Wessler, 1988). Les autres synthases n'ont pas d'activité redondante avec GBSS puisqu'elles sont incapables de synthétiser de l'amylose. Contrairement aux autres amidon synthases dont l'activité est distributive, GBSS est une enzyme processive (Denyer et al., 1999). GBSS est la première enzyme à avoir été trouvée associée de manière non covalente à l'amidon *in vivo* parce qu'elle y est, et de loin, la plus abondante. GBSS est peu active sous forme soluble, elle doit être liée au grain d'amidon pour être active (Dry et al., 1992). Cette enzyme est de petite taille comparée aux autres amidon synthases, avec un poids moléculaire de 66,6 kDa pour GBSS de pomme de terre (van der Leij et al., 1991a; Van Der Leij et al., 1991b). Elle possède dans sa séquence le domaine catalytique C-terminal comprenant deux domaines Glycosyl-Transférase (GT) de la classification CAZy, communs aux amidon et glycogènes synthases (Figure 14)(Coutinho et al., 2003). Par ailleurs, on retrouve aussi dans sa séquence les motifs K-X-G-G-L et X-X-G-G-L, qui sont respectivement localisés aux extrémités N-ter et C-ter du domaine catalytique (Cao et al., 1999). GBSS ne possède toutefois pas de domaines de liaison à l'amidon, ce qui est

étonnant du fait que cette enzyme est principalement retrouvée liée au grain d'amidon. Il a été montré chez *Arabidopsis* que l'adressage de GBSS au grain d'amidon est permis par une autre protéine, Protein Targeting to STarch 1 (PTST1) (Seung et al., 2015). L'analyse de l'expression de l'ARNm de GBSS chez le muflier à grande fleur (*Antirrhinum majus*) durant l'alternance jour/nuit a montré que GBSS est exprimée durant le jour, mais que son expression est fortement diminuée en fin de jour et absente la nuit (Mérida et al., 1999). L'expression de GBSS suivrait donc le rythme circadien. GBSS semble aussi être régulée par mécanisme de phosphorylation, une étude sur l'orge ayant montré la capacité de GBSS à être phosphorylée (Ahmed et al., 2015).



**Figure 14. Représentation des domaines structuraux des amidons synthases de pomme de terre (*Solanum tuberosum*).** Les séquences de GBSS, SS1, SS2 et SS3 utilisées pour cette figure sont extraites du Potato Genome sequencing consortium PGSC\_DM\_v3.4\_pep\_non-redundant (Avril 2017, 52.570 entrées) (Xu et al., 2011). Les séquences de SS4, SS5 et SS6 ont été fournies par François Sevestre, CNRS, UGSF UMR8576, Université de Lille, Villeneuve d'ascq, à partir du reséquençage du génome de la pomme de terre cv. Désirée (Sevestre et al., 2020). Les peptides d'adressage au plaste (TP) ont été identifiés grâce à au logiciel en ligne ChloroP 1.1 (<http://www.cbs.dtu.dk/services/ChloroP/>)(Emanuelsson et al., 1999). Les régions coiled-coil ont été identifiées à l'aide de Paircoil2 (<http://cb.csail.mit.edu/cb/paircoil2/paircoil2.html>)(McDonnell et al., 2006). Les autres régions ont été identifiées avec SMART (<http://smart.embl-heidelberg.de/>)(Letunic and Bork, 2017). Les carrés noirs et rouges représentent les domaines Glycosyl-Transférases (GT1, GT4 et GT5 selon la classification CAZy (Lombard et al., 2014)). Les rectangles jaunes représentent les modules de liaisons aux carbohydrates, les rectangles verts représentent les régions coiled-coil permettant des interactions protéine-protéine. CBM25 : Carbohydrate Binding Module 25. CC : région coiled-coil. *St* : *Solanum tuberosum*. GBSS : Granule bound starch synthase, SS : Soluble starch synthase. AA : Acides aminés.

### b. SS1

L'amidon synthase SS1 est une des trois amidon-synthases solubles responsables de la synthèse de l'amylopectine. Elle fut initialement découverte chez le maïs en 1979 (Boyer and Preiss, 1979). Elle fut ensuite retrouvée chez de nombreuses espèces végétales, sa séquence

étant très conservée au sein de ce règne (Baba et al., 1993; Mu et al., 1994; Tyynelä et al., 1995; Kossmann et al., 1999; Li et al., 1999a; Lin and Jeang, 2005)(Fujita et al., 2006). SS1 est retrouvée soluble dans le stroma du plaste, mais aussi liée aux grains d'amidon chez certaines espèces comme le blé, le maïs, l'orge, et le riz (Rahman et al., 1995; Mu-Forster et al., 1996; Borén et al., 2004; Umemoto and Aoki, 2005). SS1 possède un poids moléculaire de 70,6 kDa chez la pomme de terre (Kossmann et al., 1999). Tout comme GBSS, SS1 possède les domaines catalytiques GT1 et GT5 (Figure 14). Les tests d'activité réalisés avec SS1 de maïs recombinante ont montré que son activité diminuait avec l'augmentation du degré de polymérisation (DP) de son substrat (Commuri and Keeling, 2001). SS1 aurait donc comme substrat préférentiel les chaînes courtes de l'amylopectine. De plus, SS1 resterait fixée à son substrat au cours de la croissance du grain, ce qui expliquerait sa présence au sein du grain d'amidon (Li et al., 1999a; Delvallé et al., 2005). Chez *Arabidopsis*, l'étude des mutants dépourvus de l'activité SS1 a montré que cette enzyme a un impact sur la structure de l'amylopectine, avec une diminution des chaînes de DP 8 à 12 et une augmentation des chaînes de DP 17 à 20 (Delvallé et al., 2005). Le phénotypage de lignées de riz mutées au locus *SS1* a montré un impact similaire sur la structure de l'amylopectine ; le nombre de chaînes de DP 8 à 12 est diminué alors que celui des chaînes de DP 16 à 19 est augmenté (Fujita et al., 2006). Ces études phénotypiques confirment que SS1 allonge les chaînes courtes les plus externes, de type A ou B1, de l'amylopectine. Il existerait toutefois une redondance fonctionnelle entre SS1 et SS3, qui expliquerait le phénotype moins marqué des mutants *ss1* par rapport aux mutants *ss2* (Grzes et al., 2011). Des études menées sur le blé et le maïs ont montré que SS1 peut interagir avec d'autres protéines et former des complexes (Hennen-Bierwagen et al., 2008; Tetlow et al., 2008). Chez le maïs, SS1 interagirait avec SS3 et l'enzyme de branchement BE2a (Hennen-Bierwagen et al., 2008). Chez le blé, des complexes multimériques ont été observés entre SS1, SS2a et les enzymes de branchement BE2a et BE2b (Tetlow et al., 2008). SS1, comme GBSS, est phosphorylée lorsqu'elle est liée au grain d'amidon chez l'orge (Ahmed et al., 2015).

### c. SS2

L'amydon synthase 2 (SS2) a été décrite pour la première fois chez le pois et la pomme de terre en 1992 (Dry et al., 1992). Cette enzyme était considérée à l'époque comme une isoforme de GBSS du fait de sa forte abondance dans les grains d'amidon, et fut par conséquent nommée GBSSII. Les analyses de séquences, le fait que l'enzyme reste active sous forme soluble, et son incapacité à synthétiser de l'amylose montrèrent toutefois que cette enzyme est une amidon-synthase soluble distincte de GBSS et fut renommée SS2 (Denyer and Smith, 1992; Denyer et

al., 1993; Denyer et al., 1995a). Chez les monocotylédones, il existe plusieurs formes de SS2. Deux isoformes de SS2 sont retrouvées chez le maïs, SS2a et SS2b, et 3 isoformes sont observées chez le riz, SS2-1, SS2-2 et SS2-3 (Harn et al., 1998; Jiang et al., 2004). Chez le maïs, on retrouve principalement SS2a au niveau de l'albumen et SS2b dans les feuilles (Harn et al., 1998). Chez le riz, SS2-1 est présente dans les deux organes tout au long du développement de la plante. SS2-2 est exprimée dans les chloroplastes notamment lors des stades précoces de développement. Au contraire, SS2-3 est retrouvée dans les amyloplastés et s'exprime au cours des stades tardifs (Jiang et al., 2004). Une seule isoforme a été identifiée chez les dicotylédones comme la pomme de terre, le pois et *Arabidopsis* (Dry et al., 1992; Zhang et al., 2008). SS2 est une protéine de 85,2 kDa (chez la pomme de terre) possédant, tout comme SS1, un domaine catalytique en C-ter composé de domaines GT5 et GT1 (Figure 14)(Edwards et al., 1995). Les mutants déficients en SS2, quelle que soit l'espèce, ont un défaut d'accumulation d'amidon (Fontaine et al., 1993; Craig et al., 1998; Umemoto et al., 2002; Morell et al., 2003; Zhang et al., 2004; Zhang et al., 2008; Grzes et al., 2011). De plus la structure et la cristallinité de l'amidon sont modifiées. Le nombre de chaînes de DP 12 à 30 de l'amylopectine est fortement diminué. L'amidon synthase SS2 allongerait donc spécifiquement les chaînes de taille intermédiaire (Fontaine et al., 1993). Une redondance fonctionnelle est toutefois observée avec SS3 qui, en l'absence de SS2, est en mesure de synthétiser en partie les chaînes de taille intermédiaire de l'amylopectine (Szydlowski et al., 2011). Comme SS1, SS2 est présente dans des complexes hétéro-multimériques en association avec SS1 et BE2b dans les grains d'amidon de blé (Tetlow et al., 2008). L'isoforme SS2a chez l'orge est phosphorylée lorsqu'elle est liée au grain d'amidon, comme SS1 et GBSS (Ahmed et al., 2015).

#### *d. SS3*

SS3 fut découverte chez la pomme de terre en 1996 (Abel et al., 1996). Elle fut très rapidement soupçonnée d'être associée au grain d'amidon chez cette espèce (Larsson et al., 1996). Toutefois, l'analyse des protéines du grain par séquençage après séparation par SDS-PAGE n'a pas permis de discriminer entre SS3, SS1 et GBSS. Son association avec le grain d'amidon a par la suite été confirmée par protéomique chez le maïs (Grimaud et al., 2008). SS3 fut retrouvée chez d'autres espèces, sa séquence étant fortement conservée (Gao et al., 1998; Li et al., 1999b; Dian et al., 2005; Zhang et al., 2005). Il existe deux isoformes de SS3 chez le riz : SS3-1 et SS3-2 (Dian et al., 2005). La première est localisée aussi bien au niveau des chloroplastes que des amyloplastés, et la seconde uniquement au niveau des amyloplastés. Chez les autres espèces, une seule isoforme de SS3 est retrouvée, aussi bien au niveau des

chloroplastes que des amyloplastes (Gao et al., 1998; Li et al., 1999b; Dian et al., 2005; Zhang et al., 2005). SS3 est la plus grande isoforme de la famille des amidons synthases, avec un poids moléculaire de 139,1 kDa chez la pomme de terre (Abel et al., 1996). Elle possède un site catalytique en C-ter avec deux domaines GT, ainsi trois modules de liaison aux carbohydrates (Carbohydrate binding modules, CBMs) en N-ter (Figure 14). L'étude phénotypique de lignées de pomme de terre exprimant un ARN antisens de SS3 et de mutants KO *ss3* d'*Arabidopsis* a montré un impact modéré sur le contenu en amidon (Abel et al., 1996; Marshall et al., 1996; Zhang et al., 2005). Chez le mutant *ss3* d'*Arabidopsis*, la plante semble même accumuler plus d'amidon pendant le jour que le sauvage (Zhang et al., 2005). Il a donc été émis par les auteurs l'hypothèse que SS3 régulerait négativement l'activité des autres amidon synthases. La présence de CBMs 48 dans la partie N-ter de SS3 permet à cette dernière de se fixer à l'amidon (Senoura et al., 2007). Ce phénomène pourrait moduler les propriétés catalytiques des autres amidons synthases, expliquant sa possible fonction de régulateur (Senoura et al., 2007; Valdez et al., 2008). Les lignées mutantes du maïs (*du1*) et de *Chlamydomonas reinhardtii* (*sta3*) présentent un défaut d'accumulation d'amidon avec une diminution du contenu allant jusqu'à 70% par rapport aux références sauvages (Fontaine et al., 1993; Gao et al., 1998). La morphologie et la cristallinité des grains est aussi différente chez ces mutants. Ainsi, chez *Chlamydomonas*, on passe d'un type cristallin A chez le sauvage à un type cristallin B chez le mutant. La distribution des longueurs de chaînes de l'amylopectine est modifiée avec une plus grande proportion de longues chaînes (supérieures à DP40), et une diminution des glucanes de taille intermédiaire (Fontaine et al., 1993). Si le phénotype du simple mutant *ss3* était proche du sauvage chez *Arabidopsis*, le double mutant *ss2/ss3* a montré une croissance ralentie des plantes et un contenu en amidon considérablement réduit, plus que celui du simple mutant *ss2* (Zhang et al., 2008). L'activité de SS3 est partiellement redondante avec celle de SS2, étant capable de synthétiser les chaînes DP12 à DP25 en l'absence de cette dernière. Elle est aussi partiellement redondante avec l'activité de SS1 (Grzes et al., 2011). Par conséquent, SS3 couvrirait une plus large gamme de chaînes à allonger que SS1 et SS2, allant des chaînes courtes, type A et B1, aux chaînes plus longues, de type B2 et B3. En plus de son activité d'élongation, SS3 semble aussi être en capacité d'initier la synthèse d'un grain d'amidon chez *Arabidopsis* en l'absence de l'isoforme SS4 qui sera décrite par la suite (Szydlowski et al., 2009). SS3 a été retrouvée en complexe avec SS1 et BE2a dans l'amidon de maïs (Hennen-Bierwagen et al., 2008).

#### e. *SS4 et l'initiation de la synthèse du grain*

L'initiation de la synthèse d'un grain d'amidon est un phénomène encore mal compris. On sait aujourd'hui que cette initiation est dépendante, du moins chez *Arabidopsis thaliana*, d'une amidon-synthase et vraisemblablement d'un complexe hétéromultimérique d'initiation (Grzes et al., 2011; Seung et al., 2017; Seung et al., 2018; Vandromme et al., 2019). Ce complexe nécessite toutefois la présence préalable d'un glucane primordial d'une taille supérieure à DP10 (Seung et al., 2017). Il est probable que d'autres enzymes, supposément des amidons phosphorylases, agissent en amont pour allonger des sucres de faible DP tel que le maltose de manière à obtenir ce glucane primordial. Le siège de l'initiation dans les chloroplastes se situerait au niveau des plastoglobules (Gámez-Arjona et al., 2014). Ceux-ci sont des structures retrouvées au niveau de la membrane des thylakoïdes dans les chloroplastes. Le mécanisme d'initiation de la synthèse du grain d'amidon n'a toutefois que très peu été étudié dans le cas de l'amidon de réserve. Les seuls travaux connus portent sur le riz et se basent sur les résultats obtenus chez l'espèce modèle *Arabidopsis thaliana* (Toyosawa et al., 2016). SS4 est l'amidon synthase qui serait impliquée dans les mécanismes d'initiation de la synthèse du grain d'amidon plutôt que dans l'élongation des chaînes de l'amylopectine (Roldán et al., 2007). Elle fut décrite chez le riz en 2004 (Hirose and Terao, 2004). On retrouve deux isoformes de SS4 chez le riz : SS4-1 et SS4-2. La première est préférentiellement exprimée dans les organes de réserve et la seconde dans les tissus photosynthétiques (Hirose and Terao, 2004; Dian et al., 2005). Chez les autres plantes, on retrouve une seule isoforme de SS4 (Roldán et al., 2007; Leterrier et al., 2008; Pfister and Zeeman, 2016; Van Harselaar et al., 2017). La plupart des chloroplastes du mutant KO *ss4* d'*Arabidopsis thaliana* accumulent un unique gros grain d'amidon (quelque fois 2) de forme sphérique avec un hile plus relâché, alors qu'il y en a en moyenne 5 à 7 plus petits et plats dans les lignées sauvages (Roldán et al., 2007)<sup>138</sup>. Certains chloroplastes du mutant KO *ss4*- possèdent parfois deux grains d'amidon alors que d'autres en sont totalement exempts. Les plantes mutantes *ss4*- ont une croissance ralentie et accumulent des niveaux anormalement élevés d'ADP-Glc (Crumpton-Taylor et al., 2013; Ragel et al., 2013). La mutation au locus *SS4* ne semble toutefois pas avoir d'impact sur la structure de l'amylopectine (Roldán et al., 2007). Par ailleurs, alors que le triple mutant KO *ss1 ss2 ss3* est encore en mesure de synthétiser des grains d'amidon, plus petits mais en même nombre par plaste que le sauvage, le double mutant *ss3 ss4* en est incapable (où à de très rares exceptions et de façon stochastique dans certains chloroplastes) (Szydlowski et al., 2009). Cela montre l'implication de SS4, avec une redondance partielle de SS3, dans le processus d'initiation de la synthèse d'un grain d'amidon. Au niveau de l'albumen de riz, le double mutant *ss3-1 ss4-2* (ou *ssIIIa ssIVb*) est toujours en

mesure de produire de l'amidon, mais celui-ci est de forme sphérique au lieu d'être polyédral (Toyosawa et al., 2016). La présence d'autres isoformes de SS3 et SS4 chez le riz peut expliquer cette différence de phénotype avec *Arabidopsis* (Pfister and Zeeman, 2016). De plus, dans le cas du riz, c'est l'amidon de réserve au niveau des amyloplastés qui a été étudié alors que dans le cas d'*Arabidopsis*, il s'agissait de l'amidon transitoire au niveau des chloroplastes qui avait été étudié (Szydlowski et al., 2009; Toyosawa et al., 2016). Chez *Arabidopsis*, des études sur la localisation cellulaire de SS4 ont montré que celle-ci serait associée à la membrane des thylakoïdes, au niveau des plastoglobules, et colocaliserait avec les fibrillines FBN1a et FBN1b (Gámez-Arjona et al., 2014). Trois études plus récentes sur les interactants de SS4 d'*Arabidopsis* n'ont toutefois pas montré d'interaction avec ces fibrillines, mais avec 4 autres protéines : PTST2 ; PTST3 ; MAR binding Filament-like Protein 1 (MFP-1) et Protein involved in starch initiation 1 (PII1) (Seung et al., 2017; Seung et al., 2018; Vandromme et al., 2019). L'étude des domaines de SS4 a mis en évidence que c'est le domaine N-terminal de SS4 qui permet sa localisation au thylakoïde et qui contrôle la forme des grains chez *Arabidopsis*, alors que la Partie C-ter est responsable de l'activité d'initiation de la synthèse des grains et de l'interaction avec le complexe d'initiation, plus particulièrement avec PTST2 (Seung et al., 2017; Lu et al., 2018).

#### *f. SS5 et SS6*

Enfin, deux gènes dont la séquence est proche des amidons synthèses SS3 et SS4, SS5 et SS6, ont été décrits dans le génome de la pomme de terre (Van Harselaar et al., 2017). D'un point de vue phylogénique, les amidons synthèses se séparent en deux familles : GBSS, SS1 et SS2 forment une première famille, et SS3 et 4 forment une deuxième famille (Ball et al., 2015). SS5 et SS6 sont quant à elles phylogénétiquement affiliées à la deuxième famille (Abt et al., 2020; He et al., 2020; Sevestre et al., 2020). L'isoforme SS5 a aussi été trouvée chez *Arabidopsis* et chez le maïs (Deschamps et al., 2008; Liu et al., 2015; Abt et al., 2020). Cette amidon synthase est en mesure d'interagir avec elle-même et avec la protéine PII1, aussi appelé myosin-resembling chloroplast protein (MRC) (Abt et al., 2020). Sa mutation n'entraîne pas de différence dans la distribution de longueurs des chaînes, elle n'aurait donc pas d'activité d'élongation. Cette amidon-synthase est en mesure d'interagir avec elle-même et avec la protéine PII1. Son inactivation n'entraîne toutefois pas de différence de la distribution de la longueur des chaînes d'amylopectine, elle n'aurait donc pas d'activité d'élongation ce qui est cohérent avec l'absence de domaine GT1 et de motifs XXGGL dans SS5. Les mutants *ss5* ont un phénotype proche de celui du mutant *ss4*, avec une diminution du nombre de grains par

plaste et une augmentation de leur taille par rapport au sauvage. Néanmoins, les grains sont de forme lenticulaire dans le mutant *ss5* alors que ceux des mutants *ss4* sont sphériques. Lors de la double mutation *ss4/ss5*, le contenu en ADP-glucose est augmenté au sein des grains et la quantité de grains par plaste est fortement diminuée et ce de manière exacerbée par rapport au simple mutant *ss4* (Abt et al., 2020). A noter que les doubles mutants *ss4/ss5* sont encore capables, de manière sporadique, de former des grains d'amidon dans quelques chloroplastes. La présence de ces grains est probablement due à SS3, qui est en mesure d'initier la synthèse d'un grain d'amidon en absence de SS4 (Szydłowski et al., 2009). De par ces analyses, il semble donc que SS5 ait une activité partiellement redondante de celle de SS4 dans l'initiation de la synthèse des grains d'amidon chez *Arabidopsis*. Une autre hypothèse formulée sur le rôle du complexe SS5-MRC serait de protéger le malto-oligosaccharide primordial d'une dégradation amylolytique, prolongeant ainsi sa durée de vie et facilitant l'évènement d'initiation de la synthèse (Abt et al., 2020).

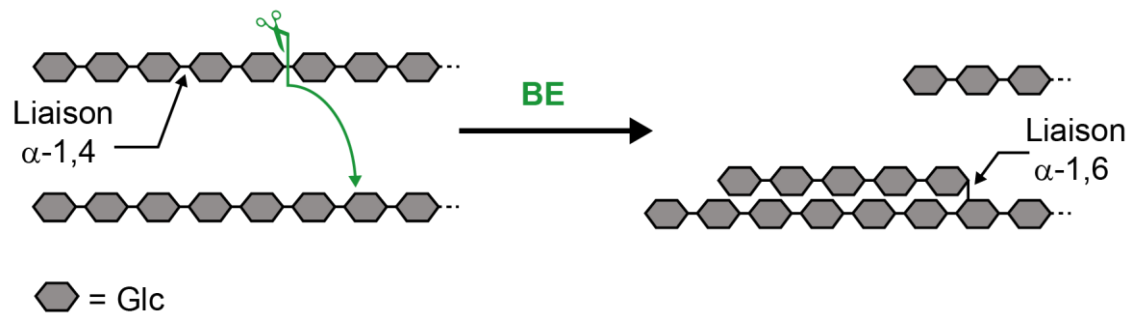
Des travaux sur la caractérisation de SS6 sont en cours chez le manioc (*Manihot esculenta*) et la pomme de terre (*Solanum tuberosum*) (communication personnelle François Sevestre, UGSF, CNRS UMR 8576, Villeneuve d'ascq, France)(He et al., 2020; Sevestre et al., 2020). Le préprint disponible sur Biorxiv concernant les amidons synthases de manioc montre que *MeSS6* n'est présente que chez les dicotylédones, sa séquence n'est toutefois par retrouvée chez la plante modèle *Arabidopsis* (He et al., 2020). Chez le manioc, les tests *in vitro* à partir d'une protéine recombinante n'ont pas montré d'activité pour *MeSS6*. Toutefois, les tests d'immunoprécipitations et de fluorescence par complémentation bimoléculaire (BiFC) montrent que *MeSS6* formerait un complexe avec l'AGPase, mais aussi avec GBSS et avec SS1, BE2, ISA1 et ISA2 (He et al., 2020). Ces données n'ont toutefois pas encore été révisée par les pairs et sont donc à considérer avec précaution.

### **3. Formation des points de branchements par les enzymes de branchements**

Les points de branchement sont formés par les enzymes de branchement (BEs). Elles créent les liaisons  $\alpha$ -1,6 par un processus de réorganisation structurale de glucanes préexistants. Ainsi l'activité de branchement consiste d'abord à cliver une liaison  $\alpha$ -1,4 d'un glucane et à transférer la partie en position non-réductrice du glucane clivé en  $\alpha$ -1,6 selon un processus inter- ou intramoléculaire (Figure 15) (Pfister and Zeeman, 2016). Les glucanes substrats des enzymes de branchement doivent avoir un degré de polymérisation (DP) supérieur ou égal à 12 (Nakamura et al., 2010). On distingue deux familles d'enzymes de branchement (Burton et al.,

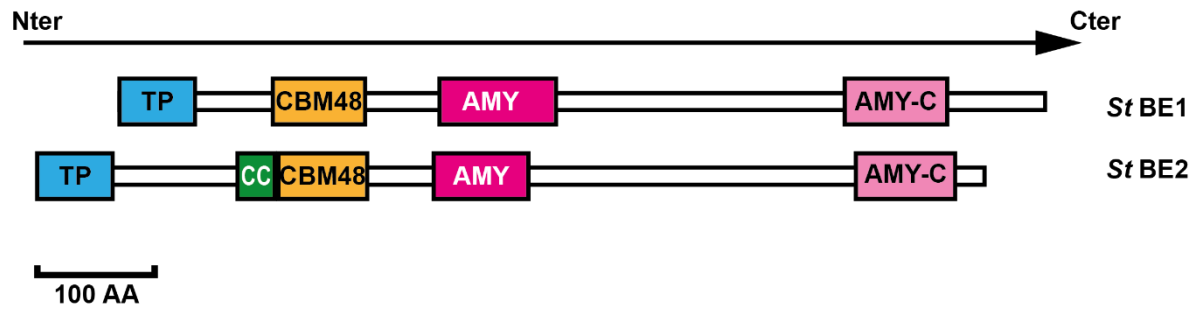


1995). La première famille est celle des enzymes de type I (BE1), et la famille des enzymes de type II (BE2). Cette classification se base sur la séquence primaire des enzymes, les BE2 ayant un domaine en N-terminal plus long que celui des BE1 (Burton et al., 1995). La nomenclature utilisée ici est celle des enzymes du maïs. Il existe toutefois plusieurs nomenclatures pour les enzymes de branchement, chaque espèce végétale ayant eu une nomenclature propre. Ainsi, si nous prenons l'exemple d'Arabidopsis, AtBE2 et AtBE3 sont, en suivant la nomenclature du maïs, toutes les deux des enzymes de type II (Dumez et al., 2006). Elles seront respectivement renommées BE2.2 et BE2.1 par la suite (Wychowski et al., 2017). AtBE1 n'appartient quant à elle ni à la famille des enzymes de type I, ni à celles de type II. Celle-ci appartiendrait à une 3<sup>ème</sup> famille qui existe aussi chez le peuplier, le riz et le maïs et qui sera nommée BE3 (Han et al., 2007; Yan et al., 2009). Les BE3 ne semblent toutefois pas avoir d'activité catalytique de branchement (Dumez et al., 2006).



**Figure 15. Formation d'un point de branchement (mécanisme intermoléculaire).** Les glucanes linéaires donneurs et accepteurs doivent nécessairement avoir un degré de polymérisation supérieur à 12. Chaque hexagone sur la figure représente une unité de glucose. La paire de ciseaux verte représente la liaison  $\alpha$ -1,4 qui sera clivée pour former le point de branchement. BE : Enzyme de branchement.

Les enzymes de branchements (BEs) de l'amidon partagent trois domaines conservés : un domaine N-ter avec un CBM48 permettant l'adhésion à l'amidon, un domaine catalytique central de type amylase, caractéristique de la famille des glycoside hydrolases 13 (GH13) de la classification CAZy, et un domaine C-ter commun à la famille des amylases (Figure 16) (Lombard et al., 2014).



**Figure 16. Représentation des domaines structuraux des enzymes de branchement de pomme de terre (*Solanum tuberosum*).** Les séquences BE1 et BE2 utilisées sont extraites de Uniprot (Accession : P30924 et Q9XGA8, respectivement). Les peptides d'adressage au plaste ont été identifiés grâce au logiciel en ligne ChloroP 1.1 (<http://www.cbs.dtu.dk/services/ChloroP/>) (Emanuelsson et al., 1999). Les régions coiled-coil ont été identifiées avec Paircoil2 (<http://cb.csail.mit.edu/cb/paircoil2/paircoil2.html>) (McDonnell et al., 2006). Les autres régions ont été identifiées avec SMART (<http://smart.embl-heidelberg.de/>) (Letunic and Bork, 2017). Les carrés roses et magentas représentent les domaines amyloses. Les rectangles orange représentent les modules de liaison aux carbohydrates, les rectangles verts (CC) représentent les régions coiled-coil permettant des interactions protéine-protéine. Les rectangles cyans représentent les peptides de transfert au chloroplaste. TP : peptide d'adressage au plaste. CBM48 : Module de liaison aux carbohydrates 48. AMY : amylose (classification CAZy) (Lombard et al., 2014). AMY-C : Amylose en C-terminus permettant le transfert des chaînes sur les points de branchement. *St* : *Solanum tuberosum*. AA : Acides aminés. On distingue deux familles d'enzymes de branchements (BEs) (Burton et al., 1995). La première famille est celle des enzymes de type I (BE1), et la famille des enzymes de type II (BE2). Cette classification se base sur la séquence primaire des enzymes, les BE2 ayant un domaine en N-terminal plus long que celui des BE1 (Burton et al., 1995). La nomenclature utilisée ici est celle des enzymes du maïs.

#### a. BE1

BE1 est exprimée dans les plastes, principalement dans les organes de réserve (Morell et al., 1997; Kim and Gultinan, 1999; Mutisya et al., 2003). Elle est associée au grain d'amidon chez le pois, le blé, le maïs, la pomme de terre et le riz (Denyer et al., 1993; Rahman et al., 1995; Larsson et al., 1996; Grimaud et al., 2008; Xing et al., 2016). Son poids moléculaire varie de 90 à 100 kDa selon les espèces, la BE1 du pois atteignant 105 kDa (Burton et al., 1995). Chez la pomme de terre, BE1 possède un poids moléculaire prédit de 99 kDa (Poulsen and Kreiberg, 1993). Les analyses de plants de maïs mutants *bel* ne montrent pas de modifications de la structure de l'amidon, aussi bien pour l'amidon transitoire des feuilles que l'amidon de réserve des graines (Blauth et al., 2002). En revanche, le mutant *bel* (nomenclature du maïs) du riz présente de petites modifications de la distribution de la longueur des chaînes de l'amylopectine en comparaison de la lignée sauvage (Nakamura, 2002; Satoh et al., 2003). Une diminution du nombre des chaînes de DP 12 - 21 et de DP > 37 fut observée chez les lignées mutantes, ainsi

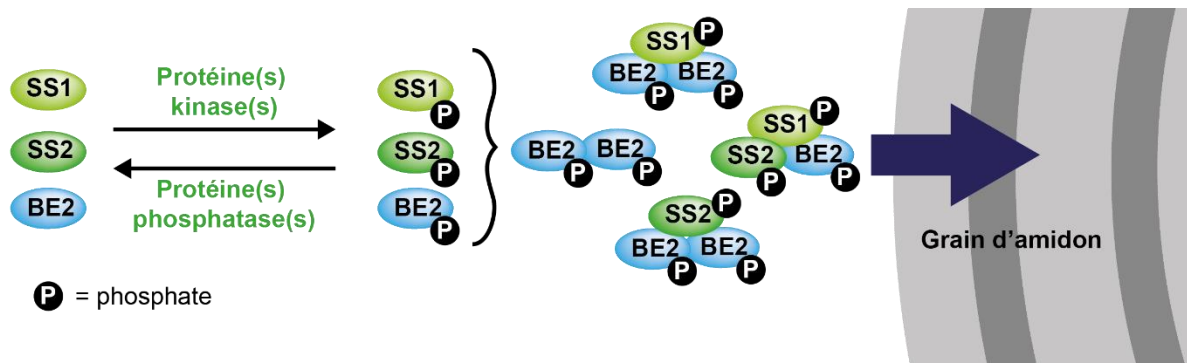
qu'une augmentation du nombre des chaînes de  $DP \leq 10$  et des  $DP 24 - 34$ . Les expériences de caractérisation de l'activité des BE1 sur des dextrans linéaires chez le maïs, le riz, le haricot rouge et la pomme de terre ont montré que cette isoforme est capable de transférer des chaînes de  $DP > 30$ , mais qu'elle transfère majoritairement des chaînes de  $DP 10 - 13$  (Takeda et al., 1993; Guan et al., 1997; Rydberg et al., 2001; Ito et al., 2004). De plus, BE1 semble avoir plus d'affinité pour l'amylose que pour l'amylopectine. Des tests plus récents sur la BE1 du riz ont montré sa capacité à transférer des chaînes de  $DP \geq 6$  et pouvant être  $> 15$  (mais avec une préférence pour les chaînes de  $DP 6 - 15$ ) à partir de chaînes d'amylopectine  $\geq DP12$  (Nakamura et al., 2010). Qui plus est, une autre étude suggère que BE1 est capable de transférer des chaînes déjà branchées (Sawada et al., 2014; Pfister and Zeeman, 2016). L'expression hétérologue des différentes formes de BE dans la levure a montré l'incapacité de BE1 à agir en absence des autres enzymes de branchement (Seo et al., 2002). Il serait donc possible que BE1 agisse en aval de BE2 dans la synthèse de l'amylopectine, permettant la synthèse des chaînes intermédiaires. Cette hypothèse fut aussi proposée dans une autre étude sur les mutants *bel* du riz (Satoh et al., 2003). Une récente étude chez des lignées de riz mutantes pour les différentes enzymes de branchement semble confirmer cette hypothèse (Sawada et al., 2018).

#### *b. BE2*

Il existe deux isoformes de BE2 : BE2a et BE2b, chez les monocotylédones telles que le riz et le maïs. Les dicotylédones telles que la pomme de terre ne possèdent qu'une seule enzyme de branchement (Fisher et al., 1996; Rahman et al., 2001). *Arabidopsis thaliana*, une dicotylédone elle aussi, fait toutefois exception à la règle puisque son génome contient deux BE2 : *BE2.1* et *BE2.2* (Wychowski et al., 2017). Chez les monocotylédones, BE2a est principalement présente dans les chloroplastes des tissus photosynthétiques, et BE2b au niveau des amyloplastides des tissus de réserve (Dang and Boyer, 1988, 1989; Gao et al., 1996). On observe toutefois dans l'albumen de blé une plus forte expression de BE2a que de BE2b (Regina et al., 2005). Les BE2 ont elles aussi été retrouvées associées aux grains d'amidon de réserve chez le blé (BE2a et BE2b), l'orge (BE2b seulement) le maïs (BE2a et BE2b) et le riz (BE2b seulement) (Borén et al., 2004; Regina et al., 2005; Umemoto and Aoki, 2005; Grimaud et al., 2008). L'isoforme BE2 fut aussi retrouvée dans l'amidon des tubercules de pomme de terre (Larsson et al., 1996). Le phénotype des mutants *be2* est beaucoup plus marqué que celui des mutants *be1*. En effet, chez les mutants *be2a* de maïs, la structure et la composition de l'amidon transitoire sont altérées, avec une augmentation des longues chaînes de l'amylopectine et une plus grande proportion en amylose (Blauth et al., 2001; Yandea-Nelson et al., 2011).

En revanche, l'amidon de réserve chez ces mutants ne présente pas d'altération remarquable. L'amidon de réserve des mutants *be2b* (*amylose extender*) de maïs et de riz, présente une augmentation de la proportion d'amylose et des chaînes d'amylopectine plus longues (Boyer and Preiss, 1978; Mizuno et al., 1993; Nishi et al., 2001). Dans le cas du blé, le simple mutant *be2b* est peu altéré en ce qui concerne la structure de l'amidon de réserve, au contraire du double mutant *be2a be2b* où le taux d'amylose excède 70% (Regina et al., 2006). Chez la pomme de terre, où une seule BE2 est présente, la caractérisation *in vitro* des activités de branchement a montré que BE2 a comme substrat préférentiel des chaînes plus courtes que BE1, aux alentours de DP6 (Rydberg et al., 2001). De plus, la surexpression de BE2 chez la pomme de terre conduit à une altération de la structure de l'amylopectine avec une forte augmentation des chaînes courtes ( $DP \leq 12$ ), notamment de DP 6 (Brummell et al., 2015). La caractérisation des enzymes de branchement d'autres espèces comme le blé, le riz et le maïs suggère aussi que BE2 agit principalement sur les chaînes courtes (Takeda et al., 1993; Guan et al., 1997; Rydberg et al., 2001; Ito et al., 2004; Nakamura et al., 2010). BE2 est aussi beaucoup plus active que BE1 sur l'amylopectine (Takeda et al., 1993). Tout comme BE1, BE2 a besoin d'un substrat accepteur suffisamment long ( $DP \geq 12$ ), laissant présager que cette isoforme reconnaît une structure sous forme d'hélice (Nakamura et al., 2010).

Chez *Arabidopsis*, il existe deux gènes codant des BEs de type II : BE2.1 et BE2.2 (Fisher et al., 1996). Les analyses des simples mutants KO *be2.1* et *be2.2* ne montrent pas de modification majeure de l'accumulation et de la structure de l'amidon, alors que le double mutant est incapable de synthétiser de l'amidon (Dumez et al., 2006). Il semble donc que ces deux enzymes aient un rôle redondant. Les études chez le blé indiquent que les enzymes de branchement sont régulées par un processus post-traductionnel de phosphorylation / déphosphorylation (Tetlow et al., 2004). En effet, les BEs de type II ayant subi une déphosphorylation voient leur activité chuter drastiquement. La même expérience avec des BEs de type I n'entraîne pas de changement d'activité. Ce type de phosphorylation sur les BE2 a aussi été observé chez le maïs et l'orge (Liu et al., 2009; Ahmed et al., 2015). Cette phosphorylation serait aussi responsable de la formation de complexes hétéromultimériques dont il a été possible de démontrer la présence au sein du grain d'amidon (Figure 17) (Tetlow et al., 2004; Liu et al., 2009; Ahmed et al., 2015). Les BE2 phosphorylées formeraient un complexe avec les amidon-synthases SS1 et SS2 pour interagir avec le grain d'amidon (Tetlow et al., 2008).



**Figure 17. Complexes d'association formés entre les enzymes de branchement et les amidon-synthases et association au grain d'amidon.** SS : amidon-synthase. BE : enzyme de branchement. Ces complexes ont été identifiés chez le blé par Tetlow et coll. (2008) (Tetlow et al., 2008). La phosphorylation de BE2 pour l'association au grain a été retrouvée chez le blé, le maïs et l'orge (Tetlow et al., 2008; Liu et al., 2009; Ahmed et al., 2015). La phosphorylation de SS1 et SS2 pour l'association au grain a été observée chez l'orge (Ahmed et al., 2015).

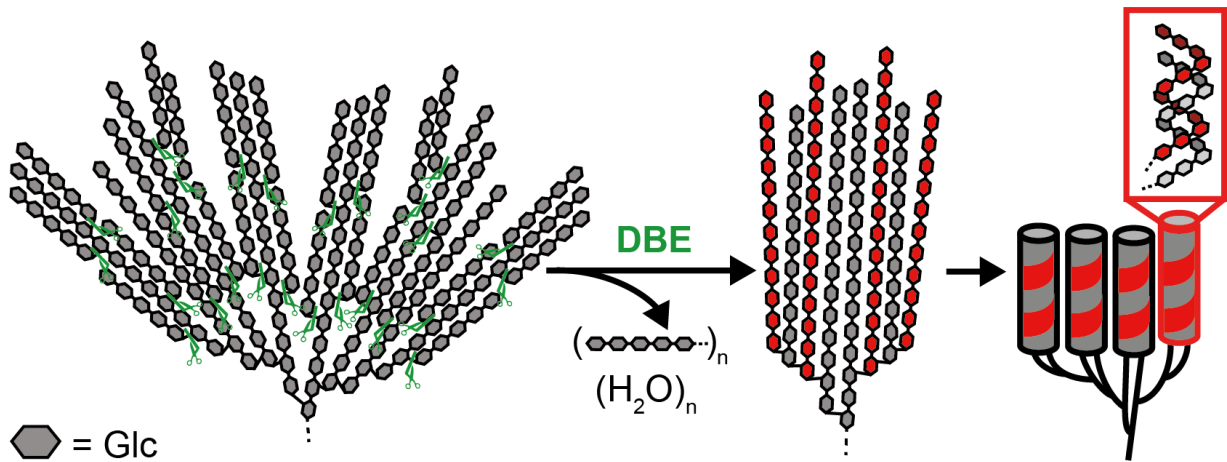
#### c. *BE3*

Une troisième isoforme, indépendante génétiquement de BE1 et des BE2, a été retrouvée dans plusieurs génomes (Han et al., 2007). Les BE3 ne semblent pas posséder d'activité de branchement. Leur inactivation chez *Arabidopsis* ne modifie pas la structure de l'amidon (Dumez et al., 2006). De plus, elles ne partagent que peu de similarités avec les BEs de type I et II. Une étude plus récente semble toutefois montrer que les mutants KO *be3* d'*Arabidopsis* possèdent un défaut d'embryogénèse, et donc que cette protéine serait cruciale pour le développement de la plante (Wang et al., 2010).

#### 4. Rôle des enzymes de débranchement dans la stabilisation de la cristallinité.

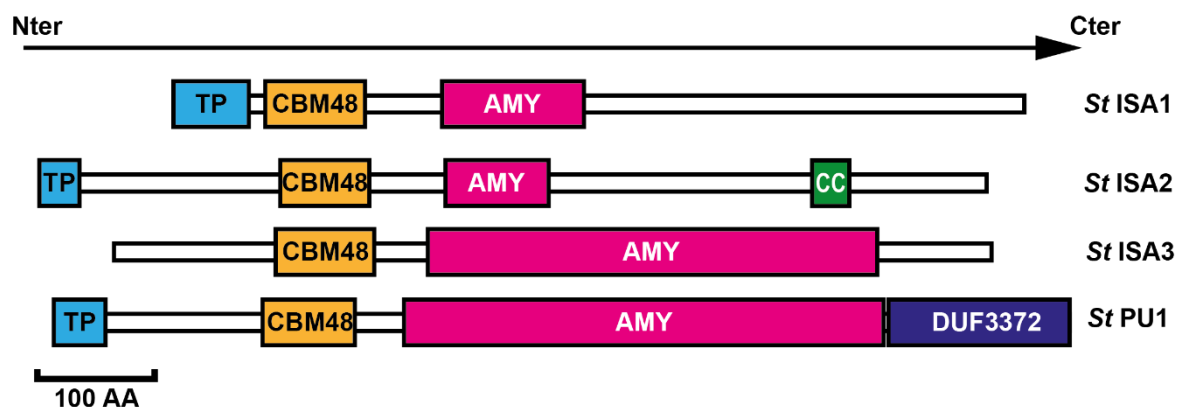
Une accumulation non contrôlée des points de branchement entrave la formation normale de la grappe d'amylopectine et par conséquent sa cristallisation et la formation du grain d'amidon. L'amidon est partiellement remplacé par un polymère en tout point semblable au glycogène et dénommé phytoglycogène. Il est nécessaire que les points de branchement soient concentrés en certaines zones de la grappe d'amylopectine pour former les lamelles amorphes et cristallines décrites en partie II (Ball et al., 1996). Certaines enzymes de débranchement vont agir pendant la synthèse de l'amidon pour retirer les points de branchement qui seraient situés en dehors de la lamelle amorphe et permettre la cristallisation des chaînes dans la lamelle cristalline. Leur activité consiste à hydrolyser des liaisons  $\alpha$ -1,6 libérant un  $\alpha$ -1,4 glucane et une molécule d'eau (Figure 18). Deux enzymes de débranchement, les Isoamylases 1 et 2 (ISA1 et ISA2), agissant en complexe, sont impliquées dans cette activité de débranchement pendant la synthèse: ISA1

est la seule des deux enzymes à posséder une activité catalytique, mais ISA2 (non catalytique) est nécessaire à la formation du complexe chez la plupart des espèces (Hussain et al., 2003).



**Figure 18. Action des enzymes de débranchement.** Chaque hexagone représente une unité de glucose. Chaque cylindre gris et rouge représente une double hélice formée par 2 glucanes linéaires. Les paires de ciseaux verts représentent les points où les enzymes de débranchement clivent la chaîne branchée en  $\alpha$ -1,6. Une chaîne sur deux a été colorée en rouge de manière à mieux distinguer les deux chaînes dans la double hélice. DBE : Enzyme de débranchement.

On distingue deux familles d'enzymes de débranchement, les isoamylases (ISA) et les limite-dextrinases, aussi appelées pullulanases (PU) (Pfister and Zeeman, 2016). Toutes les enzymes de débranchement possèdent un domaine amyrase d'hydrolyse des liaisons  $\alpha$ -1,6, ainsi qu'un module de liaison aux carbohydrates pour s'accrocher au grain (Figure 19).



**Figure 19. Représentation des domaines structuraux des enzymes de débranchement de pomme de terre (*Solanum tuberosum* St).** Les séquences de ISA1, 2 et 3 sont extraites de Uniprot (Accession : Q84YG7, Q84YG6 et Q84YG5, respectivement). La séquence de PU1 est une séquence prédite à partir de l'ADNc (Accession : XP\_006361707.1). Les peptides d'adressage au plaste (TP) sont déterminés avec ChloroP 1.1 (<http://www.cbs.dtu.dk/services/ChloroP/>) (Emanuelsson et al., 1999). Les régions coiled-coil (CC) ont été identifiées avec Paircoil2 (<http://cb.csail.mit.edu/cb/paircoil2/paircoil2.html>) (McDonnell et al., 2006). Les autres régions

ont été identifiées avec SMART (<http://smart.embl-heidelberg.de/>)(Letunic and Bork, 2017). Les carrés magentas représentent les domaines amylases. Les rectangles orange (CBM48) représentent les modules de liaisons aux carbohydrates, les rectangles verts représentent les régions coiled-coil permettant des interactions protéine-protéine. Les rectangles cyans représentent les peptides de transfert au chloroplaste. Le rectangle bleu représente le domaine non caractérisé DUF3372, spécifique des pullulanases bactériennes et végétales. AMY : amylase. DUF : domaine fonctionnel non caractérisé. ISA : Isoamylase. PU1 : pullulanase. AA : Acides aminées.

#### a. *ISA1*

L'isoamylase 1 (ISA1) est une enzyme de débranchement impliquée dans la synthèse de l'amylopectine. Elle fut découverte lors de l'étude de la mutation au niveau du locus *sugary 1* (*su1*) chez le maïs. Ce mutant synthétisait moins d'amylopectine mais accumulait en contrepartie un polyglucane soluble et branché semblable au glycogène, dénommé alors phytoglycogène (Black et al., 1966). En 1995, il fut découvert que la mutation *su1* entraînait la perte de l'activité d'une enzyme de débranchement de type isoamylase (James et al., 1995). Entre temps, une enzyme de débranchement de type isoamylase fut retrouvée chez la pomme de terre (Ishizaki et al., 1983). Plus tard, des protéines homologues à cette enzyme de débranchement furent identifiées chez *Chlamydomonas reinhardtii*, le riz, l'orge et *Arabidopsis* (Mouille et al., 1996; Zeeman et al., 1998b; Fujita et al., 1999; Kubo et al., 1999; Dauvillée et al., 2001a; Burton et al., 2002). Cette enzyme fut alors appelée DBE1, pour « Debranching enzyme 1 », puis ISA1, pour « Isoamylase 1 ». Chez *Arabidopsis thaliana*, la mutation *isa1* entraîne une réduction de 80% de la quantité d'amidon transitoire en fin de jour par rapport au sauvage et une accumulation massive de phytoglycogène (Wattebled et al., 2005). De plus, les grains d'amidon mutants possèdent une taille fortement réduite par rapport au sauvage. Dans les tubercules de pommes de terre mutantes *isa1*, la réduction de la quantité d'amidon est peu marquée (Bustos et al., 2004). L'analyse de la distribution de la longueur des chaînes de l'amylopectine des mutants *isa1* montre un enrichissement de chaînes courtes, allant de DP5 à DP13 selon les espèces végétales (Dinges et al., 2001; Burton et al., 2002; Wong et al., 2003; Wattebled et al., 2005) montrant l'impact direct de l'activité de cette enzyme sur la structure même de l'amylopectine.

La plupart des analyses en conditions natives ont montré que la protéine purifiée avait un poids moléculaire de 540 kDa, alors que les prédictions à partir de la séquence peptidique indiquent une masse de 80 kDa (Ishizaki et al., 1983; Fujita et al., 1999). Cela suggère que l'enzyme agit en complexe multimérique, probablement sous forme d'hexamères.

## b. ISA2

En 1991, deux isoformes d'isoamylases, différentes de la pullulanase, sont découvertes chez le maïs (Doehlert and Knutson, 1991). De plus, chez *Chlamydomonas reinhardtii*, une autre mutation, *sta8*, entraîne un phénotype similaire à celui des mutants *sta7* pour l'enzyme ISA1, suggérant la présence d'un autre composant responsable du débranchement de l'amylopectine, et que le multimère incluant ISA1 pourrait être en fait un hétéromultimère (Doehlert and Knutson, 1991; Dauvillée et al., 2001b; Dauvillée et al., 2001a). La protéine associée au gène *STA8*, au vu de sa séquence prédite, serait une isoforme de ISA1 mais dépourvue d'activité catalytique (Zeeman et al., 1998b; Dauvillée et al., 2001a). En 2003, deux homologues d'ISA1 ont été mis en évidence chez la pomme de terre : ISA2 et ISA3 (Hussain et al., 2003). ISA2 sera par la suite retrouvée chez le riz (Utsumi and Nakamura, 2006). Cette protéine d'environ 94 kDa ne possède pas de domaine catalytique. Les résidus d'aspartate (E) et de glutamate (D), présents au niveau du site catalytique de ISA1, sont absents chez ISA2 (Hussain et al., 2003). Il fut aussi mis en évidence que ISA1 et ISA2 forment un complexe hétéromultimérique, nommé Iso1 (Hussain et al., 2003; Utsumi and Nakamura, 2006). Chez le riz, deux formes de complexes sont présentes, un homomultimère composé seulement de ISA1 et un hétéromultimère composé de ISA1 et ISA2 (Utsumi and Nakamura, 2006). De même, chez le maïs, deux hétéromultimères de ISA1 et ISA2 ont été retrouvés, ainsi qu'un homomultimère de ISA1 (Kubo et al., 2010; Facon et al., 2013). Chez la pomme de terre, ainsi que chez *Arabidopsis* ; seule la forme hétéromultimérique contenant ISA1 et ISA2 fut retrouvée (Hussain et al., 2003; Delatte et al., 2005; Wattebled et al., 2005). Bien que n'ayant pas d'activité catalytique, ISA2 est nécessaire pour le bon fonctionnement de ISA1 dans certaines espèces végétales. Elle n'est toutefois pas nécessaire chez le riz et le maïs où les mutants *isa2* ont un amidon dont la structure proche de celle du sauvage (Kubo et al., 2010; Utsumi et al., 2011). De plus, l'insertion d'une ISA1 de maïs chez le double mutant *isa1 isa2* d'*Arabidopsis* permet le retour à un phénotype proche de celui du sauvage (Facon et al., 2013).

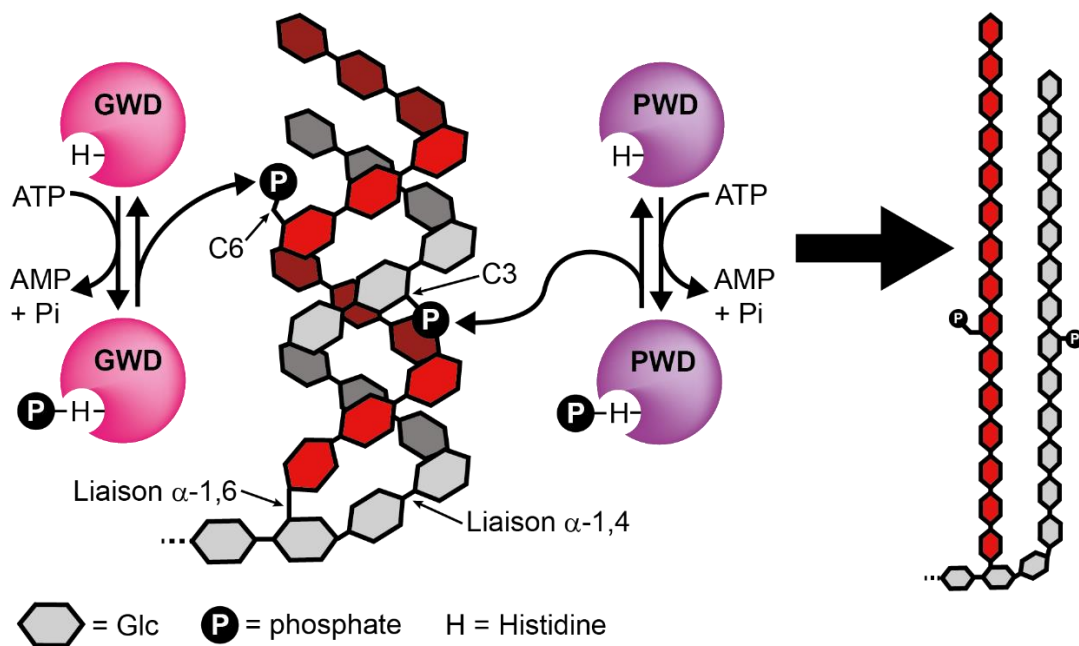
## IV. Dégradation de l'amidon

### 1. Décompaction de l'amylopectine par les dikinases de l'amidon

La structure semi-cristalline de l'amylopectine la rend plus résistante à la dégradation enzymatique. Il est donc nécessaire de décompacter les glucanes cristallins de la grappe d'amylopectine pour permettre aux enzymes de dégradation d'agir. Pour décompacter ces chaînes, certains résidus de glucose de l'amidon sont phosphorylés sur les carbones C3 et C6



(Figure 20). Cette phosphorylation est catalysée par des enzymes de la famille des dikinases, la Glucan Water Dikinase (GWD) et la Phosphoglucan Water Dikinase (PWD) (Ritte et al., 2002; Baunsgaard et al., 2005; Kötting et al., 2005). L'activité des dikinases a lieu en deux temps (Mikkelsen and Blennow, 2005). Dans un premier temps, la dikinase est activé par l'ATP : le groupement phosphate en position  $\gamma$  est transféré sur une molécule d'eau alors que le groupement phosphate en position  $\beta$  est quant à lui transféré sur un résidu d'histidine (H) du site catalytique de la dikinase. Cette première réaction libère ainsi de l'AMP et du phosphate inorganique. Dans un deuxième temps, le phosphate du résidu H est lui-même transféré sur le carbone C6 d'un résidu de glucose du glucane cible par GWD ou en position C3 par PWD. Bien qu'historiquement décrite comme impliquée dans la dégradation, il semble que la phosphorylation en position C6 ne déstabilise pas les glucanes organisés en double hélice (Blennow and Engelsen, 2010). Elle est toutefois nécessaire pour permettre l'activité de PWD et la phosphorylation en C3 qui est, quant à elle, impliquée directement dans la décompaction des doubles hélices de glucanes de l'amylopectine.



**Figure 20. Phosphorylation de l'amidon.** Chaque hexagone représente une unité de glucose. Une des deux chaînes de glucanes est colorée en rouge de manière à mieux distinguer la double hélice. GWD : Glucan Water Dikinase ; PWD : Phosphoglucan Water Dikinase. Adapté de Mahlow et coll. (2016) (Mahlow et al., 2016).

*a. GWD*

Les dikinases sont les enzymes responsables de la phosphorylation de certains résidus de glucose de l'amidon. Elles furent rapidement considérées comme des enzymes impliquées dans

la dégradation de l'amidon, car les premiers mutants pour l'une de ces protéines entraînaient un excès d'amidon dans les plastes. Ce mutant fut appelé *sex1*, pour « Starch Excess 1 » du fait de son phénotype (Caspar et al., 1991). Des analyses complémentaires montrèrent que les mutants *sex1* chez *Arabidopsis* étaient défectueux pour une protéine homologue à la protéine R1, qui fut retrouvée séquestrée dans le grain d'amidon de pomme de terre (Lorberth et al., 1998; Yu et al., 2001a). Elle fut aussi retrouvée dans des extraits de feuilles d'*Arabidopsis* (Yu et al., 2001a). Cette protéine R1 jouerait un rôle dans la dégradation de l'amidon et serait impliquée dans le sucrage à froid chez la pomme de terre (Lorberth et al., 1998). Les analyses des mutants *sex1* ont montré que ceux-ci avaient une distribution de longueur de chaîne inchangée, mais que le contenu en amylose était significativement augmenté et que le contenu en phosphoesters était fortement diminué voire nul (Yu et al., 2001a). Les études de caractérisations de la protéine R1, menées à partir d'amidon de pomme de terre, montrèrent que cette protéine est une «  $\alpha$ -Glucan water dikinase », une enzyme capable de phosphoryler l'amidon à partir de l'ATP, et produisant comme résidu de réaction de l'AMP et du Pi (Ritte et al., 2002). Le phosphate en position  $\beta$  de l'ATP étant d'abord transféré sur un résidu d'histidine (His) de GWD, puis sur une unité de glucose du glucane substrat (Ritte et al., 2002). La protéine, sous forme native, est retrouvée sous la forme d'homodimères (Mikkelsen et al., 2004). La protéine, d'une taille de 156,6 kDa chez *Arabidopsis*, possède un peptide d'adressage au plaste, un CBM, un domaine de liaison aux nucléotides et un domaine phospho-histidine (Figure 21) (Yu et al., 2001a). Suite à ces études, la protéine R1 fut renommée GWD, pour « Glucan Water Dikinase ». Les premières analyses de l'activité de GWD ont montré que cette enzyme phosphorylait les carbones C3 et C6 de certains résidus de glucose de l'amidon (Ritte et al., 2002). Des analyses plus poussées ont toutefois révélé que GWD phosphorylait uniquement l'amidon en position C6, la supposée phosphorylation en position C3 étant due à la co-élution du Glucose-3-phosphate (Glc-3-P) et du Maltose-6-phosphate (Mal-6-P) en chromatographie liquide haute performance échangeuse d'anion (HPAEC) (Ritte et al., 2006)

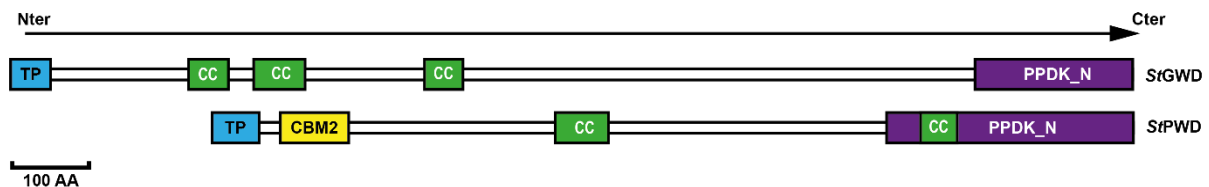
#### *b. PWD*

Peu après la caractérisation de la fonction de GWD, une deuxième isoforme de dikinase, la « Phosphoglucan water dikinase » (PWD), fut découverte (Baunsgaard et al., 2005; Kötting et al., 2005). La protéine fut identifiée par SDS-PAGE des protéines de surface des grains d'amidon phosphorylés et par la recherche de séquences homologues à GWD et leur analyse phylogénétique (Yu et al., 2001a; Mikkelsen et al., 2004; Baunsgaard et al., 2005). Chez *Arabidopsis*, PWD est localisée dans les chloroplastes et se lie au grain d'amidon durant la

dégradation (Kötting et al., 2005). La protéine d'*Arabidopsis* a une taille de 131,3 kDa. Tout comme GWD, PWD possède un CBM, un domaine de liaison aux nucléotides et un domaine phospho-histidine (Figure 21). Elle utilise le même mécanisme que GWD pour phosphoryler l'amidon : le phosphate en position  $\beta$  de l'ATP étant d'abord transféré sur un résidu d'histidine (His), puis sur une unité de glucose de l'amidon (Mikkelsen et al., 2004; Kötting et al., 2005). Tout comme les mutants *gwd-*, les mutants *pwd-* suraccumulent l'amidon. Les tests enzymatiques *in vitro* réalisés sur de l'amidon non phosphorylé issu de mutants *sex1* ont montré que la protéine PWD est incapable de phosphoryler l'amidon dépourvu de phosphates, mais que cette réaction peut avoir lieu lorsque l'amidon est préalablement incubé en présence de GWD. PWD ne phosphoryle donc que des phosphoglucanes (Kötting et al., 2005). De plus, les substrats préférentiels de PWD semblent être les longues chaînes d'amylopectine pré-phosphorylées par GWD (Baunsgaard et al., 2005). Contrairement à GWD qui phosphoryle en position C6, PWD phosphoryle en position C3 (Baunsgaard et al., 2005; Kötting et al., 2005). L'activité kinase de PWD est directement impliquée dans la décompaction de l'amylopectine. La phosphorylation de l'amidon, aussi bien en C3 et en C6, serait un moyen de déstabiliser les hélices de l'amylopectine pour permettre aux enzymes de dégradation d'accéder aux chaînes d'amylopectine (Ritte et al., 2006). Les mesures de cristallinité menées sur les doubles hélices d'amidon phosphorylées en position C3 et C6 ont montré que la phosphorylation en position C3 était bien plus efficace pour perturber la structure de la double hélice d'amylopectine (Ritte et al., 2006; Blennow and Engelsen, 2010). En effet, le groupement phosphate en position C6 s'aligne bien dans la rainure de la double hélice, et ne semble pas pouvoir permettre la décompaction de la double-hélice. Le phosphate C3, quant à lui, génère un conflit de structure dans la double hélice. Ce conflit de structure ne peut être résolu qu'en introduisant un défaut dans la structure de la double hélice, menant à une amorphisation locale de l'amylopectine (Blennow and Engelsen, 2010). La phosphorylation en C6 étant toutefois nécessaire pour l'action de PWD, GWD est donc aussi fortement liée à la régulation de la dégradation.

Chez *Arabidopsis*, une autre dikinase, proche de GWD, a été trouvée et nommée GWD2 (Baunsgaard et al., 2005). Elle ne fut toutefois retrouvée que chez certaines espèces de *Brassica* et *Citrus*, contrairement à GWD et PWD qui sont fortement conservées à travers le règne végétal (Glaring et al., 2007). GWD2 est une enzyme extra-plastidiale, probablement localisée dans le cytosol, étant capable de phosphoryler des glucanes sans phosphorylation préalable. Elle n'est toutefois pas nécessaire pour la dégradation correcte de l'amidon. Les auteurs ont émis l'hypothèse que cette enzyme est impliquée dans la dégradation de l'amidon ou de

structures proches de l'amidon durant la sénescence des feuilles (Glaring et al., 2007). Une étude plus récente réalisée sur des mutants d'*Arabidopsis* pour la protéine GWD2 a montré que cette dernière aurait un rôle dans le développement de la graine (Pirone et al., 2017).

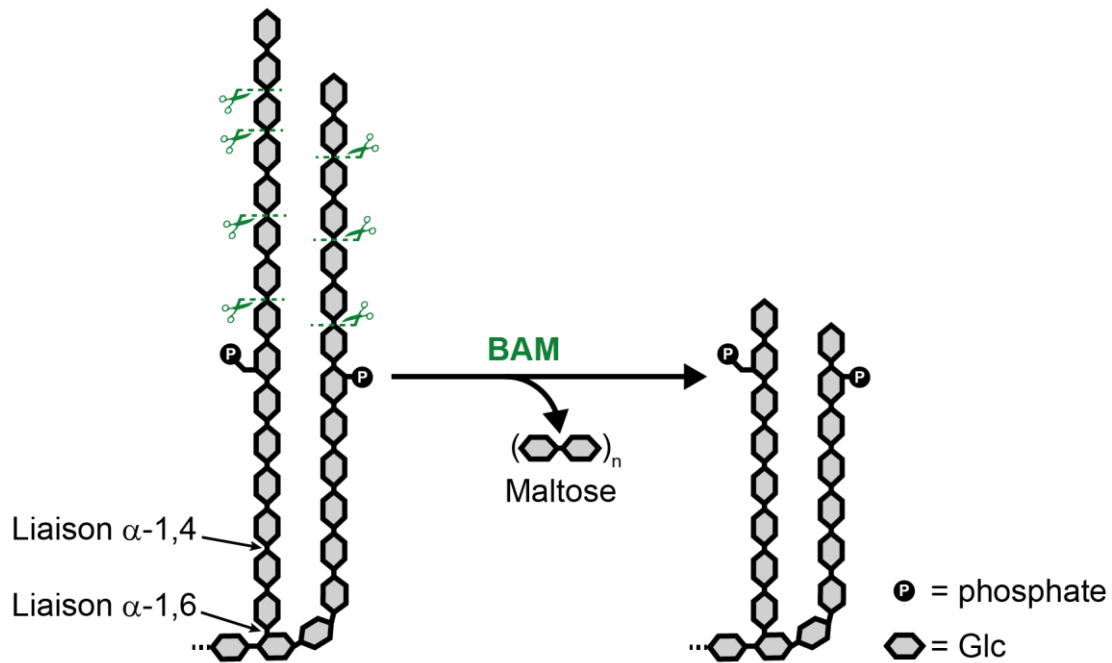


**Figure 21. Représentation des domaines structuraux des dikinases de pomme de terre (*Solanum tuberosum*, *St*).** Les séquences de GWD et PWD sont extraites du PGSC\_DM\_v3.4\_pep\_non-redundant (Avril 2017, 52.570 entrées) et annotées par Van Harselaar et coll. (2017) (Xu et al., 2011; Van Harselaar et al., 2017). Les peptides d'adressage au plaste (TP en cyan) ont été repérés avec ChloroP 1.1 (<http://www.cbs.dtu.dk/services/ChloroP/>) (Emanuelsson et al., 1999). Les régions coiled-coil (CC en vert) ont été identifiées avec Paircoil2 (<http://cb.csail.mit.edu/cb/paircoil2/paircoil2.html>) (McDonnell et al., 2006). Les autres régions ont été identifiées avec SMART (<http://smart.embl-heidelberg.de/>) (Letunic and Bork, 2017). Les rectangles violets représentent des domaines PPK, permettant le transfert de phosphate sur l'amidon à partir de l'ATP. Le rectangle jaune représente un module de liaison aux carbohydrates. PPK\_N : Pyruvate phosphate dikinase, région N-terminale. CBM2 : Carbohydrate binding module 2. AA : Acides aminés.

## 2. Dégradation des chaînes linéaires par les $\beta$ -amylases

On distingue parmi les amylases impliquées dans l'hydrolyse des liaisons  $\alpha$ -1,4 de l'amidon les  $\alpha$ -amylases (AMYS) et les  $\beta$ -amylases (BAMs). Les AMYS sont des endo-amylases en mesure d'hydrolyser l'amidon pour former des malto-oligosaccharides de tailles diverses. D'après la prédiction de domaine, seule AMY3 parmi les AMYS est adressée au plaste, et semble impliquée dans la dégradation de l'amidon chez *Arabidopsis* (Yu et al., 2005). La présence de cette dernière a été confirmée par Western blot sur les protéines de chloroplastes. Toutefois, le mutant *amy3* d'*Arabidopsis* possède un phénotype sauvage, montrant une faible implication de ces enzymes dans la dégradation. Il semble cependant que AMY3 agisse en synergie avec BAM1 lors du processus de dégradation (Thalmann et al., 2016). Les chaînes linéaires sont dégradées en grande majorité par les  $\beta$ -amylases (BAM). Celles-ci clivent les liaisons  $\alpha$ -1,4 par hydrolyse et libèrent du maltose (Figure 22). Ces enzymes agissent depuis les extrémités non-réductrices des chaînes linéaires, qu'elles vont cliver peu à peu. Les résidus de maltoses ainsi formés vont ensuite être exportés au cytosol via le transporteur MEX1 (Niittylä

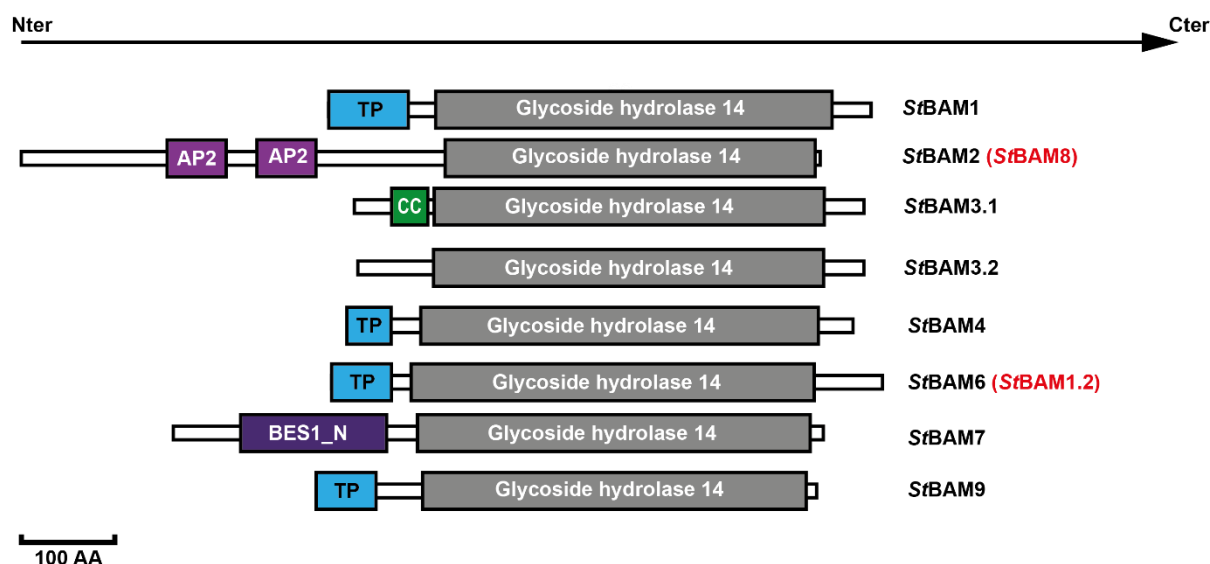
et al., 2004). Les BAMs sont toutefois incapables d’outrepasser un point de branchement ou un résidu de glucose phosphorylé. Elles s’arrêtent à 2 ou 3 résidus d’un point de branchement, et à 0 ou 1 résidu d’un phosphoester (Takeda, 1982; Delatte et al., 2006; Kötting et al., 2009). Les molécules branchées résultantes de la digestion par les BAMs sont appelées bêta-dextrines limites.



**Figure 22. Dégradation des chaînes linéaires par les  $\beta$ -amylases.** Chaque hexagone représente une unité de glucose. Les cercles noirs avec un P blanc à l’intérieur représentent les résidus de phosphate. Les paires de ciseaux verts représentent les liaisons  $\alpha$ -1,4 qui seront hydrolysées par les BAM.

Les BAMs sont des exo-amylases qui hydrolysent les chaînes d’amylose et d’amylopectine depuis leur extrémité non-réductrice pour libérer du maltose. L’inactivation de BAM1 chez la pomme de terre entraîne une diminution de la dégradation de l’amidon dans les feuilles (Scheidig et al., 2002). La caractérisation enzymatique *in vitro* de la protéine recombinante exprimée chez *E. coli*, a conduit à la production de maltose lorsqu’elle est mise en contact avec de l’amylopectine et de l’amidon de pomme de terre. Les BAMs sont des protéines dont le poids moléculaire varie de 60 à 80 kDa selon les isoformes (Fulton et al., 2008). Chez Arabidopsis, 9 gènes codent pour des BAMs. Six gènes ont été identifiés comme codant des BAMs localisées dans la fraction du plaste par fusion avec un GFP : BAM1, BAM2, BAM3, BAM4, BAM6 et BAM9 (Fulton et al., 2008; Stettler et al., 2009). BAM1, BAM3 et BAM9 sont phylogénétiquement proches et font partie de la même sous famille : BAM I (Monroe and Storm, 2018; Thalmann et al., 2019). Les autres BAMs font partie de la deuxième sous-famille :

BAM II. Seules BAM1, 2, 3, et 6 sont actives *in vitro* et impliquées dans la dégradation de l'amidon et leur activité est redondante (Fulton et al., 2008; Monroe et al., 2017; Monroe and Storm, 2018; Monroe, 2020). BAM4 et 9 seraient aussi impliquées dans la dégradation, mais en tant que régulateurs du fait de l'absence de domaine catalytique actif chez ces enzymes, du moins chez *Arabidopsis* (Fulton et al., 2008; Monroe and Storm, 2018). BAM5 est une enzyme active localisée au niveau du cytosol, plus précisément au niveau du phloème dans le cas d'*Arabidopsis thaliana* (Wang et al., 1995). BAM7 et 8 ont été localisées, après marquage à la GFP, au niveau du noyau et seraient des régulateurs de l'expression de certains gènes (Reinhold et al., 2011; Monroe and Storm, 2018). Chez la pomme de terre 10 séquences de BAMs ont été annotées : *StBAM1*, *StBAM2*, *StBAM3.1*, *StBAM3.2*, *StBAM4*, *StBAM6.1*, *StBAM6.2*, *StBAM6.3*, *StBAM7* et *StBAM9* (Figure 23)(Van Harsselaar et al., 2017).



**Figure 23. Représentation des domaines structuraux des β-amylases de pomme de terre (*Solanum tuberosum*, *St*).** Les séquences de l'ensemble des BAMs sont extraites du PGSC\_DM\_v3.4\_pep\_non-redondant (Avril 2017, 52.570 entrées) et annotées par Van Harsselaar et coll. (2017) (Xu et al., 2011; Van Harsselaar et al., 2017). *StBAM6.2* et *6.3* ont été retirées du fait de leur très faible taille, elles sont probablement incomplètes, et ne présentait pas de domaines Glycoside hydrolase lors de l'interrogation. *StBAM6.1* a été renommée *StBAM6*. Les peptides d'adressage au plaste ont été identifiés avec ChloroP 1.1 (<http://www.cbs.dtu.dk/services/ChloroP/>)(Emanuelsson et al., 1999). Les régions coiled-coil ont été identifiées avec Paircoil2 (<http://cb.csail.mit.edu/cb/paircoil2/paircoil2.html>) (McDonnell et al., 2006). Les autres régions ont été identifiées avec l'outil en ligne SMART (<http://smart.embl-heidelberg.de/>)(Letunic and Bork, 2017). Les rectangles gris représentent les domaines Glycoside Hydrolase de la famille 14, spécifiques des β-amylases. Les rectangles cyans représentent les peptides d'adressage au plaste (TP). Les rectangles violets représentent des domaines AP2 (APETALA2), des domaines de liaison à l'ADN. Le rectangle vert représente une région coiled-coil (CC). Le rectangle indigo représente le domaine BES1\_N, un facteur de transcription spécifique des plantes. BES1\_N: *Brassinosteroids insensitive 1- Ethyl methanesulfonate -Suppressor 1*, région N-terminale. AA : Acides aminés. Les noms notés en

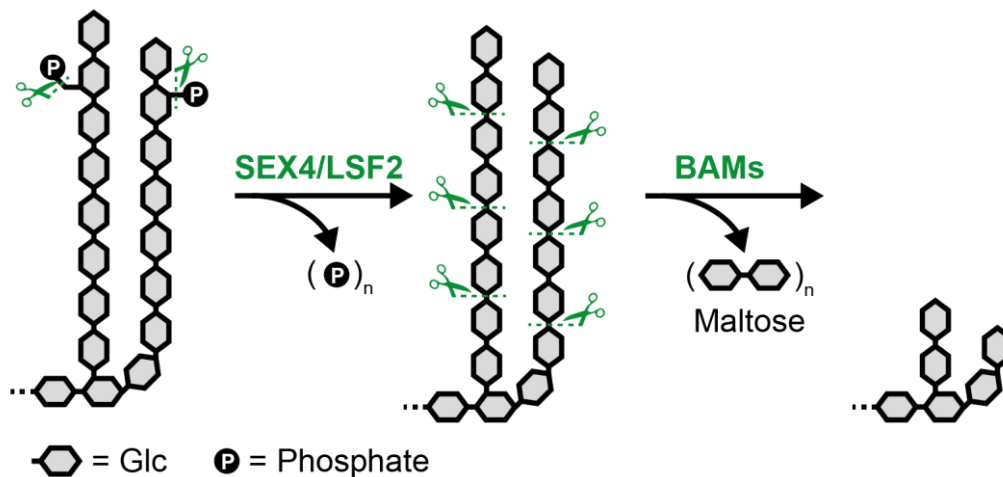
noir sont ceux de la publication de l'annotation du génome de la pomme de terre (Van Harsselaar et al., 2017). Les noms notés en rouge sont ceux corrigés après analyse phylogénétique des BAMs (voir Annexe 1).

La séquence de *StBAM2* de la publication de Van Harsselaar et coll. (2017) est plus proche phylogénétiquement de BAM8 que de BAM2 d'*Arabidopsis thaliana*. Qui plus est, la prédiction de domaine indique des domaines AP2 de liaison à l'ADN dans la séquence de *StBAM2* (Figure 23). Par conséquent, la protéine identifiée *StBAM2* est probablement un orthologue de BAM8 d'*Arabidopsis*. Les mutants *bam1* et *bam3* chez *Arabidopsis* montrent un phénotype de suraccumulation d'amidon, indiquant leur rôle dans la dégradation (Monroe and Storm, 2018). Le rôle de BAM 2 semble anecdotique comparé à celui de BAM1 et BAM3 dans la dégradation, car les mutants correspondants ont un phénotype proche de celui du sauvage (Monroe et al., 2017). Les premières analyses faites sur BAM2 laissaient penser que cette enzyme était inactive (Fulton et al., 2008). BAM4 serait elle aussi impliquée dans la dégradation chez *Arabidopsis*, mais ne possède pas d'activité enzymatique. Il a donc été proposé qu'elle aurait un rôle de régulateur (Fulton et al., 2008). L'analyse de l'expression de BAM4 dans différents tissus d'*Arabidopsis* a révélé que BAM4 est préférentiellement exprimée dans les tissus de réserve. Cependant, aucune étude n'a encore été menée au niveau des amyloplastes (Francisco et al., 2010). Les études sur BAM6 sont récentes et peu nombreuses, il semble toutefois que cette enzyme ne soit retrouvée que chez les Brassicacées, au niveau des plastes et qu'elle soit catalytiquement active (Thalmann et al., 2019; Monroe, 2020). Les études du génome de la pomme de terre ont permis d'identifier 3 séquences prédites orthologues de BAM6 (Xu et al., 2011; Van Harsselaar et al., 2017). L'analyse phylogénétique réalisée à partir des séquences identifiées chez la pomme de terre par Van Harsselaar et coll. (2017) semblent toutefois indiquer que BAM6.2 et BAM6.3 soient des fragments de séquences plutôt que des séquences de gènes entières du fait de leur brièveté. De plus, l'analyse phylogénétique des séquences de BAMs de pomme de terre et d'*Arabidopsis* montrent que BAM6.1 de pomme de terre est plus proche phylogénétiquement de BAM1 que de BAM6. Tout comme BAM4, BAM9 n'est pas active. Les mutants pour BAM9 chez *Arabidopsis* ont un phénotype proche du sauvage, mais les doubles mutants *bam3 bam9* ont un phénotype de mutant *bam3* exacerbé (Steidle, 2010; Monroe and Storm, 2018). BAM9 pourrait par conséquent jouer un rôle de régulateur de BAM3. Ces travaux sur BAM9 ont été réalisés lors d'un stage de Master et n'ont pas été publiés à ce jour. BAM9 a été localisée chez le tubercule de pomme de terre à la surface des grains d'amidon, montrant un rôle potentiel dans la dégradation de l'amidon de réserve, et dans le processus de sucrage à froid (Hou et al., 2017). Les intérêts économiques que peuvent

représenter ces protéines si elles sont impliquées dans de tels mécanismes sont encourageants pour la poursuite des études sur les BAMs dans un contexte d'amidon de réserve, où leurs rôles peuvent être différents de ceux constatés chez *Arabidopsis*.

### 3. La déphosphorylation par les phosphatases de l'amidon

Bien que la phosphorylation soit nécessaire pour initier la dégradation, la présence de groupements phosphate limite l'action des BAMs qui ne sont pas capables de passer outre lors de l'hydrolyse d'un glucane (Kötting et al., 2009). Il est donc nécessaire de retirer les résidus de phosphates pour permettre aux amylases de poursuivre la dégradation jusqu'à son terme. Les phosphatases de l'amidon hydrolysent les liaisons phosphoesters et libèrent une molécule de phosphate inorganique (Pi). L'activité de ces phosphatases est portée par un domaine catalytique qualifié « dual-specificity phosphatase » (DSP). Il appartient à une classe de domaines qui ont d'abord été décrits pour agir sur les protéines au niveau des sites cibles sérine/thréonine et tyrosine. Toutefois, deux protéines possédant un domaine DSP ayant une activité phosphatase sur de l'amylopectine phosphorylée ont été identifiées : Starch EXcess 4 (SEX4) et Like Sex Four 2 (LSF2) (Santelia et al., 2011). SEX4 est capable déphosphoryler les carbones C6 et C3 des résidus de glucose de l'amylopectine (Figure 24). LSF2 agit uniquement sur le carbone C3.



**Figure 24. Déphosphorylation de l'amidon par les phosphatases.** Chaque hexagone représente une unité de glucose. Les paires de ciseaux verts indiquent : à gauche, la position des liaisons phosphoesters qui seront hydrolysées par les phosphatases ; à droite, la position des liaisons  $\alpha$ -1,4 qui seront hydrolysées par les BAM.



### a. *SEX4*

Bien que la phosphorylation permette la décompaction des chaînes de glucanes, les enzymes de dégradation telles que les  $\beta$ -amylases et ISA3 ne peuvent outrepasser une unité de glucose phosphorylée lorsqu'elles sont en train d'hydrolyser un glucane. Les phosphatases vont donc retirer les groupements phosphates après dégradation partielle du glucane (Kötting et al., 2009). La première phosphatase décrite dans le métabolisme de l'amidon fut Starch EXcess 4 (SEX4) (Niittylä et al., 2006). Cette protéine est retrouvée au niveau du chloroplaste chez *Arabidopsis thaliana*. Le mutant *sex4* chez *Arabidopsis* accumule des glucanes phosphorylés produits au cours de la dégradation. De plus, la quantité d'amidon accumulée dans les plastes est plus grande dans les mutants *sex4*. Les mutants *sex4* ont aussi montré une forte réduction en termes de quantité et d'activité de l'enzyme de dégradation AMY3 (Zeeman et al., 1998a). La mutation *sex4* n'a toutefois pas d'effet sur la transcription de AMY3 (Comparot-Moss et al., 2010). Les tests d'activité de SEX4 ont montré la capacité de cette enzyme à déphosphoryler la surface de grains d'amidon et de phosphoglucanes solubles *in vitro* (Gentry et al., 2007; Kötting et al., 2009). Chez *A. thaliana*, SEX4 est une petite protéine de 42,6 kDa. La séquence de SEX4 est composée d'un domaine catalytique « dual-specificity phosphatase » (DSP) et d'un CBM48, permettant à la protéine de se lier à l'amidon (Figure 25) (Lombard et al., 2014). SEX4 permet de retirer les groupements phosphate en position C3 et C6 (Hejazi et al., 2010).

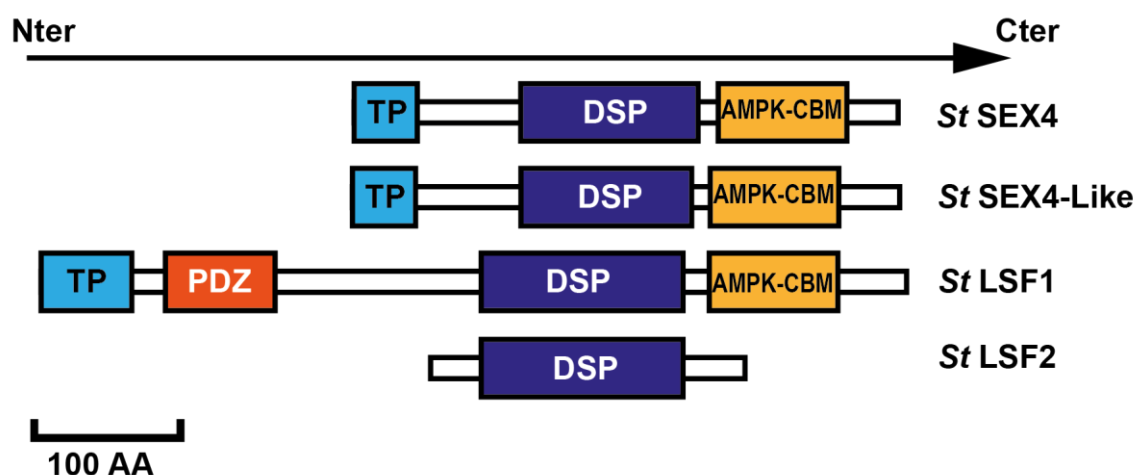
### b. *LSF1*

Les analyses du génome d'*Arabidopsis thaliana* ont permis d'identifier deux gènes homologues à SEX4, Like SEX Four 1 et 2 (LSF1 et LSF2) (Comparot-Moss et al., 2010). Les séquences de ces deux protéines possèdent aussi un domaine catalytique DSP, mais seule LSF1 possède un CBM48. LSF1 possède en plus de SEX4 un domaine « PDZ-like » qui est un domaine potentiel d'interaction protéine-protéine (Figure 25) (Santelia et al., 2011). Tout comme SEX4, LSF1 est retrouvée au niveau du chloroplaste. Le mutant *lsf1* chez *Arabidopsis* a montré une plus grande accumulation d'amidon dans les chloroplastes, comme dans le mutant *sex4*. Toutefois, il n'y a pas de réduction de l'activité et de la quantité de AMY3, et n'a pas non plus montré d'accumulation de glucanes phosphates (Comparot-Moss et al., 2010). Qui plus est, la mutation *lsf1* n'impacte pas le contenu en phosphate de l'amidon. De ce fait, il a été suggéré que l'enzyme LSF1 devait avoir une activité phosphatase différente de SEX4 et que LSF1 agirait plutôt comme une protéine phosphatase et donc dans la régulation de l'activité d'autres protéines du métabolisme de l'amidon (Comparot-Moss et al., 2010; Santelia et al., 2011). De récents tests *in vitro* mené sur LSF1 n'ont toutefois pas montré d'activité phosphatase. Qui plus

est, l'insertion d'une protéine LSF1 modifiée au niveau de son site DSP dans des plants mutés *lsf1* d'*Arabidopsis* retrouvent un phénotype sauvage. En revanche, la même expérience menée en modifiant cette fois-ci son CBM ne permet pas un retour du phénotype sauvage. Les tests d'interaction par immunoprécipitation et par méthode de marquage et de purification TAP (pour « Tandem affinity purification ») ont démontré que LSF1 formait des complexes avec BAM1 et BAM3 (Schreier et al., 2019). Il a été hypothétisé par les auteurs que LSF1 servirait de protéine d'échafaud entre les phosphoglucanes et ces  $\beta$ -amylases.

### c. *LSF2*

L'autre homologue de SEX4, LSF2, a été caractérisée en 2011 (Santelia et al., 2011). Tout comme SEX4 et LSF1, LSF2 est localisée dans le chloroplaste chez *Arabidopsis* (Schliebner et al., 2008; Santelia et al., 2011). De plus, elle a été retrouvée séquestrée dans les grains d'amidon (Santelia et al., 2011). On retrouve dans sa séquence le domaine DSP présent chez SEX4 et LSF1, mais pas le domaine CBM48. Les études structurales par cristallographie de LSF2 ont toutefois montré sa capacité à se lier aux glucanes et au phosphate au niveau de son site actif (Meekins et al., 2013). De plus, les analyses cristallographiques de LSF2 ont permis l'identification de deux autres sites de liaison aux glucanes situés à plus de 20 Å du site actif (Meekins et al., 2013). LSF2 est capable de retirer les phosphates à la surface de l'amidon et sur des glucanes phosphorylés, tout comme SEX4 (Santelia et al., 2011). De plus, les mutants *lsf2* n'accumulent pas de phospho-oligosaccharides. Des tests menés sur des glucanes phosphorylés radiomarqués au  $^{33}\text{P}$ , soit en position C3, soit en position C6, ont montré que LSF2 agissait spécifiquement sur les phosphates en position C3, alors que SEX4 qui peut agir à la fois sur le phosphate en position C3 et en position C6 (Hejazi et al., 2010; Santelia et al., 2011). La quantité d'amidon accumulé dans les plastes reste inchangée dans les mutants *lsf2* d'*Arabidopsis*, contrairement aux mutants *lsf1* et *sex4*. Le double mutant *sex4/lsf2* montre toutefois une accumulation beaucoup plus importante de l'amidon dans les plastes que le mutant *sex4*. L'activité phosphatase de LSF2 est probablement compensée par celle de SEX4 lorsque cette première est absente.



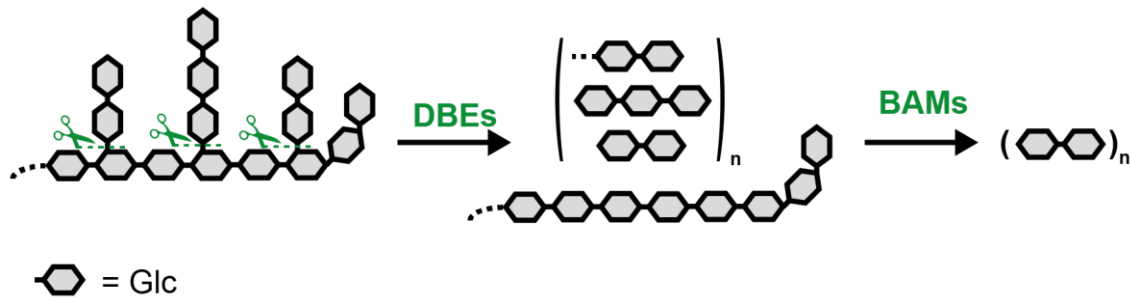
**Figure 25. Représentation des domaines structuraux des phosphatases de pomme de terre (*Solanum tuberosum*, *St*).** Les séquences de l'ensemble des phosphatases sont extraites du PGSC\_DM\_v3.4\_pep\_non-redundant (Avril 2017, 52.570 entrées) et annotées par Van Harsselaar et coll. (2017) (Xu et al., 2011; Van Harsselaar et al., 2017). Les peptides d'adressage au plaste (TP, cyan) ont été identifiés avec ChloroP 1.1 (<http://www.cbs.dtu.dk/services/ChloroP/>) (Emanuelsson et al., 1999). Les autres régions ont été identifiées avec l'outil en ligne SMART (<http://smart.embl-heidelberg.de/>) (Letunic and Bork, 2017). Les rectangles bleus représentent des domaines DSP (Dual specificity protein phosphatases), permettant de retirer un phosphate, soit en position C3 ou C6 sur l'amidon, soit sur un site Tyrosine ou Sérine/Thréonine d'une protéine. Les rectangles ocre représentent un module de liaison aux carbohydrates AMPK-CBM. Le rectangle carmin représente un domaine PDZ, un domaine de liaison aux polypeptides. AMPK-CBM : module de liaison aux carbohydrates activé par une AMP-kinase. PDZ : PSD-95, Dlg et ZO-1 (3 protéines chez lesquelles ce domaine est commun). AA : Acides aminés.

En résumé, seule SEX4 et LSF2 ont une activité de déphosphorylation de l'amidon. La première peut déphosphoryler en C6 et en C3, alors que la deuxième ne déphosphoryle l'amidon qu'en position C3. Pour l'instant, seule LSF2 a été trouvée séquestrée dans les grains d'Arabidopsis. Chez la pomme de terre, l'analyse du génome a montré la présence des gènes *SEX4*, *LSF1* et *LSF2* ; ainsi que d'une 4<sup>ème</sup> isoforme n'ayant pas encore été caractérisée : *SEX4-like* (Van Harsselaar et al., 2017). La séquence de *SEX4-like* est très proche de celle de *SEX4* et est probablement un gène dupliqué de *SEX4*.

#### 4. Les enzymes de débranchement dans la dégradation

En plus de ISA1 et ISA2 décrites précédemment dans la partie synthèse, on distingue deux enzymes de débranchement impliquées dans la dégradation de l'amidon : l'isoamylase 3 (ISA3) et la pullulanase (PU1) (Wattebled et al., 2008). Leur activité est la même que celle d'ISA1 et consiste à cliver par hydrolyse les liaisons  $\alpha$ -1,6. Les produits résiduels (des glucanes linéaires

ou des dextrans) vont ainsi pouvoir être dégradés par les BAMs en maltose (Figure 26) (Møller and Svensson, 2016).



**Figure 26. Clivage des points de branchements par les enzymes de débranchement.** Chaque hexagone représente une unité de glucose, les paires de ciseaux représentent les liaisons  $\alpha$ -1,6 hydrolysées par les enzymes de débranchement. DBEs : Enzymes de débranchement. BAMs :  $\beta$ -amylases.

#### a. ISA3

L'isoamylase 3 (ISA3) est une enzyme de débranchement impliquée essentiellement dans la dégradation. Elle fut d'abord mise en évidence dans l'amyloplaste et le chloroplaste chez la pomme de terre, (Hussain et al., 2003). Elle fut ensuite retrouvée chez d'autres espèces telles qu'*Arabidopsis thaliana* et dans l'amyloplaste chez le riz (Delatte et al., 2005; Yun et al., 2011). Tout comme ISA1, ISA3 possède une activité catalytique (Hussain et al., 2003). Chez *Arabidopsis*, son activité est partiellement redondante de celle du complexe ISA1/ISA2 (Streb et al., 2008; Wattebled et al., 2008). ISA3 n'est pas associée au complexe formé par ISA1 et ISA2 et agit indépendamment de ces deux isoformes (Hussain et al., 2003). Le rôle d'ISA3 semble être anecdotique dans la synthèse, cette enzyme serait plutôt impliquée dans la dégradation de l'amidon, l'analyse des lignées mutantes d'*Arabidopsis* pour ISA3 ayant montré que l'absence de l'enzyme entraîne une augmentation du contenu en amidon en fin de jour, ainsi qu'une diminution du taux de dégradation en fin de nuit (Delatte et al., 2005; Wattebled et al., 2005). Chez la pomme de terre, la triple mutation *isa1 isa2 isa3* entraîne une forte diminution de la quantité d'amidon dans les tubercules, mais le contenu en amidon transitoire dans les feuilles n'est pas impacté (Ferreira et al., 2017).

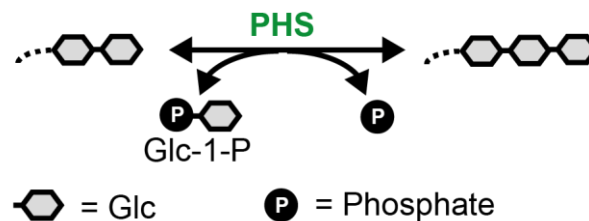
#### b. PU1

En plus des isoamylases, on retrouve une autre forme d'enzyme de débranchement chez les plantes : la pullulanase, PU1. Elle est aussi appelée « Limit-dextrinase » (LDA). Elle fut identifiée chez les micro-organismes et les plantes avant les isoamylases (Bender and Wallenfels, 1966; Gordon et al., 1975). Elle montre une plus forte affinité pour le pullulane, un

polymère bactérien regroupant des unités de maltotrioses liés en  $\alpha$ -1,6. Elle possède aussi une forte affinité pour l'amylopectine mais a peu d'affinité pour le glycogène. Tout comme ISA1, ISA2 et ISA3, PU1 est impliquée dans la synthèse et le catabolisme de l'amidon, mais de manière plus anecdotique (Dinges et al., 2003; Wattebled et al., 2005). Les mutants *pu1* ont en effet un phénotype proche de celui des plantes sauvages, mais les doubles mutants *pu1 isa3* présentent un phénotype *isa3* exacerbé (Delatte et al., 2005; Wattebled et al., 2008). La pullulanase aurait donc une activité redondante de celle de ISA3 pour la dégradation de l'amidon.

## 5. Les amidons phosphorylases

Le rôle des amidon-phosphorylases dans la dégradation de l'amidon est sujet à débat. Les phénotypes des mutants phosphorylases sont parfois opposés selon les études. Il semble que ces enzymes soient impliquées à la fois dans les mécanismes de synthèse et ceux de dégradation. En effet, elles sont en mesure de retirer une unité de glucose en  $\alpha$ -1,4 par exo-phosphorolyse (elles agissent depuis une extrémité non-réductrice en remontant progressivement un glucane vers son extrémité réductrice ou le point de branchement), formant ainsi du Glucose-1-phosphate (Glc-1-P). Elles sont toutefois aussi capables, *in vitro*, de catalyser la réaction inverse, et d'ajouter un résidu de glucose à l'extrémité non-réductrice d'un glucane à partir du Glc-1-P (Figure 27) (Steup et al., 1976). On distingue deux classes de phosphorylases dans le métabolisme de l'amidon : les amidons phosphorylases 1 (PHS1) et les amidons phosphorylases 2 (PHS2) (Kötting et al., 2005). Elles agissent préférentiellement sur des chaînes de 5 résidus de glucose (Steup, 1981).

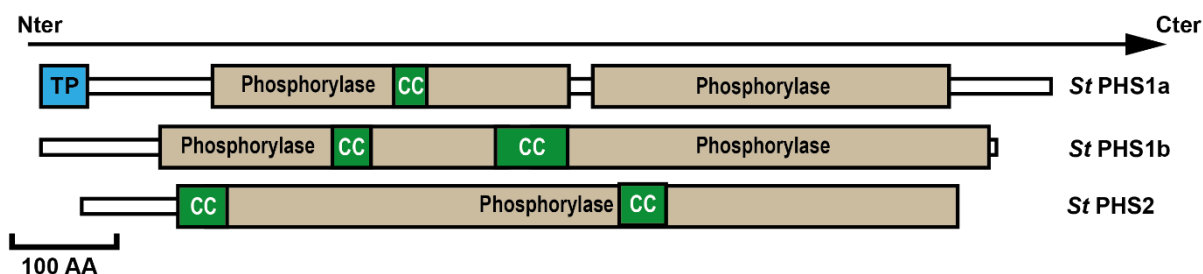


**Figure 27. Action des phosphorylases.** Chaque hexagone représente une unité de glucose. Les cercles noirs avec un P blanc à l'intérieur représentent les phosphates. PHS : Phosphorylase.

PHS1, aussi appelée phosphorylase de type L, est une enzyme plastidiale. PHS2 aussi appelée phosphorylase de type H est quant à elle cytosolique (Steup and Latzko, 1979). Chez la pomme de terre, PHS1 a un poids moléculaire de 109,5 kDa et PHS2 de 95,4 kDa (Mori et al., 1991). Ces deux formes, cytosolique et plastidiale, sont aussi présentes chez *Arabidopsis*,

la pomme de terre et *Chlamydomonas reinhardtii* (Sonnewald et al., 1995; Zeeman et al., 2004; Dauvillée et al., 2006). Une troisième isoforme, proche de PHS1, existe dans les racines du pois, mais elle est absente des tissus photosynthétiques (Steup and Latzko, 1979). Cette forme d'amidon-phosphorylase fut aussi caractérisée dans les amyloplastes de tubercule de pomme de terre et du grain de maïs. (Sonnewald et al., 1995; Yu et al., 2001b). Dans le génome de la pomme de terre, on retrouve deux isoformes de PHS1 et de PHS2 (Van Harsselaar et al., 2017). PHS1a correspond à l'isoforme principalement localisée dans les amyloplastes, PHS1b à celle localisée dans les chloroplastes. PHS2 a et b sont quant à elles deux isoformes cytosoliques.

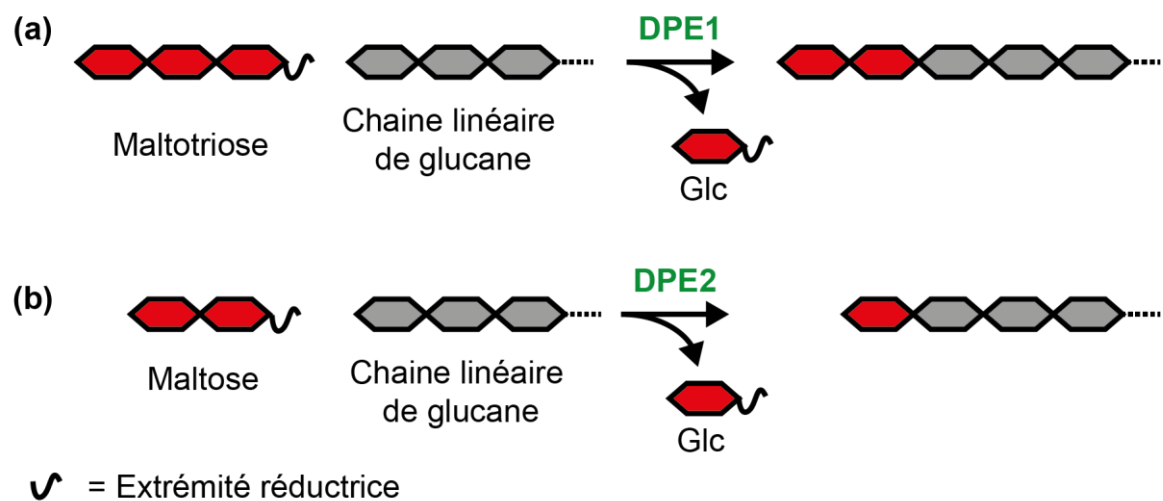
Les amidon-phosphorylases seraient impliquées dans l'initiation de la synthèse de l'amidon. En effet, celles-ci sont capables, en présence d'une concentration élevée de Glc-1-P, d'ajouter des unités de glucose sur de très courts glucanes y compris le maltose (DP2) (Sonnewald et al., 1995). De par cette capacité, il est possible que PHS1 intervienne dans les stades précoces de l'initiation de la synthèse, en allongeant le maltose pour obtenir un glucane assez long pour être reconnu par le complexe d'initiation (Seung et al., 2017). L'enzyme PHS1a, impliquée dans ce mécanisme, est principalement localisée dans les amyloplastes. Une étude menée sur l'orge a montré qu'elle était aussi présente dans le grain d'amidon sous forme phosphorylée (Ahmed et al., 2015).



**Figure 28. Représentation des domaines structuraux des phosphorylases de pomme de terre (*Solanum tuberosum*, *St*).** Les séquences de PHS1a, PHS1b et PHS2 de pomme de terre sont extraites de Uniprot (Accessions : P04045, P53535 et P32811, respectivement). Les peptides d'adressage au plaste (TP, cyan) ont été identifiés avec ChloroP 1.1 (<http://www.cbs.dtu.dk/services/ChloroP/>) (Emanuelsson et al., 1999). Les régions coiled-coil (CC, vert) ont été identifiées avec Paircoil2 (<http://cb.csail.mit.edu/cb/paircoil2/paircoil2.html>) (McDonnell et al., 2006). Les autres régions ont été identifiées avec SMART (<http://smart.embl-heidelberg.de/>) (Letunic and Bork, 2017). Les rectangles beiges représentent les domaines phosphorylases, le rectangle bleu le peptide d'adressage au plaste et les rectangles verts les domaines coiled coil. AA : Acides aminés.

## 6. Les enzymes de dismutations ou $\alpha$ -glucanotransférases (DPE)

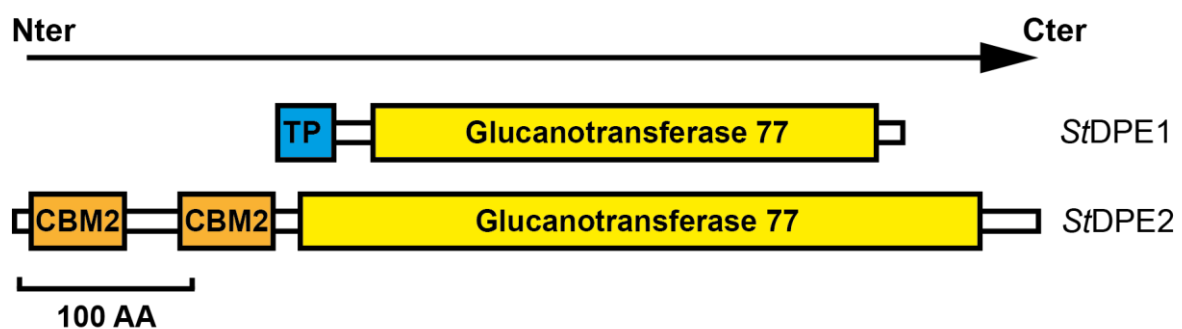
Les enzymes de dismutation (ou DPE pour disproportionating enzyme), aussi appelés alpha-glucoano-transférases jouent un rôle crucial dans la dégradation en empêchant l'accumulation de courts malto-oligo-saccharides (MOS), principalement du maltotriose et du maltose, en clivant une chaîne de glucane linéaire en  $\alpha$ -1,4 et en la transférant sur un autre glucane linéaire au niveau de l'extrémité non réductrice en  $\alpha$ -1,4 tout en relarguant une unité de glucose (Figure 29) (Critchley et al., 2001).



**Figure 29 Action des  $\alpha$ -glucanotransférases.** Chaque hexagone représente une unité de glucose. (a) Action de DPE1 à partir d'un maltotriose sur un glucane linéaire au niveau du plaste. (b) Action de DPE2 à partir d'un maltose sur un glucane linéaire au niveau du cytosol.

Tout comme les phosphorylases, il existe 2 isoformes, une plastidique et une cytosolique. La première est nommée DPE1, et la deuxième DPE2. Chez la pomme de terre, DPE1 a un poids moléculaire de 65 kDa et DPE2 de 110 kDa (Takaha et al., 1993; Lloyd et al., 2004). Toutes les deux possèdent un domaine glucanotransférase et un CBM (Figure 30). Les phénotypes mutants chez DPE1 montrent une accumulation anormale de courts MOS linéaires. Plusieurs hypothèses existent quant à la fonction de DPE1 dans le métabolisme (Critchley et al., 2001; Wattedled et al., 2003). Chez *Chlamydomonas reinhardtii*, DPE1 pourrait agir pendant la synthèse, en retransférant les chaînes clivées par les isoamylases sur une chaîne linéaire d'amylopectine, réduisant ainsi la perte de matériel à une unité de glucose (Wattedled et al., 2003). Chez le riz, elle est aussi en mesure d'utiliser l'amylose et l'amylopectine comme substrat donneur, et l'amylopectine comme substrat accepteur, démontrant un rôle potentiel dans la synthèse de l'amidon (Dong et al., 2015). Elle pourrait aussi agir pendant la dégradation,

de manière à fournir des substrats avec une meilleure affinité pour les phosphorylases et les  $\beta$ -amylases. DPE2, cytosolique quant à elle, possède comme substrat donneur préférentiel le maltose, qu'elle va dégrader en glucose tout en transférant une unité sur un substrat accepteur. Elle contribuerait ainsi à la dégradation du maltose en glucose, qui serait ensuite converti en Glc-6-P par l'hexokinase. Le rôle des glucanotransférases est fortement lié à celui des phosphorylases. Ces deux enzymes vont permettre d'allonger des chaînes qui ne peuvent pas être prises en charge par les enzymes de dégradation, et contribuent elles même directement à la dégradation en relarguant du glucose, ou du Glc-1-P, qui servira ensuite à la synthèse du saccharose pour le transport et la remobilisation de l'énergie au sein de la plante.



**Figure 30. Représentation des domaines structuraux des phosphorylases de pomme de terre (*Solanum tuberosum*, *St*).** Les séquences de DPE1 et DPE2 de pomme de terre sont extraites de Uniprot (Accessions : Q06801 et Q6R608, respectivement). Les peptides d'adressage au plaste (TP, cyan) ont été identifiés avec ChloroP 1.1 (<http://www.cbs.dtu.dk/services/ChloroP/>) (Emanuelsson et al., 1999). Les régions coiled-coil (CC, vert) ont été identifiées avec Paircoil2 (<http://cb.csail.mit.edu/cb/paircoil2/paircoil2.html>) (McDonnell et al., 2006). Les autres régions ont été identifiées avec SMART (<http://smart.embl-heidelberg.de/>) (Letunic and Bork, 2017). Les rectangles jaunes représentent les domaines glucanotransférases, les rectangles oranges les domaines CBM2 et le rectangle bleu le peptide de transfert au plaste. AA : Acides aminés.

## 7. Les protéines sans activité catalytique connue impliquées dans le métabolisme de l'amidon

Récemment, d'autres protéines ne possédant pas d'activité catalytique connue ont été localisées dans le stroma du plaste et la mutation du gène les codant impacte le métabolisme de l'amidon. Parmi ces protéines se trouvent les protéines d'adressage au grain, nommées PTST pour Protein targeting to starch, ne possédant pas d'activité catalytique mais ayant des domaines structuraux permettant des interactions protéine-protéine, et protéine-glucane (Seung et al., 2015; Seung et al., 2017). Deux autres protéines ont été découvertes chez *Arabidopsis*, Early

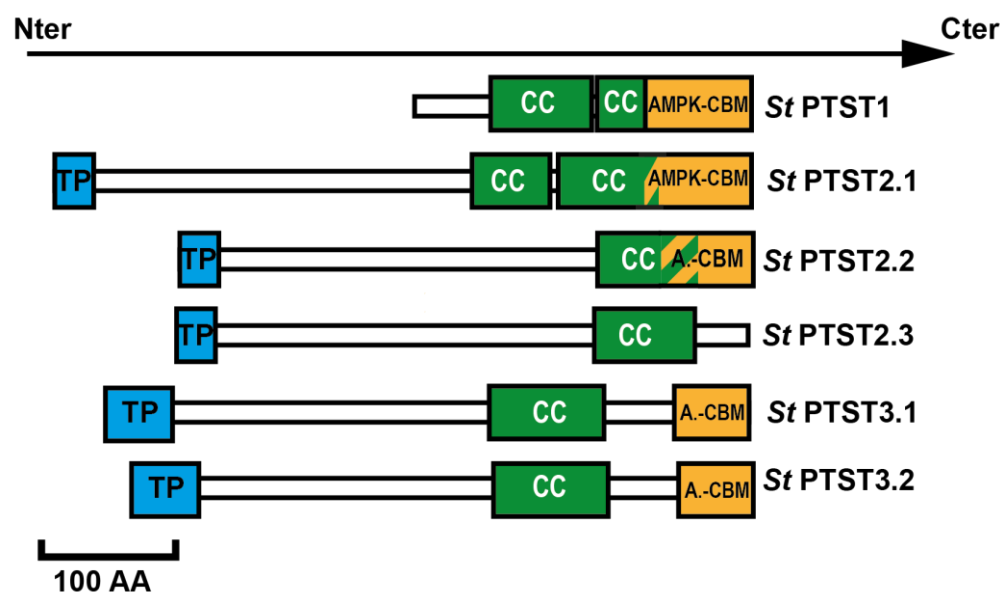


StarVation 1 (ESV1) et son homologue Like ESV1 (LESV) (Feike et al., 2016). Ces deux protéines n'ont pas d'activité catalytique connue, elles semblent toutefois impliquées dans la régulation de la dégradation avec des activités antagonistes. Il est proposé que ces protéines jouent un rôle dans le maintien de la structure de l'amidon. Enfin, de récentes études ont permis de mettre à jour chez *Arabidopsis* deux protéines non catalytiques impliquées dans l'initiation de la synthèse du grain d'amidon, MFP-1 et PII1 (Seung et al., 2018; Vandromme et al., 2019).

## 8. Les protéines d'adressage « Protein Targeting to Starch »

### a. *PTST1*

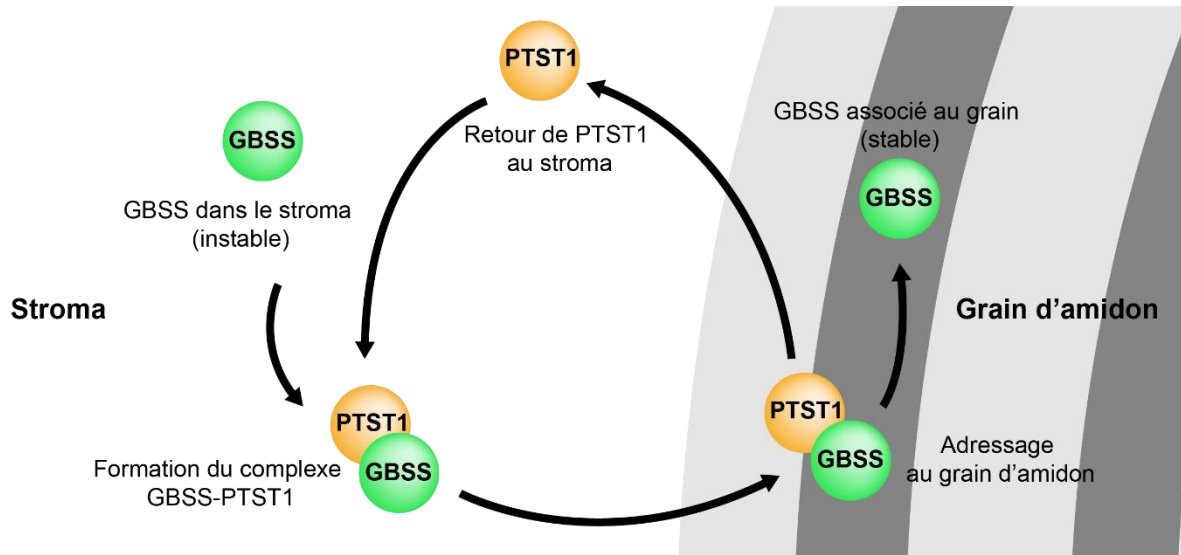
Protein Targeting to STarch (PTST1) fut décrite pour la première fois en 2008, chez *Arabidopsis thaliana*, comme une protéine potentiellement impliquée dans le métabolisme de l'amidon (Lohmeier-Vogel et al., 2008). Elle fut toutefois caractérisée plus en détail en 2015 (Seung et al., 2015). Cette protéine est fortement conservée au sein des plantes vertes. C'est une petite protéine de 26 kDa, possédant un peptide de transit, un domaine coiled-coil et un module de liaison aux carbohydrates, mais elle ne possède pas d'activité enzymatique (Figure 31).



**Figure 31. Représentation des domaines structuraux des protéines d'adressage de pomme de terre (*Solanum tuberosum*).** Les séquences de PTST utilisées pour cette figure sont extraites du PGSC\_DM\_v3.4\_pep\_non-redundant (Avril 2017, 52.570 entrées)(Xu et al., 2011). Elle ont été retrouvées et annotées à partir de la séquence de PTST1 d'*Arabidopsis thaliana* (Seung et al., 2015). Les séquences de PTST2.1, PTST 2.2, PTST 2.3, PTST3.1 et PTST3.2 (Accessions : XP\_006367347.1, XP\_006367349.1, XP\_006367350.1, XP\_006354369.1 et XP\_015167316.1, respectivement) sont des séquences prédites obtenues à partir d'un BLASTp sur les séquences de PTST2 et 3 de tomates (*Solanum Lycopersicum*) présentes dans la

publication de Seung et collaborateurs de 2017 (Seung et al., 2017). Les peptides de transfert ont été identifiés avec ChloroP 1.1 (<http://www.cbs.dtu.dk/services/ChloroP/>) (Emanuelsson et al., 1999). Les régions coiled-coil ont été identifiées à l'aide de Paircoil2 (<http://cb.csail.mit.edu/cb/paircoil2/paircoil2.html>) (McDonnell et al., 2006). Les autres régions ont été identifiées avec l'outil en ligne SMART (<http://smart.embl-heidelberg.de/>) (Letunic and Bork, 2017). Les rectangles cyans représentent les peptides de transfert au chloroplaste. Les rectangles verts représentent les régions coiled-coil. Les rectangles ocres représentent un module de liaison aux carbohydrates AMPK-CBM. TP : Peptide de transfert. CC : coil-coiled. AMPK-CBM/A.-CBM : module de liaison aux carbohydrates activé par une AMP-kinase. *St. Solanum tuberosum*. AA : Acides aminées.

Les mutants *ptst*, chez qui la protéine PTST1 est absente, ne produisent plus d'amylose. Ce phénotype est habituellement celui des mutants *waxy*, déficitaires en GBSS. De plus, le contenu de GBSS est fortement diminué dans le grain d'amidon lorsque *PTST1* est muté. Des expériences d'immunoprécipitation ont permis de confirmer l'interaction entre GBSS et PTST1, et le rôle de PTST1 dans l'adressage de GBSS au grain d'amidon. PTST1 est principalement retrouvée dans le stroma des chloroplastes chez *Arabidopsis thaliana*, mais elle se trouve également en association avec l'amidon bien qu'en faible quantité (Seung et al., 2015). Les auteurs ont proposé que PTST1 faisait la navette entre le grain et le stroma, et qu'elle ne serait présente que de manière transitoire dans le grain d'amidon (Figure 32) (Seung et al., 2015). PTST1 a récemment été retrouvée en association avec GBSS chez le manioc et chez l'orge (Bull et al., 2018; Zhong et al., 2019). Chez le manioc, les mutants *ptst1* présentent une réduction de la quantité d'amylose, là où l'amylose était totalement absente chez les mutants d'*Arabidopsis*. Cette différence peut venir du fait que l'amidon étudié dans le manioc est de l'amidon de réserve, et non de l'amidon transitoire (Bull et al., 2018). Les mutants *ptst1* de l'orge n'accumulent plus d'amidon. Cela semble indiquer que PTST1 n'adresserait pas seulement GBSS au grain chez l'orge, mais jouerait aussi un rôle essentiel dans la synthèse d'amidon au niveau du grain (Zhong et al., 2019).



**Figure 32. Adressage de GBSS au grain par PTST1.** Ce modèle est basé sur le mécanisme proposé par Seung et collaborateurs en 2015 (Seung et al., 2015). PTST1: Protein Targeting To Starch 1. GBSS: Granule Bound Starch Synthase.

#### b. *PTST2 et 3*

Les analyses phylogéniques, menées à partir de la séquence de *PTST1* ont permis l'identification de deux homologues chez *Arabidopsis* : *PTST2* et *PTST3* (Seung et al., 2017). Tout comme *PTST1*, les deux protéines correspondantes sont fortement conservées au sein du règne végétal, et possèdent aussi un domaine coiled-coil et un CBM. La protéine FLO6 chez le riz, dont le mutant est connu pour altérer l'initiation de la synthèse du grain, s'avère être un homologue de *PTST2* (Peng et al., 2014; Chia et al., 2019a). *PTST2* et 3 sont plus grandes que *PTST1*, La première possède une masse moléculaire de 51,4 kDas et la seconde de 61,0 kDas, après clivage du peptide de transit. Du fait de leur similarité avec *PTST1*, il fut d'abord hypothétisé que ces protéines avaient pour rôle d'adresser les autres SSs au grain. Les mutants *ptst2* et *ptst3* d'*Arabidopsis thaliana* ont montré un défaut de dégradation de l'amidon pendant la nuit. L'activité de *PTST2* et *PTST3* est partiellement redondante. Le double mutant *ptst2-ptst3* montrait un profil de dégradation de l'amidon similaire à celui observé pour le mutant *ss4*. De plus, les mutants *ptst2* et *ptst2-ptst3* ne montraient la présence que d'un unique grain dans la plupart des chloroplastes, beaucoup plus gros que celui des lignées sauvages, phénotype aussi retrouvé chez les mutants *ss4*. Le contenu en ADP-glucose est aussi fortement augmenté dans le double mutant *ptst2-ptst3*, tout comme le mutant *ss4*. Les expériences d'immunoprécipitations ont montré que *PTST2* interagissait avec *SS4*, et que *PTST3* interagissait avec *PTST2*. Grâce à son CBM, *PTST2* est capable de se lier à de longues chaînes

de malto-oligosaccharides (MOS), mais pas aux chaînes courtes (moins de 10 unités de glucose). PTST2 serait donc probablement capable de se lier aux chaînes formant une structure en hélice. Cela a amené les auteurs à penser que PTST2 et PTST3 forment un complexe avec SS4 de manière à pouvoir adresser SS4 à son substrat, qui serait un ou des MOS avec un degré de polymérisation (DP) d'au moins 10 unités de glucoses. Ces deux protéines jouent donc un rôle crucial dans l'initiation de la synthèse du grain d'amidon. Chez la pomme de terre, les analyses que j'ai réalisées par blastp des séquences de PTST2 et 3 de tomates présentes dans la publication de Seung et coll. (2017) permettent d'identifier 3 isoformes de PTST2 et 2 isoformes de PTST3 (Seung et al., 2017). Toutes, hormis PTST2.3, possèdent un module de liaison aux carbohydrates ainsi qu'un domaine coiled-coil (Figure 31). Il est probable que ces protéines aient la même fonction que leurs orthologues chez *Arabidopsis* au niveau des feuilles de pommes de terre.

## 9. Early StarVation proteins

### a. *ESV1*

La protéine Early StarVation 1 (*ESV1*) a été découverte chez *Arabidopsis thaliana* en 2016 (Feike et al., 2016). L'analyse phylogénétique de son gène de structure montre qu'elle est conservée chez les angiospermes, les monocotylédones et les algues vertes. La séquence de *ESV1*, d'un poids moléculaire prédit de 49 kDa, est riche en tryptophane et possède une région riche en proline en position C-ter. Elle est à la fois présente dans le stroma des chloroplastes et associée au grain d'amidon. Elle serait un régulateur négatif de la dégradation. Les mutants *esv1* entraînent une diminution rapide et non linéaire de l'amidon transitoire durant la nuit, les réserves d'amidon étant épuisées avant l'aube. Une autre étude, réalisée en 2018, a montré une implication de cette protéine dans le mécanisme de phosphorylation de l'amidon, permis par les dikinases GWD et PWD (Malinova et al., 2018). Lors d'incubations simultanées de *ESV1* avec GWD ou PWD, une diminution de la phosphorylation fut observée dans le premier cas, alors qu'une augmentation de la phosphorylation fut observée dans le deuxième cas. À l'heure actuelle, *ESV1* n'a pas d'activité enzymatique connue. Elle ne possède pas de domaines ayant été précédemment annotés, son activité est donc difficilement prévisible. Il a toutefois été proposé qu'elle jouerait un rôle central dans l'organisation des polymères de glucanes au sein des grains d'amidon.

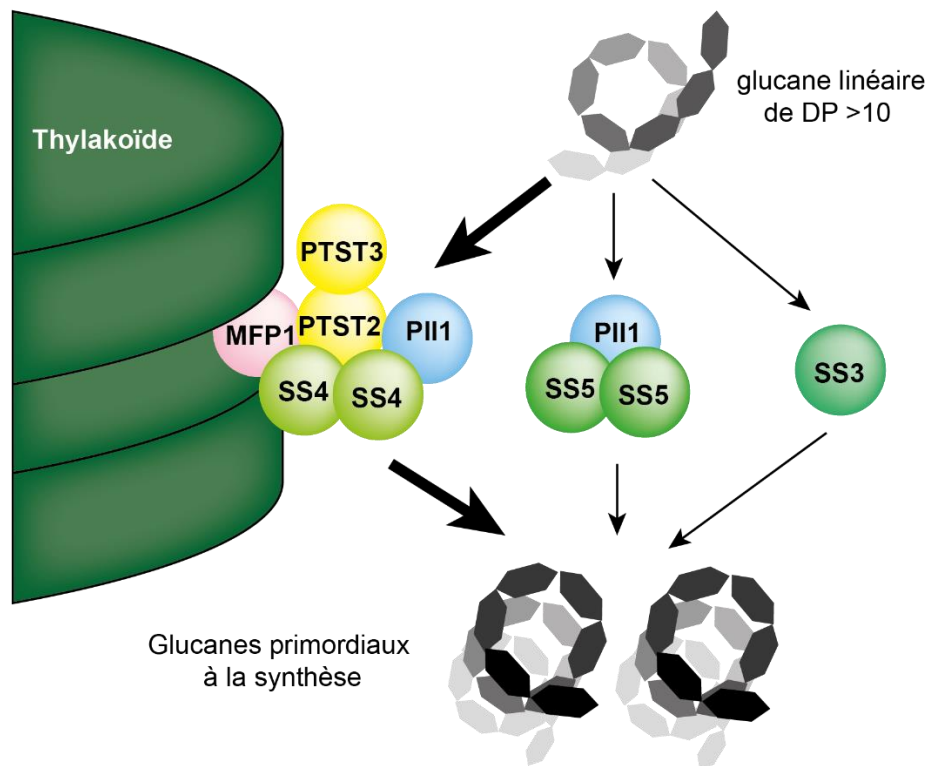
### *b. LESV*

La protéine Like-ESV1 (LESV) fut découverte en même temps que ESV1 chez *Arabidopsis thaliana* (Feike et al., 2016). Elle est elle aussi conservée chez la plupart des plantes. Elle est présente à la fois dans le stroma des chloroplastes et dans le grain d'amidon. Elle est un peu plus grande que ESV1, avec une taille de 66 kDa, mais possède toutefois une forte homologie de séquence avec ESV1. Toutes deux partagent une région riche en tryptophane, mais elle ne possède pas de région riche en proline en C-ter. Elle possède aussi un peptide de transit en position N-ter qui est absent chez ESV1 d'*Arabidopsis thaliana* mais qui a été retrouvé chez des homologues d'autres espèces telles que la vigne, et le manioc. Elle aussi est impliquée dans les mécanismes de dégradation mais semble avoir un rôle antagoniste à celui de ESV1. Tout comme cette dernière, LESV n'a pour le moment pas d'activité enzymatique connue. Il semblerait que ces deux enzymes agissent sur la structuration des glucanes, permettant ou empêchant l'accessibilité des enzymes de dégradation en modifiant leur structure. Une autre isoforme potentielle, plus proche de la séquence de ESV1 que de celle de LESV, est présente dans le génome de la pomme de terre (données non publiées, communication de François Sevestre et Mara Berti, Université de Lille, CNRS, UGSF UMR8576, Université de Lille, Villeneuve d'Ascq). L'étude de leur structure par cristallographie RX de ces deux protéines est actuellement en cours chez *Arabidopsis* et la pomme de terre de manière à identifier des sites catalytiques potentiels et/ou des domaines de liaisons à l'amidon ou à d'autres protéines (données non publiées, communication de Coralie Bompard, CNRS, UGSF UMR8576, Université de Lille, Villeneuve d'Ascq).

### **10. Autres protéines récemment découvertes comme impliquées dans l'initiation : PII (MRC) et MFP1.**

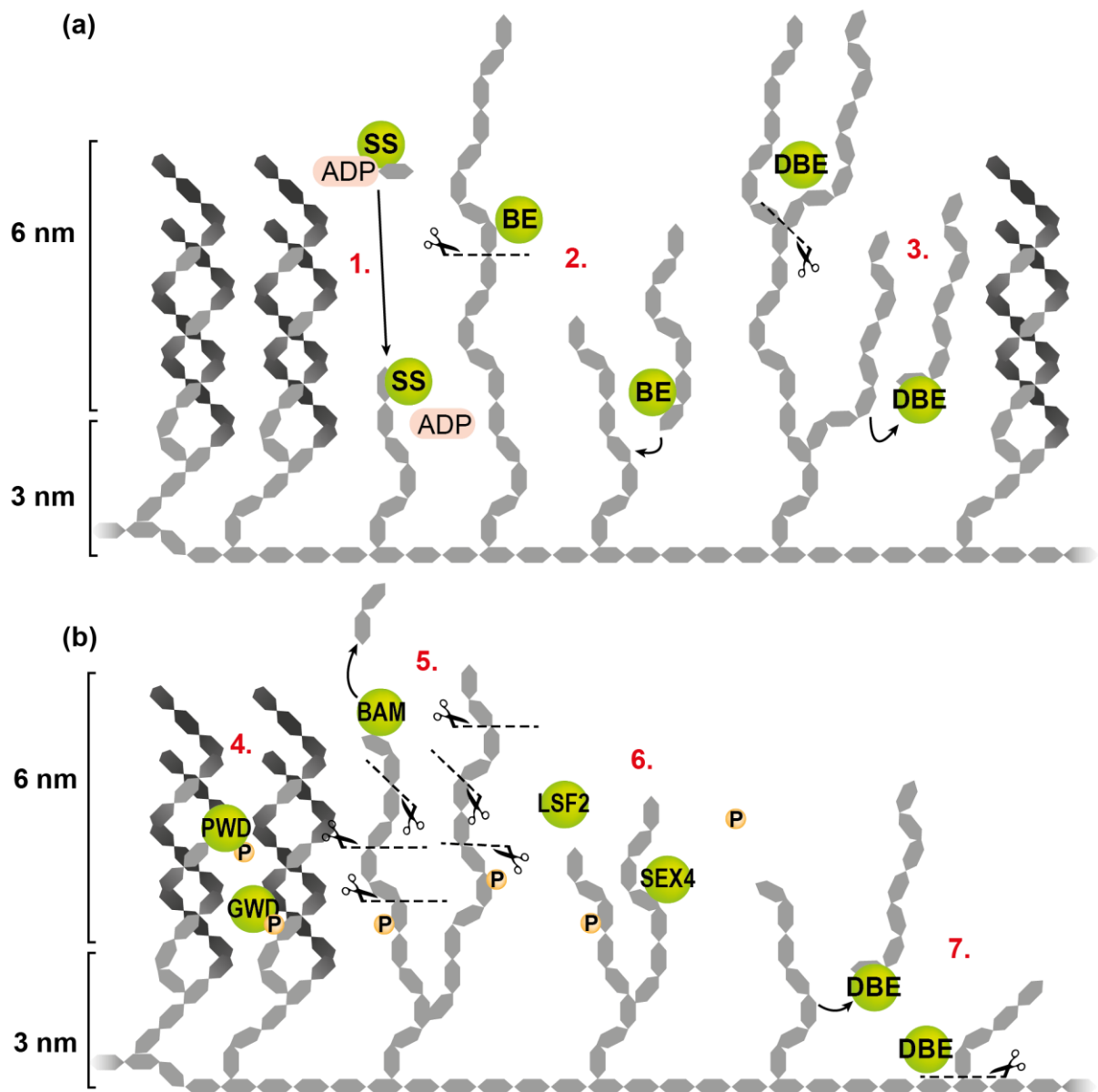
Après identification des protéines PTST2 et PTST3 impliquées dans l'initiation de la synthèse de l'amidon, Seung et collaborateurs ont mené des expériences de co-immunoprécipitation avec PTST2 pour identifier d'autres interactants (Seung et al., 2018). En parallèle, Vandromme et collaborateurs ont utilisé la technique de double-hybride avec SS4 comme protéine test pour identifier de nouveaux interactants potentiels de SS4 (Vandromme et al., 2019). Dans le premier cas, deux interactants ont été retenus et nommés MFP1 et MRC. Dans le deuxième cas, un interactant a été retenu, qui a été nommé PII1. Il s'avère toutefois que MRC et PII1 sont la même protéine (identifiant At4g32190). La protéine MFP1, pour « MAR BINDING FILAMENT- LIKE PROTEIN1 » fut d'abord décrite comme une protéine de liaison

aux nucléotides, avec un domaine « coiled-coil » présageant une interaction avec d'autres protéines (Jeong et al., 2003). C'est une protéine de 82 kDa hautement conservée au sein des différents règnes, et localisée dans le noyau. Toutefois, chez *Arabidopsis*, elle fut observée au niveau de la membrane des thylakoïdes. La Protéine MRC, pour « myosin-resembling chloroplast protein », rebaptisée « Protein involved in starch initiation » (PII1), ne possède aucune activité connue. L'analyse de sa séquence a toutefois montré la présence de plusieurs domaines coiled-coil. Cette protéine de 90 kDa est localisée au niveau du chloroplaste. Les mutant *mrc/pii1* et *mfp1* chez *Arabidopsis* ont le même phénotype que les mutants *ptst2* et *ss4* sur le nombre de grains d'amidon par plaste. On ne retrouve dans ces mutants qu'un seul grain, plus large que ceux de la lignée sauvage. La morphologie des grains de ces mutants est toutefois différente des mutants *ss4*, les premiers ayant des grains plats, alors que le dernier possède des grains sphériques. Les études d'interactions ont montré que PII1 interagissait à la fois avec SS4 et PTST2, alors que MFP1 interagit exclusivement avec PTST2. PII1 est aussi en mesure d'interagir avec SS5 (Abt et al., 2020). De ce fait, chez *Arabidopsis*, il a été hypothétisé que MFP1 est responsable de la localisation de PTST2 à la membrane des thylakoïdes, et que PII1 est responsable de la consolidation du complexe d'initiation de la synthèse du grain d'amidon ; composé à l'heure actuelle de SS4, PTST2, PTST3, MFP1 et PII1 (Seung et al., 2018; Vandromme et al., 2019). L'hypothèse proposée est que le complexe SS5-PII1, tout comme le complexe formé par SS4, PTST2/3, PII1 et MFP1, permettrait l'initiation de la synthèse de grain (Seung et al., 2017; Seung et al., 2018; Abt et al., 2020). Le complexe SS5-PII1 ne serait toutefois pas canonique à la synthèse de l'amidon chez les plantes sauvages. Le complexe formé avec SS4 serait privilégié lors de la synthèse de l'amidon (Figure 33).



**Figure 33. Les différentes voies possibles pour l'initiation de la synthèse du grain.** Les flèches en gras indiquent la voie de synthèse privilégiée. SS5 pouvant interagir avec elle-même, il est probable qu'elle soit sous forme de multimère avec PII1, mais cela n'a pour le moment pas été démontré. Ces différentes voies sont basées sur les publications de Szydlowski et coll. (2009), Seung et coll. (2017), Seung et coll. (2018), Vandromme et coll. (2019) et Abt et coll. (2020)(Szydlowski et al., 2009; Seung et al., 2017; Seung et al., 2018; Vandromme et al., 2019; Abt et al., 2020).

Le métabolisme de l'amidon est un mécanisme complexe, nécessitant l'intervention d'une dizaine d'enzymes. Ces enzymes agissent en synergie, le substrat d'une enzyme étant le produit d'une autre (Figure 34). Elles interagissent donc ensemble, de manière indirecte, mais aussi de manière directe comme l'a montré l'identification de complexes protéiques. Ces enzymes interagissent aussi avec le grain, et une partie d'entre-elles a déjà été retrouvée dans le protéome de grains d'amidon.



**Figure 34. Schéma résumé de la synthèse et de la dégradation de l'amidon.** (a) synthèse de l'amidon. 1. les amidons synthase allongent les chaînes de l'amylopectine en  $\alpha$ -1,4. 2. Les enzymes de branchements créent les points de branchements en  $\alpha$ -1,6. 3. Les enzymes de débranchements retirent certains points de branchements de manière à former les lamelles cristallines. (b) Dégradation de l'amidon. 4. Les dikinases vont décompacter les chaînes linéaires de l'amylopectine en les phosphorylant. 5. Les BAMs vont dégrader les chaînes linéaires à partir de l'extrémité non réductrice jusqu'à atteindre un phosphoester ou un point de branchement. 5. Les phosphatases vont retirer les groupements phosphates. 6. Les enzymes de débranchements impliqués dans la dégradation vont retrier les points de branchements en  $\alpha$ -1,6.

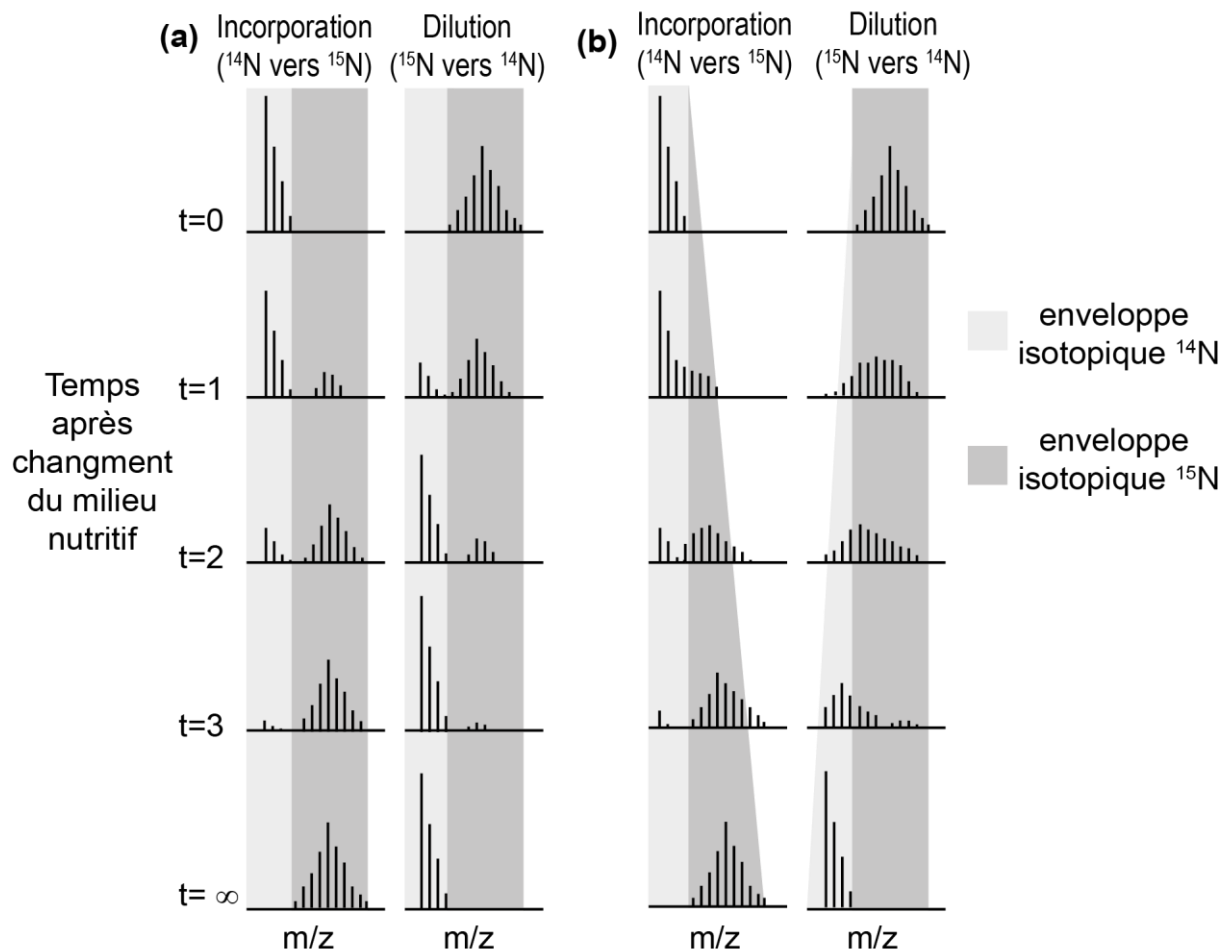


## V. L'analyse des protéines par spectrométrie de masse

Une partie des protéines associées au grain d'amidon ont été localisées à partir d'analyses protéomiques. La protéomique s'intéresse à identifier l'ensemble des protéines dans un organisme, un organe, un tissu, une cellule, ou un organite. C'est en 1996 que le terme « protéome » est utilisé pour la première fois, bien que le concept de protéomique existait depuis 1975 grâce au développement du gel d'électrophorèse en 2 dimensions (2D-PAGE)(O'Farrell, 1975; Wilkins et al., 1996). L'analyse par gel 2D permet des études comparatives sur la présence et de la quantité de l'ensemble des protéines entre plusieurs échantillons donnés. Ce type d'expérience, appelée 2D-DIGE (pour « 2 dimensional difference gel electrophoresis ») a notamment permis d'identifier de potentiels marqueurs du sucrage à froid chez la pomme de terre. Le sucrage à froid correspond à la production de sucres solubles dans les tubercules au cours de leur conservation à des températures inférieures à 6 °C (Fischer et al., 2013). Parmi ces marqueurs, on retrouve GBSS dont la quantité est corrélée positivement avec la sensibilité au sucrage à froid, ainsi que des inhibiteurs de protéase et de l'invertase qui seraient quant à eux corrélés négativement avec cette sensibilité. Dans ces travaux, les protéines ont été quantifiées par mesure de l'intensité des spots sur le gel 2D-PAGE (Fischer et al., 2013), alors que l'identification a été faite par découpe des spots et leur analyse en spectrométrie de masse après extraction et digestion en peptides. Cette identification ne permettait pas d'identifier des protéines inconnues de la banque de donnée utilisée. Ce type d'analyse est appelé bottom-up, dans le sens où l'on part du bas vers le haut, de la séquence du peptide vers la séquence d'une protéine entière connue. Si cette méthode est efficace pour identifier les protéines, elle demande en revanche beaucoup de travail en amont, et prend beaucoup de temps pour l'analyse des spots de plusieurs gels dans le cas d'une étude comparative. De plus, la quantification par mesure d'intensité des spots est peu reproductible entre deux gels et elle est biaisée par la possibilité d'avoir plusieurs protéines présentes dans un même spot (Zhang et al., 2013). Il est possible de réaliser l'identification et la quantification des protéines en même temps par la méthode de protéomique shotgun bottom-up. Dans cette méthode, L'ensemble des protéines de l'échantillon sont digérées en même temps, sans séparation préalable, puis les peptides sont analysés à l'aide d'un spectromètre de masse couplé à une chromatographie liquide, ou plus récemment à une électrophorèse capillaire. Cette méthode permet un gain de temps considérable lors de la préparation et l'analyse des échantillons par rapport à la méthode 2D-DIGE. La méthode shotgun permet par ailleurs la quantification relative des protéines sans marquage, à partir du nombre de spectres correspondant à un peptide (PSM, pour « peptide

spectrum matches »), et de l'intensité du signal du pic correspondant au peptide identifié sur chacun de ses spectres (Zhang et al., 2013). La validation de l'intégrité d'une protéine nécessite toutefois la réalisation d'un gel SDS-PAGE, qui permet de s'assurer dans une certaine mesure que la protéine est bien présente, qu'elle n'est pas tronquée ou dégradée.

L'approche protéomique permet en outre de suivre les variations quantitatives des protéines de manière dynamique dans un organisme vivant. Ceci peut être réalisé *in planta* par marquage avec un isotope stable. Cette stratégie est appelée SILIP (pour « Stable Isotope labelling *in planta* ») (Fan et al., 2016). La méthode est adaptée des méthodes SILAM et SILAC (pour « Stable Isotope Labelling with Amino acids in Mammals/Cell culture »), respectivement mises en œuvre chez les mammifères ou à l'aide de cultures cellulaires (Ong et al., 2002). Les deux premières méthodes ont été développées initialement pour réaliser des études quantitatives comparatives entre deux échantillons en mélange, l'un marqué avec des isotopes lourds stable et l'autre non (Zhang et al., 2013). Le suivi du taux d'incorporation d'isotope stable lors d'un marquage permet de suivre la dynamique de synthèse des protéines. Ce type d'expérience a déjà été réalisé chez l'organisme photosynthétique unicellulaire *Chlamidomonas reinhardtii* et chez la plante *Arabidopsis thaliana* pour établir le turnover des protéines (Sauer et al., 2014; Fan et al., 2016). Il apparaît que l'incorporation de l'azote ne suit pas le même cheminement chez ces deux organismes. En effet, chez *Chlamydomonas reinhardtii*, qui est un organisme unicellulaire, les massifs de pics évoluent de manière indépendante due à la faible proportion de matériel de réserve et de la consommation rapide des nutriments (Figure 35a)(Sauer et al., 2014). En revanche, chez les organismes complexes pluricellulaires comme *Arabidopsis thaliana* ou la pomme de terre, la présence de tissus de réserve et de différentes voies de synthèse implique un turnover différent, où le massif de pic correspondant aux protéines marquées se déplace vers les masses correspondantes au peptide totalement marqué dans le cas de l'incorporation, et inversement dans le cas de la dilution (Figure 35b)(Fan et al., 2016).



**Figure 35. Incorporation et dilution du marquage en fonction du temps après changement du milieu nutritif.** (a) Turnover des protéines chez une espèce photosynthétique dont les réserves d'acides aminés sont faibles et l'assimilation des nutriments rapide (cas de *Chlamydomonas reinhardtii*) (Sauer et al., 2014). Les schémas simulant des spectres de masses représentent l'évolution d'un même peptide dans des échantillons récoltés à différents temps. *Incorporation* ( $^{14}\text{N}$  vers  $^{15}\text{N}$ ) correspond à un organisme ayant été cultivé sur un milieu avec une abondance naturelle de chaque isotope, puis placé sur un milieu enrichi en  $^{15}\text{N}$ . *Dilution* ( $^{15}\text{N}$  vers  $^{14}\text{N}$ ) représente un organisme ayant été cultivé sur un milieu enrichi en  $^{15}\text{N}$ , puis placé sur un milieu avec une abondance naturelle des différents isotopes. (b) Turnover des protéines pour un organisme dont les réserves en acides aminés, ou des éléments de synthèses de ces acides aminés, sont élevées (cas de *Arabidopsis thaliana* et de la pomme de terre) (Fan et al., 2016). Schéma reproduit à partir de la publication de Fan et collaborateurs de 2016 (Fan et al., 2016).

Dans le domaine de l'amidon, les analyses du protéome associé au grain par spectrométrie de masse ont permis d'identifier certaines protéines (Rahman et al., 1995; Borén et al., 2004; Grimaud et al., 2008; Stensballe et al., 2008; Xing et al., 2016). Aujourd'hui, une bonne partie des protéines du métabolisme de l'amidon sont connues pour être associées au grain (Tableau 1).

**Tableau 1 protéines identifiées comme associées au grain d'amidon.**

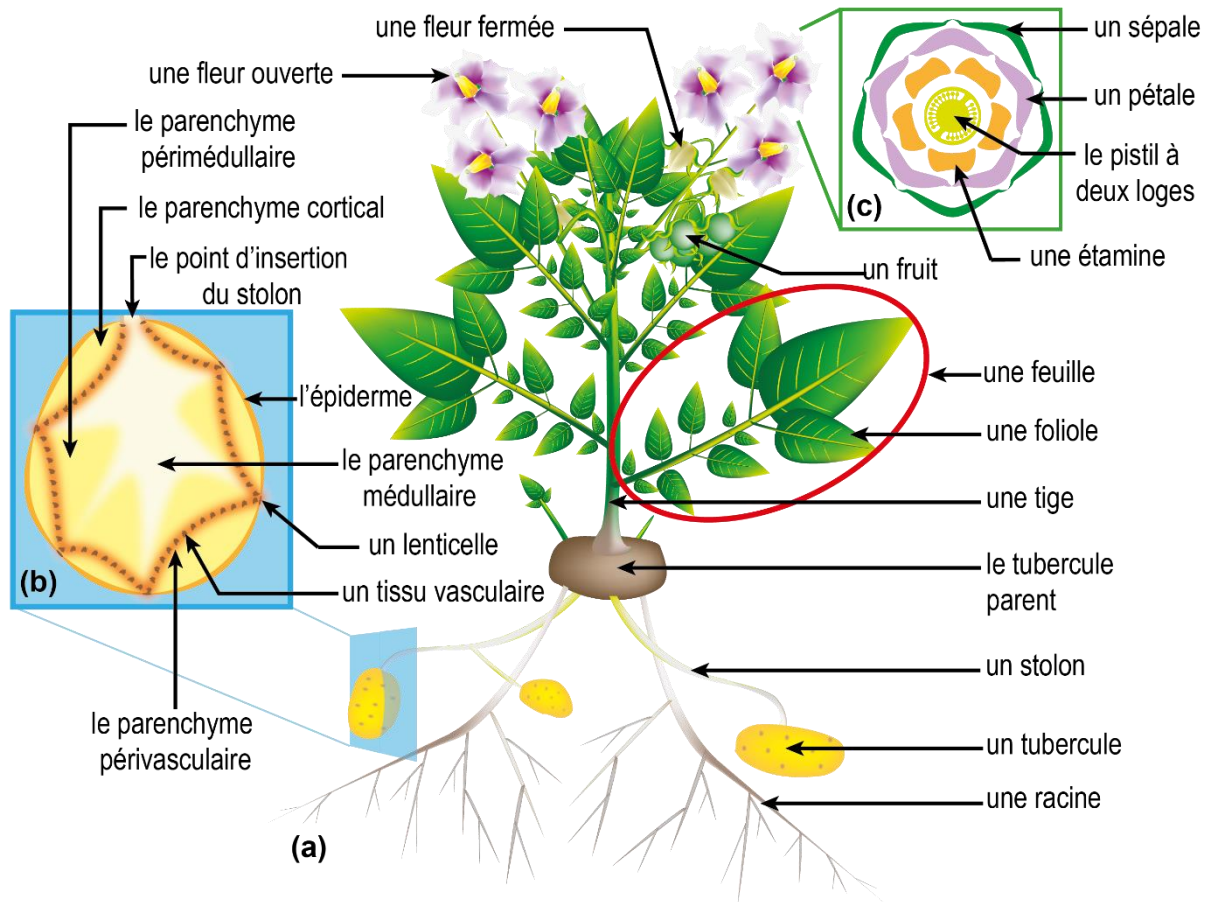
Espèces	Protéines identifiées comme associées au grain d'amidon	Références
Pomme de terre	GBSS; SS2 ; GWD	(Stensballe <i>et al.</i> , 2008)
Maïs	GBSS; SS1; SS2a ; SS3 ; BE2a ; BE2b; BE1 ; PHS1	(Grimaud <i>et al.</i> , 2008)
Riz	GBSS; SS1; SS2a, SS3, BE2b; PHS1; PUL; ISA-2	(Xing <i>et al.</i> , 2016)
Pois	GBSS; SS2; BE2; BE1; GWD	(Denyer <i>et al.</i> , 1993); (Ritte <i>et al.</i> , 2002)
Orge	GBSS; SS1; SS2a; BE2b; BE1c	(Boren <i>et al.</i> , 2004); (Peng <i>et al.</i> , 2000)
Blé	GBSS; SS1; SS2a; BE2a; BE2b; BE2c	(Rahman <i>et al.</i> , 1995); (Regina <i>et al.</i> , 2005)
Arabidopsis	PTST1, PTST2, PTST3, ESV1, LESV, SEX4; LSF2	(Seung <i>et al.</i> , 2015); (Feike <i>et al.</i> , 2016); (Seung <i>et al.</i> , 2017); (Sokolov <i>et al.</i> , 2006); (Santelia <i>et al.</i> , 2011)

## VI. La pomme de terre (*Solanum tuberosum*)

La pomme de terre, ou *Solanum tuberosum*, est une plante de la famille des solanacées, qui regroupe une grande famille de plantes, toutes dicotylédones, de morphologies très diverses. On retrouve dans cette famille comme plante d'intérêt économique, hormis la pomme de terre, la tomate (*Solanum lycopersicum*), l'aubergine (*Solanum melongena*) et les piments (*Capsicum*). La pomme de terre est originaire de la cordillère des Andes, ou elle faisait partie, avec le maïs, des aliments de base de l'empire Inca. Elle fut importée en Europe au XVIème siècle à la suite de la conquête de l'Amérique. Sa culture en Europe ne commencera cependant qu'au début du XVIIème siècle et de manière sporadique, avec une culture officielle en France en 1620. C'est à la fin du XVIIIème siècle que la pomme de terre connaîtra un engouement dans toute l'Europe. Du fait de son haut rendement nutritif et de son faible coût, elle participera à la libération des Etats de la disette, améliorant ainsi la qualité de vie des peuples Occidentaux. Au XIXème siècle, la pomme de terre commence sa conquête du monde via les

empire colonial. Elle est aujourd'hui consommée et cultivée sur l'ensemble du globe. Elle pourrait dépasser les limites de la Terre avant la fin du XXI<sup>ème</sup> siècle, plusieurs projets prévoyant sa culture sur la station spatiale internationale et sur le sol de Mars (Wheeler, 2017). La feuille de pomme de terre est composée d'un nombre impair de folioles, le plus souvent 7 à 9 grands folioles par feuilles, avec de petites folioles s'intercalant par paire entre les grandes folioles (Figure 32). La lymphe des feuilles possède des trichomes à sa surface. Les tiges aériennes de la pomme de terre sont succulentes et peuvent être relativement lignifiées en fin de développement à la base. Elles peuvent atteindre 1 m de hauteur. L'inflorescence de la pomme de terre naît à l'extrémité de la tige et possède en général 7 à 15 fleurs. Les fleurs ont un diamètre de 4 à 5 centimètres, avec une symétrie radiale en pentamère, composée, de l'extérieur vers l'intérieur, de 5 sépales, 5 pétales, 5 étamines et d'un pistil central composé d'un ovaire supère à 2 loges (Figure 32). Les fruits de la pomme de terre sont des baies vertes ressemblant à des tomates cerises. Ils contiennent toutefois des alcaloïdes et des protéines toxiques, et ne sont donc pas comestibles. Les tiges souterraines, appelés stolons, se développent plus tardivement, et marquent le début de la tubérisation. La caractéristique la plus marquante chez la pomme de terre, et qui fait tout son intérêt économique, est sa capacité à produire des tubercules à partir de ces tiges souterraines. Ces tubercules contiennent beaucoup d'amidon, faisant de la pomme de terre l'espèce végétale ayant le meilleur rendement en termes d'amidon produit par surface de terre. Les tubercules de la pomme de terre lui permettent d'être pérenne et de survivre à l'hiver, malgré le fait qu'elle soit caduque et perde ses feuilles et ses tiges aériennes lors des saisons froides. C'est à partir de ce tubercule que vont se développer les tiges et les racines de la pomme de terre par germination au retour à des températures clémentes après une période de dormance plus ou moins longue selon les variétés. Sur la coupe transversale d'un tubercule, on distingue, de l'extérieur vers l'intérieur, l'épiderme ou la peau, qui protège le tubercule, le parenchyme cortical, l'anneau vasculaire, le parenchyme périvasculaire, le parenchyme péri-médullaire et le parenchyme médullaire (Figure 36). L'anneau vasculaire, ainsi que le parenchyme médullaire, tous deux contenant une grande quantité de tissus vasculaires, xylème et phloème, prennent naissance au niveau du talon, zone interstitielle entre le stolon et le tubercule, et se rejoignent au niveau de l'épiderme sur des zones appelées lenticelles, qui sont les points de bourgeonnement du tubercule lors de la germination. L'ensemble des parenchymes possèdent en leur sein une très grande quantité d'amidon. Ce dernier représente environ 18% du poids frais d'un tubercule de pomme de terre, et environ 90% de son poids sec (le tubercule de pomme de terre étant composé à 80% d'eau).

Le tubercule pomme de terre produit des grains d'amidon avec une grande hétérogénéité de taille, les plus petits mesurant 5  $\mu\text{m}$  et les plus gros pouvant atteindre 100  $\mu\text{m}$ .



**Figure 36. Schéma d'un plant de pomme de terre.** (a) Dessin d'un plant de pomme de terre. (b) Schéma d'une coupe transversale de tubercule de pomme de terre. (c) diagramme floral de la fleur de pomme de terre.

## VII. Objectifs de la thèse

### 1. Établissement du protéome complet de l'amidon de pomme de terre

Le premier objectif de la thèse fut d'établir le protéome complet de l'amidon de pomme de terre de manière à identifier l'ensemble des protéines associées au grain d'amidon. Avant le début de ce projet, le premier protéome établi sur l'amidon de pomme de terre avait permis d'identifier trois protéines associées au grain d'amidon : GBSS, SS2 et GWD (Stensballe et al., 2008). Toutefois, d'autres études réalisées sur le pois, le blé, l'orge et le maïs ont révélé la présence d'autres enzymes au sein du grain d'amidon, telles que SS1, SS3, les enzymes de branchement, les phosphorylases et les phosphatases (Denyer et al., 1993; Peng et al., 2000;

Borén et al., 2004; Grimaud et al., 2008; Xing et al., 2016). Qui plus est, les récentes études des protéines PTST1, ESV1 et LESV chez *Arabidopsis thaliana* ont révélé que ces protéines étaient présentes dans la fraction insoluble des plastes, et donc dans le grain d'amidon (Seung et al., 2015; Feike et al., 2016). De plus, ces investigations ont montré que des protéines non-catalytiques associées au grain d'amidon pouvaient aussi être impliquées dans son métabolisme et dans sa régulation, et que certaines protéines minoritaires pouvaient encore être caractérisées. Par conséquent, établir le protéome exhaustif de l'amidon devrait mener à l'identification de fonctions encore inconnues dans le métabolisme de l'amidon. Le génome de *Solanum Phureja*, une pomme de terre diploïde, a été séquencé en 2011 (Xu et al., 2011). De plus, 77 locus génomiques codant pour des enzymes du métabolisme de l'amidon ont été annotés, avec plusieurs nouvelles isoformes potentielles (Van Harsselaar et al., 2017). Grâce à ces puissants outils génomiques à notre disposition et aux nouvelles découvertes faites avec *Arabidopsis thaliana*, nous avons pu établir une banque de données complète pour l'identification des protéines liées au grain d'amidon. Grâce aux méthodes de protéomique quantitative, nous avons aussi pu établir la stœchiométrie de ces protéines, et étudier l'évolution de leur contenu quand la surface des grains est préalablement traitée pour en retirer les protéines adsorbées.

## **2. Fluctuation de la stœchiométrie basée sur la morphologie des grains.**

Une fois le protéome complet de l'amidon de pomme de terre établi, ainsi que sa stœchiométrie, nous nous sommes intéressés à la stabilité de ce protéome. Nous avons cherché à savoir si les protéines identifiées étaient toujours les mêmes en fonction de la morphologie des grains et si la concentration de chaque protéine prise individuellement fluctuait à l'intérieur du grain d'amidon. En effet, l'amidon montre une forte hétérogénéité structurale et morphologique. Cette disparité s'observe au niveau de l'origine botanique, de la localisation tissulaire dans une espèce donnée, ainsi qu'au sein d'un même plaste, avec par exemple l'amidon de blé qui possède deux types de grains (A et B) au sein du même plaste (Tanaka et al., 2017). Les grains de type A étant localisés au centre du plaste et étant beaucoup plus gros que les grains de type B localisés à l'extrémité des stromules, des extensions de la membrane des plastes (Tanaka et al., 2017). Chez la pomme de terre, on retrouve une grande diversité de morphologie de grains avec des diamètres variant de 5 à 100  $\mu\text{m}$  (Tester et al., 2004) Une étude précédente indiquait l'existence de variation de structures et de composition entre des grains de grande taille et ceux de petite taille (Wang et al., 2016). Les premiers avaient plus d'amylose et de chaînes longues ( $\geq$  DP 37), alors que les deuxièmes ont plus de courtes chaînes et sont de

type cristallin A les autres grains étant de type B. De manière à évaluer l'hétérogénéité potentielle des protéines liées au grain d'amidon en fonction de la structure et de la morphologie des grains, nous avons étudié la taille et la forme des grains ainsi que la composition et la structure de grains d'amidon fractionnés en fonction de leur diamètre moyen. L'amidon a été séparé en 4 fractions distinctes : A, B, C et D (A étant la fraction contenant les grains les plus petits et D contenant les grains les plus gros). Les contenus en amylose, la cristallinité de l'amidon, la distribution des longueurs des chaînes de l'amylopectine et le contenu en phosphate ont été déterminés ainsi que le contenu en protéines et leur stœchiométrie.

### **3. Dynamique d'association des protéines au grain d'amidon**

L'étude réalisée sur la stœchiométrie des protéines liées du grain en fonction de sa morphologie a permis de montrer qu'il existait bien des fluctuations dans le contenu en protéines en fonction de la taille des grains. Cette découverte a soulevé la question de savoir si ces fluctuations étaient des indicateurs d'un état de croissance du grain d'amidon, avec comme postulat que les petits grains sont des grains en début de croissance et que les gros grains sont en fin de croissance. Ces fluctuations peuvent aussi être indicatrices d'un déterminisme, chaque taille ayant un protéome propre. De manière à pouvoir répondre à cette question, le développement de méthodes de suivi de l'association des protéines au grain d'amidon fut nécessaire. Les protéines sont synthétisées chez les plantes à partir de la photosynthèse et des sels de nitrates, ces derniers constituant la source d'azote essentielle à la synthèse des acides aminés (Xuan et al., 2017). De manière à pouvoir suivre la synthèse et l'incorporation des protéines au grain d'amidon de pomme de terre, nous avons donc réalisé des expériences de marquage à l'azote 15 ( $^{15}\text{N}$ ). Le  $^{15}\text{N}$  est un isotope stable que l'on retrouve naturellement en très faible quantité dans la nature, l'isotope  $^{14}\text{N}$  représentant 99,64 % de l'azote présent sur terre. Le fait d'enrichir un milieu en sels de nitrates marqué au  $^{15}\text{N}$  à un temps donné permet de suivre la synthèse et l'incorporation dans le grain des protéines par spectrométrie de masse du fait du glissement des massifs de pics correspondant aux peptides marqués. Pour déterminer à quel moment de la tubérisation les protéines du métabolisme de l'amidon sont synthétisées et se lient au grain, nous avons réalisé trois expériences de marquage au  $^{15}\text{N}$ . Dans la première expérience, le marquage est effectué au début de la tubérisation, lorsque les stolons viennent d'apparaître, et est arrêté au milieu de la tubérisation. Dans la deuxième expérience, le marquage est effectué quelques temps après l'apparition des stolons et arrêté quelques temps



avant la récolte. Enfin, dans la 3<sup>ème</sup> expérience, le marquage est commencé au milieu de la tubérisation et continue jusqu'à la récolte.

#### **4. Impact d'une protéine sur l'ensemble du protéome, le cas des mutants *gbss***

La plupart des protéines du métabolisme de l'amidon n'agissent pas de manière individuelle mais sous forme de complexe. La présence ou l'absence d'une protéine devrait par conséquent modifier la stœchiométrie de l'ensemble du protéome. Au sein du grain d'amidon, GBSS est la protéine majoritaire (Grimaud et al., 2008). L'inactivation de l'expression de GBSS devrait donc en toute logique diminuer drastiquement la quantité des protéines liées à l'amidon. Il fut récemment montré chez *Arabidopsis* que GBSS est adressée au grain par PTST1 (Seung et al., 2015). Bien que majoritairement retrouvée dans la fraction soluble, PTST1 est aussi associée au grain. L'hypothèse émise est que la localisation de PTST1 oscille entre le stroma et le grain pour amener GBSS à son substrat. PTST1 ne serait liée au grain que lorsqu'elle est elle-même complexée à GBSS. Si c'est bien le cas, l'absence de GBSS devrait aussi entraîner une diminution, voire une absence de PTST1 dans le grain. La mutation *sex1*, qui entraîne la perte de l'activité de GWD, entraîne aussi une augmentation significative du contenu en amylose (Yu et al., 2001a). Il est donc possible qu'il existe un lien entre l'activité de GBSS et l'activité de GWD. De manière à pouvoir étudier l'influence de GBSS sur l'ensemble du protéome, des mutants de pomme de terre cv. Désirée avec des niveaux d'expression variables de GBSS ont été sélectionnés (Veillet et al., 2019). Ces mutants ont été produits par l'IGEPP du CRG de l'INRA de Ploudaniel avec qui nous collaborons sur ce sujet. La morphologie des grains, la distribution de la longueur de chaînes de l'amylopectine, le contenu en phosphate et le protéome de ces amidons ont été étudiés. La quantification de l'ensemble des protéines dans chaque cas a permis de comparer les quantités de chacune en présence et en absence de GBSS.

#### **5. Protéomique à l'échelle du grain d'amidon**

La synthèse d'un grain d'amidon est un évènement nécessitant une prise de décision cellulaire pour pouvoir être initié, et un enchainement d'évènements d'élongation, de branchements et de débranchements permettant d'aboutir à un grain. Un certain nombre d'expérimentations ont déjà été mises en place pour simuler la formation de points de branchement par « chimie clic » ou « click-chemistry » (Harada et al., 1972; Marmuse et al., 2005; Putaux et al., 2006; Kaneko et al., 2007; Nepogodiev et al., 2007; Yanase et al., 2007; Grimaud et al., 2013). De ces expériences ont été synthétisés de longs polymères non branchés,

imitant l'amylose, ou des particules dendritiques branchées ou non branchées synthétisées à partir d'une base de glycogène enrobant une nanoparticule (O'Neill and Field, 2015). Des sphérulites imitant l'amidon ont aussi pu être synthétisés à partir d'amylopectine et d'amylose commerciaux de pomme de terre (Ma et al., 2011). Ces deux dernières méthodes ont permis de définir des systèmes 3D pour l'étude des enzymes de synthèse de l'amidon à l'échelle d'un événement de synthèse. Toutefois, ces modèles restent limités car la disposition des glucanes ne reflète pas celle d'un grain d'amidon natif. Une autre manière de travailler est de réaliser la synthèse de l'amidon dans un organisme qui n'est normalement pas en mesure de le synthétiser en insérant des gènes dans son ADN. Ce type d'expérience a d'ores et déjà été réalisé chez la levure *Saccharomyces cerevisiae* par insertions de séquences codant des protéines du métabolisme de l'amidon d'*Arabidopsis thaliana* (Pfister et al., 2016). Le modèle a d'abord été délesté de l'ensemble des gènes permettant la synthèse du glycogène, puis ont été inséré les séquences de l'AGPase d'*Escherichia coli*, les amidon-synthases SS1 à 4 d'*Arabidopsis thaliana*, ainsi que BE2.1, BE2.2, ISA1 et ISA2 de la même espèce. La levure ainsi modifiée était en mesure de produire une matrice de glucanes insolubles de structure proche de celle du grain d'amidon. Le but de l'ensemble de ces expérimentations est, *in fine*, d'être en mesure de synthétiser de l'amidon *in vitro*.

Travailler à l'échelle du grain unique permettrait de mieux comprendre les événements de sa synthèse et, en parallèle d'autres expériences citées précédemment et d'études cinétiques de chaque enzyme, d'établir un modèle mathématique permettant la synthèse de l'amidon *in vitro*. La stœchiométrie des protéines n'est pas constante selon la morphologie des grains d'amidon. Toutefois, établir si la stœchiométrie reste la même pour des grains de même morphologie passe par l'analyse à l'échelle d'un seul grain. En effet, les grains d'amidon sont synthétisés dans le tubercule à différents instants, dans différents tissus, sous différentes conditions de stress (les conditions de stress peuvent varier au long du développement du tubercule). Par conséquent, il est probable que des événements de synthèse distincts puissent aboutir à une même morphologie de grain. Pour le moment, les analyses que nous avons réalisées ont été faites à partir d'une quantité conséquente d'amidon (300 mg). Cela apporte un biais, du fait que la stœchiométrie de chaque grain ne peut pas être distinguée, et que les résultats représentent une moyenne de l'hétérogénéité présente dans l'échantillon.

Grâce à des calculs théoriques, nous avons pu établir que, pour un grain d'amidon de grande taille (100  $\mu\text{m}$ ), en estimant la quantité de protéines à l'intérieur de ce grain à  $2.2 \mu\text{g}.\text{mg}^{-1}$  et la densité d'un grain d'amidon à  $1,5 \text{ g}.\text{cm}^{-3}$  (le poids du grain est donc estimé à  $0,785 \mu\text{g}$ ), les cinq

protéines majoritaires de l'amidon seraient toujours détectables avec les approches de spectrométrie de masse que nous utilisons au laboratoire. En effet, les spectromètres de masse à transformée de Fourier Orbitrap, tel que le Q-Exactive plus que nous utilisons ont une résolution permettant de détecter des molécules à l'échelle de l'attomole (amol), le minimum de détection étant environ à 100 amol. Grâce aux expériences précédentes ayant permis d'estimer la stœchiométrie des protéines de l'amidon de pomme de terre, nous pouvons estimer que GBSS représente 80 % des protéines du grain, SS2 10 %, GWD, SS3 et LESV 1 %. En termes de quantité de matière, cela représente 20 786 amol pour GBSS, 2 028 amol pour SS2, 106 amol pour GWD, 124 amol pour SS3 et 272 amol pour LESV. GBSS et SS2 devraient donc être aisément détectables à l'échelle d'un grain alors que SS3, GWD et LESV devraient se trouver à la limite de détection.

# Matériels et Méthodes

## **I. Matériels**

Toutes les solutions ont été préparées avec de l'eau ultrapure purifiée avec un système MilliQ™ Academic (Merck Millipore, Burlington, MA, USA). Tous les produits chimiques et solvants ont été achetés chez Merck, anciennement Sigma Aldrich (Saint-Louis, MO, USA) et utilisés tels quels sans purification.

### **1. Amidon utilisé pour l'étude du protéome complet, de la stœchiométrie des protéines liées au grain et pour les études grain unique.**

Les grains d'amidon ont été isolés de tubercules de *Solanum tuberosum*, cultivar (cv.) Monalisa, cultivés en pleine terre à Villeneuve d'Ascq (50.607735, 3.143431), France, entre Mars et Juillet 2014. Les tubercules (environ 70 g) ont été lavés à l'eau du robinet puis pelés avant extraction. Les tubercules ont été broyés avec une centrifugeuse à fruits dans 200 mL d'eau ultrapure. Les extraits obtenus sont ensuite filtrés sur filtres en nylon d'une largeur de maille de 100 µm. L'extrait filtré est ensuite laissé à sédimenter pendant 3h. Le surnageant est ensuite jeté et le culot d'amidon est resuspendu dans 500 mL d'eau Milli-Q. L'amidon a été lavé 3 fois successivement avec 1 L d'eau Milli-Q et stocké dans l'éthanol 20% à 4 °C.

### **2. Amidon utilisé pour les études de dynamiques d'association des protéines au grain**

Les grains d'amidon ont été isolés de tubercules de *Solanum tuberosum*, cultivar Désirée, produits à partir d'explants venant du Centre de ressource biologique BrACySol (IGEPP INRA). Les plantules ont été cultivées sur un milieu 19 Murashig et Skoog (MS) incluant des vitamines et 8 g.L<sup>-1</sup> d'Agar de chez Duchefa®. Les plantules ont grandi en chambres de croissance GroBank avec 16h de lumière (80 µmol.m<sup>-2</sup>.s<sup>-1</sup>) et 8h d'obscurité pendant 3 semaines à 20 °C. Les plantules ont été repiquées toute les 3 semaines par transplantation de la partie terminale coupée au niveau d'un méristème secondaire. Les plantules de 3 semaines sont ensuite transplantées dans des pots de 2 L avec un terreau pauvre en nitrate. 5 conditions ont été définies pour cette expérience, et 2 plantes par conditions ont été utilisées. 2 plantes ont été utilisées comme contrôles négatifs et ont été arrosées avec du milieu MS incluant des vitamines de chez Duchefa® pendant 10 semaines. 2 autres plantes servant de contrôle positif, ont été arrosées avec un milieu MS maison incluant des vitamines de même composition que le milieu MS 19 Duchefa® mais dont les sels de nitrates (KNO<sub>3</sub>, NH<sub>4</sub>NO<sub>3</sub>) ont été enrichis en <sup>15</sup>N. 6 autres plantes utilisées pour les expériences d'incorporation et de chasse ont d'abord été arrosées avec un milieu MS incluant des vitamines (Duchefa®), avec des isotopes naturels,

pendant 4 semaines. Après cette période les stolons ont commencé à apparaître, indiquant l'imminence de la tubérisation. A partir de ce moment, 2 plantes ont été arrosées avec les milieux MS enrichi en  $^{15}\text{N}$  pendant 3 semaines. Ensuite, les nutriments dans le terreau ont été lessivés avec une grande quantité d'eau, puis les plantes ont été arrosées avec du milieu MS contenant des isotopes naturels pendant 3 semaines. 2 autres plantes ont été arrosées avec un milieu MS contenant des isotopes naturels pendant 3 semaines. Les nutriments dans le terreau ont été lessivés avec une grande quantité d'eau, puis les plantes ont été arrosées avec du milieu MS enrichi en  $^{15}\text{N}$  pendant 3 semaines. Enfin, les 2 plantes restantes ont d'abord été arrosées avec des isotopes naturels pendant 1,5 semaines, puis arrosées avec du milieu MS enrichi en  $^{15}\text{N}$  pendant 3 semaines après lessivage des nutriments. Enfin, les plantes sont arrosées les 1,5 semaines restantes avec un milieu MS contenant des isotopes naturels après que le terreau ait été une nouvelle fois lessivé. Après ces 10 semaines, les tubercules ont été récoltés et l'amidon a été extrait des tubercules (environ 10 g de tubercule par plantes). Les tubercules ont été lavés à l'eau du robinet, puis pelés et coupés en cubes d'environ 1 cm de côté. Les dés de tubercules sont ensuite écrasés au pilon avec environ 20 mL d'éthanol 20 % (v/v). Les extraits de tubercules obtenus sont ensuite filtrés sur filtre de nylon de 100  $\mu\text{m}$  de largeur de maille, le filtrat est transféré en tube de 50 mL. Les extraits filtrés sont ensuite centrifugés à 5,000 g pendant 15 min. Le surnageant est ensuite retiré et le culot est resuspendu dans 50 mL d'eau Milli-Q. La suspension d'amidon est ainsi lavée successivement 3 fois à l'éthanol 20 % (v/v), 5 fois au SDS 2 % (m/v) et 3 fois à l'éthanol 20 % (v/v). Environ 1 g d'amidon de chaque plante a ensuite été séché à l'étuve à 40 °C.

### **3. Amidon utilisé pour l'étude de l'influence de GBSS sur l'ensemble du protéome**

Les amidons utilisés pour l'étude de l'influence de GBSS sont issus de pommes de terre cv. Désirée ayant été mutées pour le gène codant GBSS. Ces pommes de terre ont été fournies sous formes de tubercules par le GRC BraCySol (IGEPP) INRA de Ploudaniel dans le cadre d'une collaboration (Veillet et al., 2019). Les mutants choisis pour cette étude sont les mutants 17T.716.146 et 17T.701.010. L'amidon a été extrait des tubercules de la même manière que décrit précédemment en section 2 : « Amidon utilisé pour les études de dynamiques d'association des protéines au grain ».

## **II. Méthodes**

### **1. Traitements des grains d'amidon**

Les grains d'amidon purifiés de *Solanum tuberosum* (cv. Monalisa) ont été soumis à différents traitements de manière à retirer les protéines de surface. L'un des traitements effectués fut de les laver 5 fois au SDS 2% qui va dénaturer et décrocher les protéines en interaction avec la surface des grains. L'autre traitement a été de les incuber en présence de thermolysine qui va digérer les protéines de surface.

### **2. Morphologie des grains d'amidon.**

La morphologie des grains d'amidon est déterminée par granulomorphométrie avec un granulomorphomètre Ipac 1 (Occhio, Angleur, Belgique). Les paramètres retenus pour la mesure de taille et de forme des grains sont l'« ISO Area Diameter » et l'« Ellipsoid roundness ». Après la mesure, un filtre numérique a été appliqué. Celui-ci exclut les particules dont le diamètre est inférieur à 4  $\mu\text{m}$ . De plus, les particules combinant une rotundité ellipsoïde inférieure à 46 % et un diamètre compris entre 4 et 18  $\mu\text{m}$  (nous avons déterminé empiriquement que ces particules sont en quasi-totalité des débris cellulaires) ont aussi été exclues. Environ 3 000 grains ont été analysés pour chaque expérience.

### **3. Fractionnement des grains d'amidon en fonction de leur taille.**

Une partie des grains d'amidon issus de *Solanum tuberosum* cv. Monalisa ont été fractionnés à l'aide de filtres de nylons avec une largeur de maille de 50, 31 et 11  $\mu\text{m}$ . Les filtrations sont réalisées de manière séquentielle en commençant par le filtre 11  $\mu\text{m}$ . Les fractions non retenues sur les filtres sont séquentiellement nommées fraction A, B et C. La fraction retenue sur le filtre 50  $\mu\text{m}$  est nommée fraction D. La qualité du fractionnement est estimée par mesure de la morphologie des grains avec un granulomorphomètre Ipac 1 (Occhio, Angleur, Belgique). Les paramètres utilisés sont les mêmes que ceux décrits précédemment dans la section 2.

### **4. Capture des grains d'amidon en capillaire pour les analyses avec 30 grains et 1 grain.**

Les suspensions de grains d'amidons issus de la fraction D de *Solanum tuberosum* cv. Monalisa ont été dilués de manière successive dans de l'éthanol 20 % (v/v) jusqu'à être capable

de dénombrer les grains au microscope dans une goutte de 20  $\mu\text{L}$ . Dans le cas de la capture de 30 grains d'amidon, la goutte est laissée à sécher quelques minutes de manière à concentrer les grains d'amidon. Dans le cas de la capture d'un grain unique, d'autres dilutions sont réalisées jusqu'à être en mesure d'isoler un grain. Le(s) grain(s) a/ont été capturé(s) à l'aide d'une colonne capillaire transparente en silice (diamètre interne : 100  $\mu\text{m}$ , diamètre externe : 360  $\mu\text{m}$ , longueur : 2 cm). La colonne est mise en contact avec la goutte de suspension d'amidon. Par capillarité, le liquide à la surface de la colonne entre à l'intérieur et le(s) grain(s) à proximité sont emportés avec le liquide. Les grains ainsi capturés ont été contrôlés, dénombrés et mesurés à l'aide d'un microscope optique (Optika, Italie), et du logiciel Image J.

### **5. Imagerie des grains d'amidon par microscopie électronique à balayage.**

L'imagerie par microscopie électronique à balayage a été réalisée au sein du CERVAM de Grenoble par Jean-Luc PUTAUX. Une goutte de suspension d'amidon de *Solanum tuberosum* (cv. Monalisa), préalablement fractionnée et diluée, a été déposée sur une plaque métallique couverte avec une bande C/Ni et laissée à sécher. Les échantillons ont ensuite été recouverts avec du Au/Pd dans un pulvérisateur cathodique Baltec MED 020 (Leica, Wetzlar, Allemagne) et observés en mode « secondary electron » sur un microscope FEI Quanta FEG 250 (Hillsboro, OR, USA) opérant à une tension d'accélération de 2.5 kV.

### **6. Etude de la cristallinité des grains d'amidons par diffraction aux rayons X.**

L'étude de la cristallinité par diffraction aux rayons X a été réalisée au sein du CERVAM de Grenoble par Jean-Luc PUTAUX. Les fractions de grains ont été placées dans une chambre maintenant une humidité relative de 93 % pendant 5 jours. Les poudres sont ensuite versées dans des capillaires en verres de 0.7 mm (diamètre externe) scellées à la flamme. Tous les échantillons sont passés aux rayons-X par transmission dans une chambre à vide Warhus avec une radiation  $\text{CuK}\alpha$  filtré Ni ( $\lambda=0.1542$  nm) en utilisant un générateur Philips PW3830 (Eindhoven, Pays-Bas) opérant à 30 kV et 20 mA. Les motifs 2 dimensions de diffractions aux rayons X sont enregistrés sur des plaques d'imagerie Fujifilm exposées pendant 2h et lues hors-ligne avec analyseur de bioimagerie Fujifilm BAS 1800-II (Minato, Tokyo, Japon). Les profils de diffractions sont calculés avec les moyennes de rotations des motifs 2 dimensions.



## 7. Dosage de l'amylose.

Le dosage de l'Amylose a été adapté à partir du kit k-Amyl Megazyme (Wicklow, Irlande) (Gibson et al., 1997). 1.5 mg d'amidon a été gélatinisé dans du diméthylsulfoxyde (DMSO) 100 %, et précipité dans l'éthanol 95 % de manière à retirer les lipides. L'amidon est ensuite resolubilisé dans le DMSO 100 %. L'amylopectine est précipitée avec la concanavaleine A (ConA) en suivant les instructions du manuel. Après centrifugation, le surnageant contenant l'amylose a été récupéré, et le culot contenant l'amylopectine a été resolubilisé dans le DMSO 100 %. La ConA a ensuite été désactivée par chauffage à 100 °C. L'Amylose, l'amylopectine et l'amidon total ont ensuite été digérés à l'aide d'alpha-amylases et d'amyloglucosidases, puis le D-Glucose a été dosé à partir du réactif de détermination du glucose (GOPOD) disponible dans le kit K-AMYL de Megazyme. Ce réactif contient une glucose oxydase, une peroxydase et de l'acide p-hydroxybenzoïque. L'absorbance est lue à 510 nm avec un spectrophotomètre SPECTROstar Nano (BMG LABTECH's, Allmendgrün, Allemagne). En parallèle, une coloration à l'iode est réalisée sur les fractions d'amylose, d'amylopectine et d'amidon total. L'absorbance a été lue entre 500 et 700 nm pour déterminer la longueur d'onde au maximum d'absorbance ( $\lambda_{max}$ ) pour tous les complexes glucanes-I<sub>2</sub>.

## 8. Distribution des longueurs de chaînes d'amylopectine.

La distribution de longueurs de chaînes est faite à partir de 2 mg d'amidon purifié. Cet amidon est dans un premier temps gélatinisé par chauffage à 100°C dans l'eau pendant 10 minutes en mélangeant régulièrement à l'aide d'un vortex.

La distribution de longueur de chaînes de l'amidon a été déterminée par électrophorèse capillaire assisté par fluorescence (FACE). 2 milligrammes d'amidon de chaque fraction, en suspension dans 250 µL d'eau Milli-Q, ont été gélatinisés par chauffage à 100 °C, avec homogénéisation régulière au vortex, jusqu'à ce que la solution devienne limpide. Après cela, 20 unités (U) d'isoamylase et 1 U de pullulanase ont été ajoutées et la solution fut incubée à 42 °C toute la nuit. Les Protéines ont ensuite été précipitées par chauffage à 95 °C pendant 15 minutes. Le surnageant a été récupéré après centrifugation. Les échantillons ont été dessalés sur une colonne d'extraction en phase solide (SPE), puis lyophilisés. Les échantillons secs ont été resolubilisés dans 2 µL de cyanoborohydrure de sodium 1 M dans du tétrahydrofurane (THF), et 2 µL d'acide 8-aminopyrene-1,3,6-trisulfonique (APTS) à 104,16 mg.mL<sup>-1</sup> dans l'acide acétique 15 % (v/v). Les échantillons furent incubés toute la nuit à 42 °C, et dilués avec 46 µL

d'eau Milli-Q. Après la première dilution, les échantillons ont été redilués 50 fois dans l'eau Milli-Q avant injection dans une électrophorèse capillaire Beckman Coulter PA800-plus Pharmaceutical Analysis System (Beckman Coulter, Brea, CA, USA) équipée d'une colonne capillaire en silice (diamètre interne (50  $\mu\text{m}$ ; diamètre externe: 360  $\mu\text{m}$ ; longueur: 60 cm). Le capillaire a été rincé et rempli de gel Beckman Coulter *N*-linked carbohydrate separation gel buffer dilué au 1/3 avec de l'eau Milli-Q avant injection (3 minutes à 30 psi). Les échantillons ont été injectés par pression à 0,5 psi pendant 10 sec. La migration a été effectuée à 30 kV pendant 1h à 20 °C. L'excitation a été faite avec un module laser (LIF) à une longueur d'onde de 488 nm, et l'émission à 512 nm a été mesurée à l'aide d'un détecteur de fluorescence. La fenêtre de détection est localisée 10 cm au-dessus de la sortie de la colonne capillaire.

## **9. Contenus en phosphates.**

Le contenu en phosphates pour chaque fraction a été déterminé par FACE. 1 mg d'amidon de chaque fraction a subi une hydrolyse acide avec 100  $\mu\text{L}$  d'acide trifluoroacétique (TFA) 2 M, à 95 °C pendant 1 h et 20 min, précisément. Un total de 20  $\mu\text{L}$  de l'hydrolysate fut ensuite dilué avec 200  $\mu\text{L}$  d'eau Milli-Q pour arrêter la réaction. L'échantillon a ensuite été séché au concentrateur SpeedVac™ (Eppendorf™ Concentrator Plus, Eppendorf, Hamburg, Germany). Les échantillons ainsi séchés furent ensuite dérivatisés avec de l'APTS comme décrit précédemment dans la section 8 : "distribution des longueurs de chaînes d'amylopectine". Les échantillons dérivatisés ont ensuite été dilués avec 185  $\mu\text{L}$  d'eau Milli-Q avant injection dans l'électrophorèse capillaire Beckman Coulter PA800-plus Pharmaceutical Analysis System (Beckman Coulter, Brea, CA, USA), avec le même capillaire que celui utilisé dans la section 8. Le capillaire a été rincé et rempli avec un gel Beckman Coulter *N*-linked carbohydrate separation gel buffer avant injection (3 minutes à 30 psi). Les échantillons ont été injectés par pression à 0,5 psi pendant 10 sec. La migration a été réalisée à 30 kV pendant 10 minutes. La détection a été faite à 488 nm comme décrit en section 8.

## **10. Extraction des protéines**

Les protéines du grain ont été extraites à partir de 300 mg d'amidon. 10 aliquots de 30 mg d'amidon ont été mis en suspension dans 500  $\mu\text{L}$  de tampon d'extraction (Tris HCl 0,2 M à pH 6,8 ; SDS 2 % (w/v) ; 50 mM de dithiothreitol (DTT)). Les aliquots sont ensuite chauffés à 100 °C jusqu'à gélatinisation des grains. Après gélatinisation, les échantillons sont laissés pendant 10 minutes à 100 °C et centrifugés pendant 10 min à 13 000 tours par minute. Les surnageants

contenant les protéines sont récupérés, mis en commun et concentrés sur une colonne Amicon® (Millipore, Billerica, MA, USA) avec un filtre d'acétate de cellulose 50 kDa dans un tube de 15 mL. Les tubes ont été centrifugés à 7 500 tours par minutes pendant 30 minutes de manière à concentrer les échantillons. 5 µL de chaque échantillon, dilué 10 fois, a été utilisé pour réaliser un dosage à l'acide bicinchoninique (BCA; Thermo Fisher Scientific, Waltham, MA, USA) en suivant les instructions du fabricant.

## **11. Electrophorèse en gel de polyacrylamide contenant du sodium dodécyl sulfate (SDS-PAGE)**

Une solution stock d'acrylamide (50 mL) a été préparée en utilisant 14,6 g d'acrylamide et 390 mg de N, N, N',N' méthylène-bis-acrylamide comme agent de liaison. 330 mg de Dextran (MW 500 000) ont aussi été incorporés à la solution précédente de manière à améliorer la séparation dans le gel en se basant sur le protocole développé dans une précédente étude (Ainseba-Chirani et al., 2011). La concentration en polyacrylamide était de 4 % dans le gel de concentration (avec 14 mM de SDS et 0,5 M de tris-base, pH 6,8) et de 7 ou 10 % dans le gel de séparation (avec 20 mM de SDS, 1,1 M de tris-base et 0,4 M de tris-HCl, pH 8.8). Les échantillons de protéines ont été dissouts dans du tampon de dénaturation contenant 0,375 M de tris, 2 mM de dithiothréitol (DTT), 1,4 M de SDS, 2 % de glycérol (v/v), et 0.02 % de bleu de bromophénol (m/v). Des marqueurs de tailles de faible et haut poids moléculaires ont été utilisés pour la calibration du gel. L'électrophorèse a été menée en utilisant un système SE 600 Ruby (GE Lifescience, Velizy-Villacoublay, France), à 40 mA dans un tampon de migration composé de 0,025 M de tris, 0,192 M de glycine et 0,1 % (m/v) de SDS. Les gels ont été fixés dans une solution à 20 % (v/v) d'éthanol et 7 % (v/v) d'acide acétique pendant 20 min. Les gels ont ensuite été incubés toute la nuit sous agitation dans une solution de coloration contenant du Ru(BPS)<sub>3</sub>. Les gels ont enfin été lavés dans une solution à 25 % (v/v) d'éthanol, 14% (v/v) d'acide acétique pour être digitalisés à l'aide d'un scanner Typhoon® 9000 (GE Lifescience, Velizy-Villacoublay, France), L'acquisition des images a été réalisée à une résolution de 100 nm, photomultiplicateurs 580 V avec le laser bleu les filtres correspondant à 488 nm pour la longueur d'onde d'excitation et 610 nm pour la longueur d'onde d'émission.

## **12. Digestion à la trypsine en gel**

Les bandes de protéines ont été coupées dans le gel et décolorées dans de l'acétonitrile à 50 % (v/v), 25 mM de bicarbonate d'ammonium (ABC) avant de les déshydrater dans de

l'acétonitrile à 100 % (v/v). Les protéines ont été réduites en incubant les bandes de gels dans une solution à 10 mM de DTT, 25 mM d'ABC à 56 °C pendant 1h. Elles ont ensuite été alkylées dans une solution avec 55 mM d'iodoacétamide, 25 mM d'ABC à l'obscurité et à température ambiante pendant 1h. Les bandes de gels ont de nouveau été deshydratées avec une solution à 50% d'acétonitrile et 25 mM d'ABC, deux fois pendant 30 min, puis une fois pendant 5 min dans 100 % (v/v) d'acétonitrile. Les bandes de gels ont ensuite été séchées pendant 3 minutes à température ambiante avant réhydratation avec une solution de trypsine pour séquençage (10 ng.µL<sup>-1</sup> dans 25 mM d'ABC) pendant 30 min sur de la glace. Les bandes dans la trypsine, une fois réhydratées, ont été incubées toute la nuit à 37 °C. Les peptides ainsi obtenus ont été extraits des bandes avec de l'acétonitrile 100 % (v/v) contenant 0,1 % (v/v) d'acide trifluoroacétique (TFA) pendant 30 min. Les extraits ont finalement été séchés dans un concentrateur sous vide SpeedVac® et dissous dans une solution d'acide formique à 0,1 % (v/v) pour analyse en spectrométrie de masse.

### **13. Digestion des protéines par « enhanced Filter Aided Sample Preparation » (eFASP)**

50 µg de protéines ont été réduites, alkylées et digérées avec le protocole de digestion eFASP (Erde et al., 2014). Des unités d'ultra-filtration Amicon® 10 kDa (Millipore, Billerica, MA, USA) ont été incubées toute la nuit avec 5 % (v/v) de TWEEN-20 (T20, Sigma-Aldrich, Saint-Louis, MO, USA). Après incubation, les unités de filtration ont été rincées par immersion dans de l'eau Milli-Q 3 fois. La digestion par la méthode eFASP a été réalisée à partir de 50 µg d'échantillon (masse estimée à partir du dosage réalisé avec le kit BCA). Ces échantillons ont été mélangés à 50 µL de tampon de solubilisation et de réduction (SDS 4 % (m/v), acide déoxycholique (DCA) 0,2 % (m/v), DTT 50 mM, ammonium bicarbonate (ABC) 200 mM pH 8,8) toute la nuit à 4 °C. Le lendemain, les échantillons réduits ont été mélangés à 200 µL de tampon d'échange (8 M d'urée, DCA à 0,2 %, 100 mM d'ABC pH 8,8), puis transférés dans les unités Amicon®. Après centrifugation à 13 000 tours par minute pendant 30 min, le filtrat a été jeté, et 200 µL de tampon d'échange est redéposé dans chaque unité de filtration. Et la centrifugation a repris pour 30 minutes de plus. Cette étape fut répétée 2 fois. Les protéines réduites dans l'unité de filtration sont ensuite alkylées par ajout de 100 µL de tampon d'alkylation (8 M d'urée, 50 mM d'iodoacétamide, et 100 mM d'ABC pH 8,8), et incubées à 37°C pendant 1h à l'obscurité. Les unités de filtration sont ensuite centrifugées 30 min à 13 000 tours par minute. Le filtrat fut jeté, et 200 µL de tampon d'échange ont été ajoutés dans chaque

unité. Cette étape a été répétée 1 fois. Après que l'échange de tampon ait été effectué, 3 passages de 200  $\mu\text{L}$  de tampon de digestion (50 mM d'ABC pH 8.8, 0.2 % (m/v) de DCA) dans les unités de filtrations, 50 ng de trypsine/Lys C dans 100  $\mu\text{L}$  d'ABC 50 mM pH 8.8 furent ajoutés dans les unités Amicon<sup>®</sup>. La digestion a été réalisée toute la nuit à 37 °C. Les peptides ainsi obtenus sont récupérés en transférant les unités de filtrations dans un nouveau tube collecteur, qui a ensuite été centrifugé à 13 000 tours par minute pendant 20 min. Pour compléter la récupération des peptides, deux autres passages avec 50  $\mu\text{L}$  d'ABC 50 mM pH 8.8 ont été réalisés. 200  $\mu\text{L}$  d'acétate d'éthyle ont été ajoutés au filtrat et l'ensemble a été transféré en tube 2 mL. 2,5  $\mu\text{L}$  de TFA 6 M ont été ajoutés pour précipiter les peptides. 800  $\mu\text{L}$  d'acétate d'éthyle ont été ajoutés, puis l'ensemble a été homogénéisé au vortex. Les tubes ont ensuite été centrifugés à 13 000 tours par minute pendant 10 minutes. La phase organique (au-dessus) a été retirée et cette étape fut répétée 2 fois. Les tubes sont ensuite placés ouverts 5 minutes sur bloc chauffant à 60 °C de manière à évaporer l'acétate d'éthyle résiduel. Les sels volatils sont ensuite retirés par séchage au concentrateur SpeedVac<sup>™</sup> (Eppendorf<sup>™</sup> Concentrator Plus, Eppendorf, Hamburg, Allemagne). Cette étape a été répétée après ajout de 100  $\mu\text{L}$  de méthanol 50 % (v/v). 10  $\mu\text{g}$  de peptides secs furent ensuite resuspendus dans de l'acide formique 0,1 % (v/v).

#### **14. Extraction et digestion des protéines de l'amidon par C18 in Stage Tip (iST)**

Les grains d'amidons ayant été capturés en capillaires ont été élués dans des tubes 1,5 mL contenant 50  $\mu\text{L}$  d'un tampon de dénaturation, réduction et alkylation (Urée 8 M, 10 mM de tris (2-carboxyethyl) phosphine (TCEP) ; 40 mM de 2-chloroacétamide, 100 mM de tris base pH 8,5). Les échantillons ont été chauffés à 95 °C pendant 10 minutes, en remettant en suspension régulièrement à l'aide d'un vortex. Par la suite, les échantillons sont laissés 10 minutes de plus à 95 °C sans agitation. 150  $\mu\text{L}$  de tampon de digestion (Acétonitrile à 10 % (v/v) dans 25 mM de tris pH 8,5) ont été ajoutés à l'échantillon, avec 1  $\mu\text{g}$  de mix trypsine Lys-C. Les échantillons ont été incubés pendant 3h à 37 °C. La digestion est arrêtée par l'ajout de 100  $\mu\text{L}$  de TFA à 2 % (v/v). Les peptides ont ensuite été purifiés à l'aide d'un cône de pipette contenant un disque C18. Le cône a préalablement été activé avec 200  $\mu\text{L}$  de méthanol, puis équilibré avec 200  $\mu\text{L}$  d'acétonitrile à 80 % (v/v) dans de l'acide acétique à 0,1 % (v/v), suivi de 200  $\mu\text{L}$  d'acide acétique à 0,1 % (v/v). Après cela, l'échantillon a été chargé dans le cône. 2 lavages ont été effectués avec 200  $\mu\text{L}$  d'acide acétique à 0.1 % (v/v). Enfin, les peptides purifiés ont été élués à l'aide de 2  $\times$  100  $\mu\text{L}$  d'acétonitrile à 80 % (v/v) dans de l'acide acétique à 0,1 % (v/v) à l'intérieur d'un tube de 1,5 mL. Les tubes ont été séchés au SpeedVac<sup>®</sup>, puis remis en

suspension dans 2 $\mu$ L d'acide formique à 0,1 % (v/v) pour pouvoir être analysés en spectrométrie de masse.

### **15. Analyses en spectrométrie de masse par acquisition dépendante de données (DDA)**

Un instrument HPLC nanoflux (U3000 RSLC, Thermo Fisher Scientific, Waltham, MA, USA) a été couplé en ligne à un spectromètre de masse Q Exactive plus Orbitrap (Thermo Fisher Scientific) muni d'un électronébuliseur comme source d'ionisation. 1  $\mu$ L de la mixture de peptide est chargé dans la colonne de préconcentration (Thermo Fisher Scientific, Acclaim PepMap100 C18, granulométrie 5  $\mu$ m, 300  $\mu$ m de diamètre interne  $\times$  5 mm de longueur) en utilisant une boucle d'injection partielle pendant 5 minutes à un débit de 10  $\mu$ L.min<sup>-1</sup> avec le tampon A (5 % d'acétonitrile et 0.1% d'acide formique). L'échantillon injecté a ensuite été séparé sur une colonne C18 (Acclaim PepMap100 C18, granulométrie 3  $\mu$ m, 75 mm de diamètre interne  $\times$  500 mm de longueur) avec un gradient linéaire de 5 à 50 % de tampon B (75 % acétonitrile et 0.1% d'acide formique), avec un débit de 250 nL.min<sup>-1</sup> et une température de 45 °C. Le temps total d'une série LC MS/MS fut de 270 min. Les données de MS ont été acquises sur le spectromètre de masse Q-Exactive plus (Thermo Fisher Scientific) en utilisant la méthode « data-dependant top 20 », choisissant de manière dynamique les ions précurseurs les plus abondants lors du « survey scan » (350-1600 m/z) pour la fragmentation par dissociation collisionnelle à plus forte énergie (HCD). La durée de l'exclusion dynamique fut de 60 secondes. L'isolation des ions précurseurs a été faite sur une fenêtre de 1,6 m/z et les analyses MS/MS ont été acquises avec une masse de départ de 80 m/z. Les analyses MS, ou « survey scans », ont été acquises à une résolution de 70 000 pour m/z=400. L'« Automatic Gain Control target » (AGC), correspondant au nombre d'ions accumulés dans le piège à ions, est réglé sur 10<sup>6</sup> ions avec un temps de remplissage maximum de 100 millisecondes. La résolution des spectres MS/MS après HCD a été réglée sur 70 000 pour un m/z = 200. L'AGC a été réglé sur 10<sup>5</sup> ions avec un temps de remplissage maximum de 200 millisecondes. L'énergie de collision normalisée était de 28 eV. Le taux de remplissage, qui spécifie le pourcentage minimum de la valeur cible pouvant potentiellement être atteinte au temps de remplissage maximum, a été défini à 0.4 %. Les instruments ont été utilisés en mode « peptide recognition », c'est à dire pour des charges comprises entre 2 et 8 inclus, l'exclusion des ions précurseurs monochargés et non assignés ayant été autorisée.

## 16. Identification des protéines associées au grain d'amidon avec PEAKS

Pour les premières expériences réalisées sur *Solanum tuberosum* cv. Monalisa de manière à identifier l'ensemble des protéines du grain d'amidon, les données brutes acquises ont été analysées avec PEAKS Studio 7.0 (Bioinformatics Solutions Inc.) en utilisant une banque de données incluant les séquences protéiques du PGSC\_DM\_v3.4\_pep\_non-redundant (Avril 2017, 52.570 entrées), et avec l'ajout des protéines du métabolisme connues mais non annotées ou mal annotées dans le PGSC en se basant sur l'article de Van Harsselaar et collaborateurs de 2017 (Zhang et al., 2012; Van Harsselaar et al., 2017). La tolérance pour la masse des peptides a été réglée à 10 ppm et 0,01 Da pour la MS/MS. Les modifications variables incluses pour l'analyse des masses furent les suivantes : Oxydation de M, Y, H, dé-amidation de N, Q, carbamidométhylation de C, phosphorylation de Y, S, T, pyro-G. Pour une plus grande confiance dans les résultats, le seuil du « False discovery rate » (FDR) a été réglé à 1%. Les protéines avec un score “-10lgP” minimum de 20 et avec deux peptides uniques ont été retenues.

## 17. Quantification avec MaxQuant et Perseus

Les protéines issues des expériences réalisées à partir de *Solanum tuberosum* cv. Monalisa, ainsi que celles des mutants *waxy* du cv. Désirée, ont été quantifiées avec le programme Maxquant 1.5.3.30. L'interrogation a, dans un premier temps, été faite avec la banque de donnée PGSC\_DM\_v3.4\_pep\_non-redundant modifiée avec les protéines non-annotées et mal-annotées. Dans un deuxième temps, après que le protéome complet du grain d'amidon ait été établi, l'interrogation a été faite avec une banque de données customisée contenant seulement l'ensemble des protéines identifiées précédemment, de manière à réduire le temps d'analyse et à éviter les doublons existants dans la banque de donnée PGSC. Les scores iBAQ et LFQ obtenus pour chaque analyse ont été utilisés sur le programme Perseus 1.6.0.2 pour obtenir l'abondance de chaque protéine. Le paramètre retenu pour la quantification réalisée lors de l'établissement du protéome complet a été “molecule/total molecule.10^6”. Lors de l'étude réalisée sur les différentes fractions, le paramètre retenu a été “Concentration [nM] iBAQ”. Dans ce dernier cas, la concentration a ensuite été divisée par la concentration de l'ensemble des protéines dans le grain d'amidon (en  $\mu\text{g.mg d'amidon}^{-1}$ ) de manière à obtenir un résultat en  $\text{fmol.mg d'amidon}^{-1}$ . La valeur d'accroissement est ensuite obtenue en normalisant la concentration d'une protéine donnée dans une fraction par rapport à celle de la même protéine dans la fraction A dont la valeur est fixée à 1. Les valeurs LFQ sont transformé en  $\log_2$ , puis filtrées et utilisées sur Perseus 1.6.0.2 pour réaliser les tests statistiques. Pour les comparaisons

multiples, une analyse ANOVA avec une correction basée sur le score FDR (FDR=0,05) a été faite. Pour la comparaison entre deux groupes, des *t-tests* ont été réalisés, où une *p-value* inférieure à 0,05 permettait de considérer la différence comme statistiquement valide. Les corrélations entre les différentes variables ont été définies par des coefficients de Pearson à l'aide d'un script écrit sur le programme R 3.6.0.

### 18. Calcul du taux d'incorporation de <sup>15</sup>N.

Lors des expériences d'incorporation/chasse au <sup>15</sup>N réalisées sur *Solanum tuberosum* cv. Désirée, les peptides issus des contrôles non marqués ont été identifiés à partir du programme Thermo® Proteome Discoverer 2.4. Le déroulement des opérations choisi pour l'interrogation a été "Basic Consensus" pour le consensus et "Q-Exactive Basic Sequest" pour le traitement. La banque de données utilisée a été la banque de données maison composée des protéines ayant été identifiées lors de l'établissement du protéome complet. Les fichiers "Peptide spectrum matches" (PSMs) ont été utilisés pour sélectionner 3 peptides pour chaque protéine identifiée de manière à calculer le taux d'incorporation du <sup>15</sup>N. Les massifs de pics des m/z des ions précurseurs ont été visualisés sur le programme Thermo® Xcalibur. Les chromatogrammes, les spectres de masses et les tableaux de valeurs des intensités de pics m/z correspondant au peptide d'intérêt (en se basant sur l'incrément de masse tenant compte de la charge de l'ion précurseur) ont été exportés de Thermo® Xcalibur vers Microsoft® Office Excel pour la mise en forme des figures. L'abondance relative de chaque pic dans le contrôle non-marqué, ajusté à partir de la valeur d'intensité du pic monoisotopique, a été utilisée pour discriminer l'abondance naturelle des isotopes lourds de la quantité de <sup>15</sup>N incorporé dans la plante lors des expériences de marquage. Par la suite, l'abondance relative des pics issue des massifs de pics de l'ion précurseur ont été calculées. Enfin, le taux d'incorporation a été calculé à partir de cette formule

$$:taux\ d'incorporation = \frac{\sum_{i=1}^n \binom{n}{i} i \times (abondance\ relative\ du\ pic)_{+i}}{\text{nombre d'azote dans la séquence du peptide}}$$

Où *n* correspond au nombre de pics présents dans le massif de pics du peptide étudié.



# Résultats

## I. Etablissement du protéome complet du grain d'amidon de pomme de terre

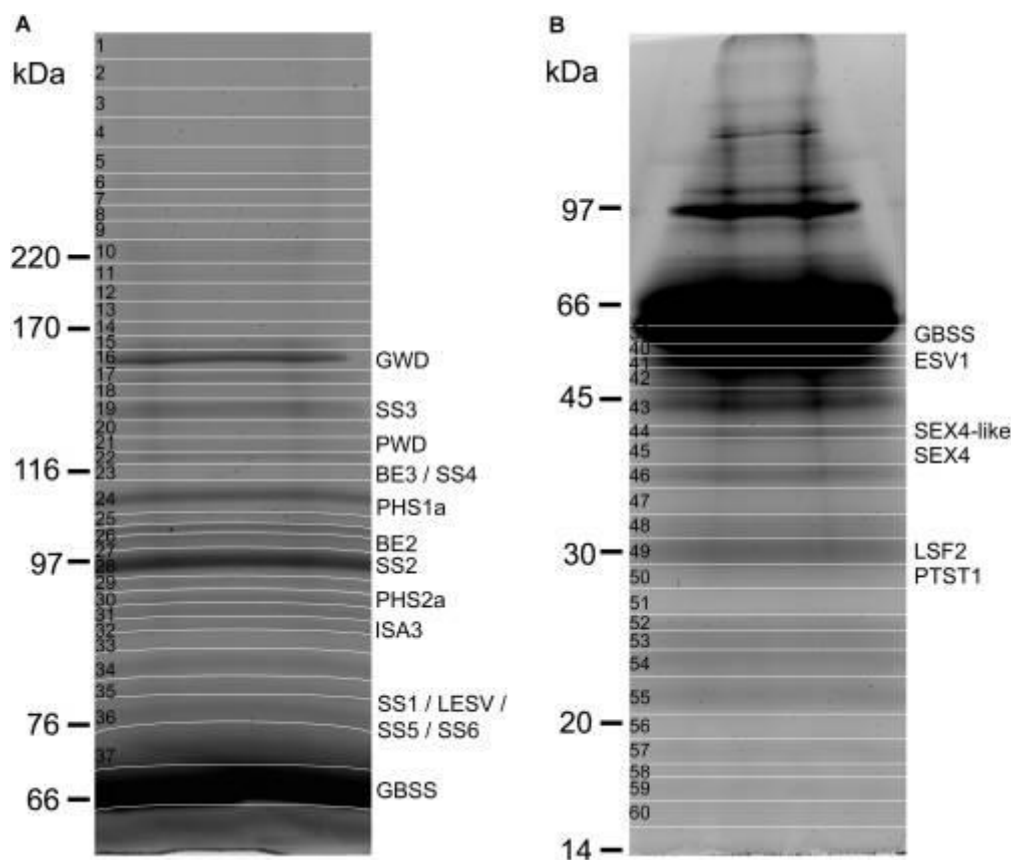
L'établissement du protéome complet de la pomme de terre était un projet initié par Nicolas Szydlowski et Jérémy Verbeke en 2014. La culture des pommes de terre, l'extraction de l'amidon de pomme de terre, le traitement des grains d'amidon, les gels SDS-PAGE et l'identification des bandes ont été réalisés par leurs soins. L'analyse en spectrométrie de masse et l'interprétation des résultats à partir des logiciels d'identification et de quantification ont été réalisés par les membres de la plateforme de protéomique du MSAP, Stéphanie Flament et Fabrice Bray. Par la suite, j'ai réalisé les expériences de quantification des protéines par protéomique shotgun, c'est-à-dire par l'analyse de l'ensemble des protéines extraites de l'amidon sans séparation préalable de celles-ci en gel. Cela a permis l'établissement de la stœchiométrie des protéines du grain d'amidon de la pomme de terre. Ces travaux ont donné lieu à une publication en 2018 dans le journal *Frontiers in Plant Science*.

[Publication 1: Proteome Analysis of Potato Starch Reveals the Presence of New Starch Metabolic Proteins as Well as Multiple Protease Inhibitors.](#) (voir Annexe 1)

### 1. Protéines du métabolisme de l'amidon.

Les protéines du grain d'amidon ont été analysées par SDS-PAGE et imagerie de fluorescence après isolation des protéines à partir de 300 mg d'amidon de pomme de terre cv. Monalisa (Figure 37.). Les bandes de poids moléculaires supérieurs à 60 kDa et inférieurs à 45 kDa ont été respectivement séparées sur des gels de polyacrylamide de 7 et 10 %. Le poids moléculaire apparent de chaque bande a été déterminé à partir de leur distance de migration relative (Rf) en utilisant une gamme étalon linéaire construite à partir des Rfs des protéines de l'échelle de masse (Tableau 2). Soixante bandes de gel ont été coupées en suivant le motif représenté sur la figure 37. Les protéines contenues dans les bandes ont été digérées par la trypsine en gel et les peptides extraits de chaque bande ont été analysés par spectrométrie de masse. L'identification des protéines a été réalisée à l'aide de la banque de données du PGSC (Xu et al., 2011). Certaines enzymes connues du métabolisme de l'amidon ont une séquence tronquée ou une mauvaise annotation dans le génome publié (par exemple, *SS3*, *PHS1a*, *ISA3*) (Xu et al., 2011; Van Harsselaar et al., 2017). Par conséquent, nous avons ajouté manuellement les identifiants iTAG Sotub09g011090.1.1 (BE2), Sotub02g012780.1.1 (AMY3-like), Sotub07g025820.1.1 (BE1.2) et les références UniProt, [P30924](#) (BE1), [Q43846](#) (SS3), [P32811](#) (PHS2a), [Q84YG5](#) (ISA3), et [P04045](#) (PHS1a) à la banque de donnée. Le seuil de

confiance de l'identification des protéines (-10LgP) a été défini empiriquement à 100 et seules les protéines identifiées avec au moins 2 peptides uniques ont été prises en considération.



**Figure 37. Gels d'électrophorèse en polyacrylamide des protéines liées au grain d'amidon de pomme de terre.** 300 mg d'amidon ont été utilisés pour l'isolation des protéines. (A) Après concentration sur colonne d'ultrafiltration Amicon® 50 kDa, les échantillons de protéines ont été chargés sur un gel de polyacrylamide à 7 % et soumis à une migration par électrophorèse à 25 mA pendant 5h. Le gel a été coloré avec une solution de Sypro ruby (0,5 mg/L) avant imagerie par fluorescence. Le gel a été coupé en plusieurs bandes pour des analyses complémentaires en se basant sur l'abaque représenté par des lignes blanches. (B) Même qu'en (A) avec une concentration en polyacrylamide de 10 %.

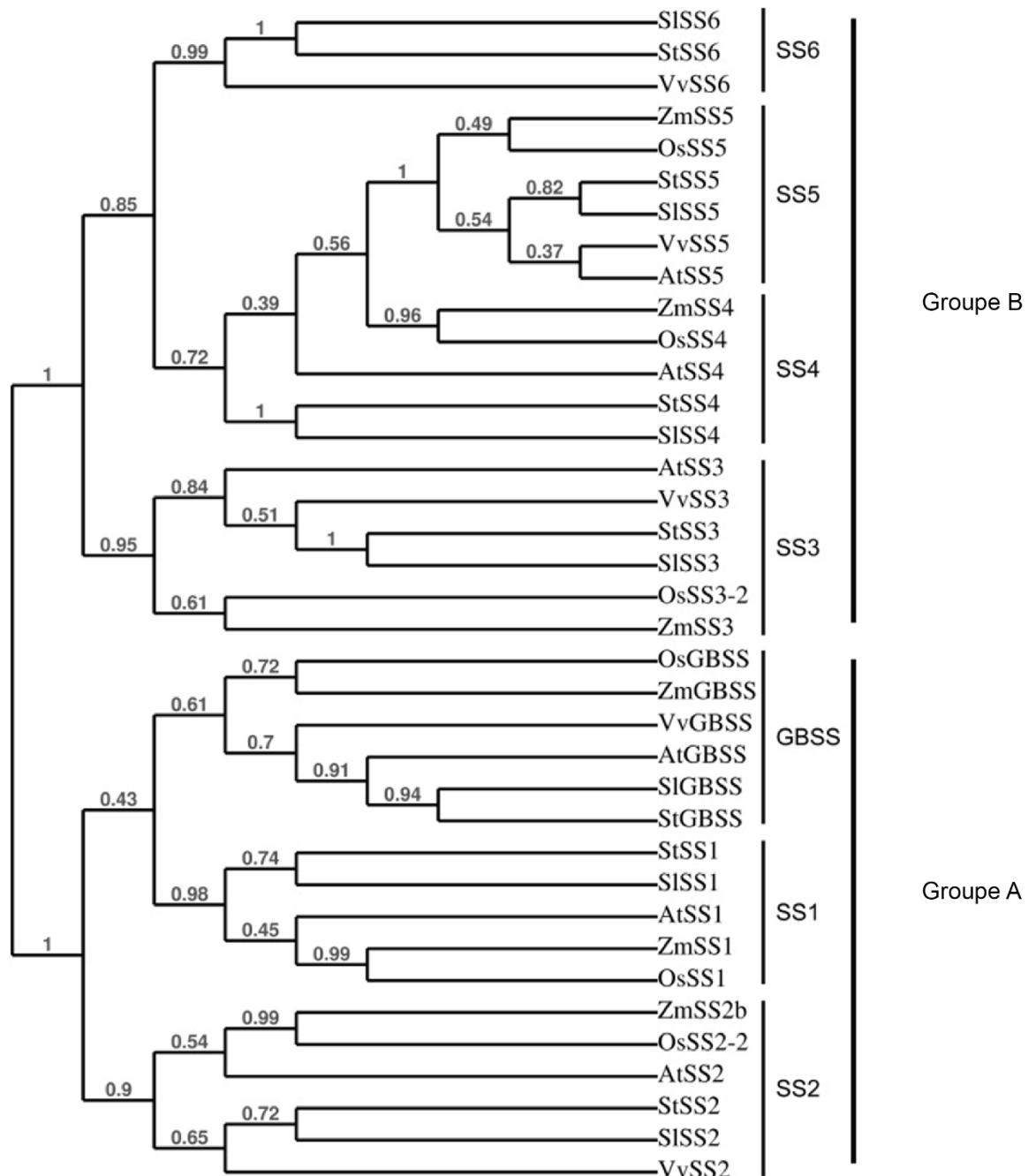
**Tableau 2. Protéines du métabolisme de l'amidon identifiées sur le gel SDS-PAGE.**

Isoforme	Identifiant PGSC/UniProt/iTAG	MW prédit (kDa)	MW observé (kDa)	Numéro de bande	Nombre de peptides uniques	Couverture (%)
GWD	PGSC0003DMP400013565	163,2	160,7	17	78	39
SS3	<a href="#">Q43846</a>	139	142,7	19	57	37
PWD	PGSC0003DMP400029040	132,1	129,3	21	27	25
SS4	PGSC0003DMP400014590	113,8	122,9	23	4	4
BE1	PGSC0003DMP400017627	104,1	122,9	23	9	9
PHS1a	<a href="#">P04045</a>	109,4	118,2	24	23	28
BE2	Sotub09g011090.1.1	100	105,7	27	54	41
SS2	PGSC0003DMP400002384	85,2	103	28	203	87
PHS2a	<a href="#">P32811</a>	95,1	96	30	15	21

ISA3	<a href="#">Q84YG5</a>	87,4	88,9	32	31	45
SS1	PGSC0003DMP400032329	70,6	74,1	36	46	46
LESV	PGSC0003DMP400028510	63,4	74,1	36	72	77
SS6	PGSC0003DMP400023948	69,5	74,1	36	9	29
SS5	PGSC0003DMP400053324	76,5	74,1	36	10	27
GBSS1	PGSC0003DMP400021398	66,5	62,8	39	870	94
ESV1	PGSC0003DMP400017605	48,9	51,8	41	6	17
SEX4-like	PGSC0003DMP400047519	41,5	42,8	44	12	38
SEX4	PGSC0003DMP400026742	41,5	38,3	45	11	33
LSF2	PGSC0003DMP400050640	24,9	30,7	49	13	50
PTST1	PGSC0003DMP400053308	28,5	28,6	50	21	58

Un total de 36 protéines a été identifiées dans cette étude (Tableau 2 et 3). Vingt d'entre elles participent au métabolisme de l'amidon (Tableau 2). GBSS est la protéine majeure du grain d'amidon de pomme de terre (Stensballe et al., 2008). La bande présentant une très forte intensité de fluorescence sur le gel et d'un poids moléculaire estimé de 62,8 kDa, correspond à GBSS comme ceci fut confirmé par spectrométrie de masse (Figure 37. Tableau 2). Dans une étude précédente du protéome de l'amidon de pomme de terre, seules SS2 et GWD avaient été identifiées par spectrométrie de masse en association avec le grain (Stensballe et al., 2008). SS1, SS2, SS3, BE1, BE2, GWD, et PHS1 avaient aussi été décrites chez d'autres espèces tel que le maïs, le blé, le riz, le pois et l'orge (Grimaud et al., 2008; Stensballe et al., 2008). Nous avons retrouvé l'ensemble de ces isoformes dans nos analyses (Tableau 2). Nous avons par ailleurs observé 2 isoformes de SSs supplémentaires : SS5 et SS6, avec respectivement 10 et 9 peptides uniques, ainsi qu'une protéine partageant une forte homologie de séquence avec SEX4, annotée SEX4-like (Tableau 1)(Van Harsselaar et al., 2017). Les deux gènes SS5 et SS6 avaient été annotés chez la pomme de terre mais leur fonction n'avait à ce moment pas encore été étudiée (Deschamps et al., 2008; Liu et al., 2015; Van Harsselaar et al., 2017). Une recherche par BLAST en utilisant la séquence de SS6 de pomme de terre (XP\_006353746.1) comme référence, révèle la présence d'orthologues annotés chez la tomate *Solanum lycopersicum* (NP\_001234387.1) et la vigne *Vitis vinifera* (NP\_001268021.1). De plus, cette analyse permet d'identifier 90 séquences orthologues avec un score allant de 572 à 1301 et une E-value de 0 dans les génomes de nombreuses espèces de plantes, indiquant que ce gène est conservé au sein du règne végétal. L'analyse phylogénétique de l'ensemble des séquences de SSs confirme que SS6 est une classe d'amidon synthase à part entière (Figure 38). Tout comme SS3 et SS4, SS6 appartient au groupe B des amidons synthases. On retrouve une forte similarité de séquence entre SS4 et SS6 mais cette dernière ne possède pas la longue extension en N-ter habituellement observée dans les séquences de SS4 (Leterrier et al., 2008). De plus, la séquence de SS6 contient

les deux motifs K-X-G-G-L et X-X-G-G-L, fortement conservés chez les SSs (Figure 39) (Cao et al., 1999).



**Figure 38. Arbre phylogénétique montrant les relations entre les SSs de plantes.** L'arbre a été construit avec PhyML (basé sur le maximum de vraisemblance) et les limites de confiances ont été assignées en amorçant les alignements avec 100 essais. At, *Arabidopsis thaliana*; St, *Solanum tuberosum*; Zm, *Zea maize*; Os, *Oryza sativa*; Vv, *Vitis vinifera*; S1, *Solanum lycopersicum*. S1SS6: [NP\\_001234387.1](http://www.ncbi.nlm.nih.gov/nuccore/NP_001234387.1), StSS6: [XP\\_006353746.1](http://www.ncbi.nlm.nih.gov/nuccore/XP_006353746.1), VvSS6: [NP\\_001268021.1](http://www.ncbi.nlm.nih.gov/nuccore/NP_001268021.1), ZmSS5: [NP\\_001123603.1](http://www.ncbi.nlm.nih.gov/nuccore/NP_001123603.1), OsSS5: [XP\\_015626202.1](http://www.ncbi.nlm.nih.gov/nuccore/XP_015626202.1), StSS5: PGSC0003DMP400053324, S1SS5: [XP\\_019067616.1](http://www.ncbi.nlm.nih.gov/nuccore/XP_019067616.1), VvSS5: [NP\\_001290013.1](http://www.ncbi.nlm.nih.gov/nuccore/NP_001290013.1), AtSS5: [NP\\_569018.1](http://www.ncbi.nlm.nih.gov/nuccore/NP_569018.1), ZmSS4: [NP\\_001123590.1](http://www.ncbi.nlm.nih.gov/nuccore/NP_001123590.1), OsSS4: [XP\\_015639005.1](http://www.ncbi.nlm.nih.gov/nuccore/XP_015639005.1), AtSS4: [NP\\_193558.3](http://www.ncbi.nlm.nih.gov/nuccore/NP_193558.3), StSS4:

PGSC0003DMP400014590, SISS4: [NP\\_001234617.2](#), AtSS3: [NP\\_001184965.1](#), VvSS3: [XP\\_002269011.2](#), StSS3: [Q43846](#), SISS3: [NP\\_001234623.1](#), OsSS3-2: [XP\\_015650668.1](#), ZmSS3: [NP\\_001104881.2](#), OsGBSS: [XP\\_015644490.1](#), ZmGBSS: [NP\\_001105001.3](#), VvGBSS: [XP\\_010660257.1](#), AtGBSS: [NP\\_174566.1](#), SIGBSS: [NP\\_001311458.1](#), StGBSS: PGSC0003DMP400021398, StSS1: PGSC0003DMP400032328, SISS1: [XP\\_010318024.1](#), AtSS1: [NP\\_197818.1](#), ZmSS1: [NP\\_001104892.1](#), OsSS1: [XP\\_015644241.1](#), ZmSS2b: [NP\\_001106014.1](#), OsSS2-2: [XP\\_015627452.1](#), AtSS2: [NP\\_186767.1](#), StSS2: PGSC0003DMP400002383, SISS2: [XP\\_004232219.1](#), VvSS2: [XP\\_010661072.1](#).

```

STGBSS 95  KTGGIGDVLGGLPPLAARGHRVMTISPRYDQ-----YKDAWDTSV---AVEVKVGDS 144
STSS1 145  KTGGIGDVCGLPMLAARGHRVMVSPRYLNGGSPSEKAYANAVDLVDV---RATVHCFGD 201
STSS2 290  KTGGIGDVAGALPKALARRGHRVMVAPRYDN-----YPEPQDSGV---RKIYKVDGQ 339
STSS3 794  KVGGLGDVVTLSRAVQDLNHNVDIILPKYDCLK-----MNNVKDFRFHK---NYFWGG 844
STSS5 211  SVGSLAPYVTGLSRALQRKGNLVEVILPKYASLN-----LNEVHGLREVEAEFHSYFNGQ 265
STSS4 518  KVGGLGDVVAGLGLKALQEKHGLVEIVLPKYDCMQ-----YESIKDMKVLDDVVESYFDGR 572
STSS6 139  KVGGLGDVVTGLGRACLNRGKHKVDVMI PFYECIP-----KHCINELALMK-TYNSYHDGN 192

STGBSS 145  IEIVRFFHCYKRGVDRVFDVHMPFLEKVGKGTGSKIYGPKAGLDYLDNELRFSLLCQAAL 204
STSS1 202  AQEVAFYHEYRAGVDWVFDHSSYCRP--G---TPYGDIYGAFG-DNQFRFTLLSHAAC 254
STSS2 340  DVEVTFYQAFIDGVDFVFDISHMFRHI--G---NN---IYGGNRVDILKRMVLFCKAAI 390
STSS3 845  -TEIKVWFGKVEGLSVYFLEPQN--GL-----FSKGCVYG-CS-NDGERFGFFCHAAL 892
STSS5 266  LHGNRIWTGVVCGIGVTFIEPLYSAF-----FGCENIYG-YS-NDFERFTYFSRSL 316
STSS4 573  LYKNNIWTGIVEGLPVYFIEPQHPGKF-----FCRQLYG-EH-DDFKRFSFFSRVAL 623
STSS6 193  WVACKAYRGEVSGVPVILIEPSN--HF-----FKGKNIYGGSY-NELDAYLFFSRACL 242

STGBSS 265  AYQGRFSFSDFPLLNLDPDEFGRSDFDFIDGY---EKPVKGRKINWMMKAGILESHRVVTVS 320
STSS1 310  AHQGVPAVTYNNLGLPPQWYGAWEWIFPTWARAHALDTGETVNVVLKGAIAVADRILTVS 369
STSS2 447  AHQGRGPLEDFS YVDLPPHYMDPFKLYDP-----VGGEHFNI FAAGLKTADRVTVS 498
STSS3 941  EFGADLIG-----RAMTADKATTVS 961
STSS5 365  ESQCVEQPEK LALCGLDPYGLHCSDR LQDN-----NKSHLVNVLKAGVVYSNNVIIMS 417
STSS4 672  EYQGTAAASELTSCGLDAYHLNKPDRMQDN-----SANDRINPVKGAIVFSNIVTTVS 724
STSS6 291  EHYGECRQEQLSKFGLDGSAYATEDKAVDD---RTVGHNPERLSLLKGGIVYSNAIVTVS 347

STGBSS 321  PYYAQELVSAVDKGVELDSV--LRKTCITGIVNGMDTQEWNPATDKYTDVK---YDITT 374
STSS1 370  QGYSWEITTPE-GGYGLHELLSSRQSVLNGITNGIDVNDWNPSTDEHIASHY---SIND 424
STSS2 499  HGYSWELKTSQ-GGWGLHQIINENDWKLQGIVNGIDTKEWNPELDVLHQSDGYMNYSLDT 557
STSS3 962  PTYSQEVSG-----NPVIAPHLHKFHGIVNGIDPDIWDPLNDKFIPIPI---YTSEN 1009
STSS5 418  SMQTKGQIIHA-TSHGLEPTLTIHKDKLVVAPPGFDSAWDPSVDMFLPQN---YSA-- 470
STSS4 725  PTYAQEVRSQA-GGKGLHATINSHSKKFAGILNGIDTDWNPASDNFLKVQ---YSAS- 778
STSS6 345  PTYLKETLCSG---WLSGALMRNRDKYSGILNGIDTEMWNPATDIYLPK---FDAS- 398

STGBSS 375  VMDAKPLLKEALQAAVGLPVDK-----KIPLIGFIGRLEEQKGS DILVAAIHKFIG 425
STSS1 425  -LSGKVQCKTDLQKELGLPIRP-----DCPLIGFIGRLDYQKGVDIILSAIPELMQ 474
STSS2 556  LQTGKPKQKAALQKELGLPVRD-----DVPLIGFIGRLDPQKGVDLIAEAVPMMMG 608
STSS3 1010  VVEGKTAAKEALQRKLGKQAD-----LPLVGIITRLTHQKGIHLIKHAIWRTLE 1059
STSS5 471  DLKGSVCKVSLQQHLGLQEA-----SIVLVGCIFSDISDIELENLKTLIWMASR 521
STSS4 779  DIDGKIQNKALRRLGLSSDF-----RQPLVSCITRLVPQKGVHLIRHAIYRTLE 830
STSS6 399  KTEGKRICKQFVQRGLGLPFQGIKHGICVADQIPLVVCITRLVAQKGLHLITHAIKHVEE 458

STGBSS 481  RFEPCLIQHLAMRYGTVPICASTGGIGDVTVKEGY--TG---FHMGA FNVECDVVDPAVD 535
STSS1 530  RFEPCLNQLYAMRYGTIPVHSTGGIGDVTVKDFN--PY---AQEGNGEGTGWTFSPLOTS 584
STSS2 664  RFEPCLNQLYAMRYGTIPVHSTGGIGDVTVQPFD--PF---NE----SGLGWTF SRAEA 714
STSS3 1120  IFEPCLTQLTAMRYGSIPIVHSTGGIGDVTVFD--VDHDKERAQQCGLEPNGFSFDGADA 1177
STSS5 579  FDGTVLQIPLKAMRYGAMPILIDFTD---SKYGH SVDRDLEGT-----EFS-----RYIND 626
STSS4 889  IFEPCLTQMIAMRYGSIPIVHSTGGIGDVTVFD--VDDDSIPL----QFRNGFTFVTADE 942
STSS6 517  MYEPCGLAQMIGMRYGAVPIVHSTGGIGDVTVFD--MDDQSH-T----EIANGFVFE GIDE 569

```

**Figure 39. Alignement de séquences des domaines catalytiques des amidon synthases de *Solanum tuberosum*.** L'alignement des protéines a été réalisé à partir de l'outil Clustalω2 disponible sur le site de l'EBI (<https://www.ebi.ac.uk/Tools/msa/clustalo/>). Les encadrés noirs

indiquent les motifs K-X-G-G-L et X-X-G-G-L. Les peptides uniques identifiés par spectrométrie de masse sont surlignés en gris pour les séquences de SS4, SS5 et SS6. St, *Solanum tuberosum*, StGBSS: PGSC0003DMP400021398, StSS1: PGSC0003DMP400032328, StSS3: [Q43846](#), StSS4: PGSC0003DMP400014590, StSS5: [XP\\_006357749.1](#), StSS6: [XP\\_006353746.1](#).

Deux isoformes de PHS2, PHS2a et PHS2b, ont été identifiées dans le génome de la pomme de terre (Van Harsselaar et al., 2017). Cette enzyme, aussi appelée PHS2, est une forme cytosolique d'amidon phosphorylase participant au catabolisme de l'amidon dans le cytosol (Fettke et al., 2004; Fettke et al., 2005). La protéine a aussi été identifiée dans le stroma du plaste par analyse protéomique, suggérant une localisation sub-cellulaire multiple (Zybailov et al., 2008; Helm et al., 2014). De manière intéressante, nous avons observé PHS2a en association avec le grain d'amidon dans cette étude (Tableau 2). De plus, PTST1, une protéine responsable de l'adressage de GBSS au grain d'amidon, ainsi que ESV1 et LESV, 2 protéines régulatrices impliquées dans la dégradation de l'amidon à travers un mécanisme encore inconnu ont récemment été identifiées chez *Arabidopsis* (Seung et al., 2015; Feike et al., 2016). Le génome de la pomme de terre contient des orthologues de *PTST1*, *ESV1* et *LESV*. Nous avons observé les protéines correspondantes dans cette étude à des distances de migration relative en accord avec leurs masses moléculaires prédites de respectivement 28,6, 51,8 et 74,1 kDa. (Figure 37 et Tableau 2). Parmi les enzymes de dégradation, PWD, SEX4, ISA3 et LSF2 ont été détectées avec un nombre de peptides uniques allant de 11 à 31 (Tableau 1). Ceci corrobore le fait que ces enzymes interagissent avec le grain d'amidon.

## 2. Les autres protéines identifiées au sein du grain d'amidon

Outre des protéines du métabolisme de l'amidon, nous avons identifié 10 séquences d'inhibiteurs de protéases (Tableau 3). Des analyses phylogénétiques nous ont permis d'établir une classification comprenant 4 groupes dans lesquels le niveau d'identité des protéines est compris entre 77.55 % et 99.09 % (Figure 40a). Neuf de ces protéines appartiennent à la famille des « potato Kunitz-type proteinase inhibitor (PKPI). 2 de ces inhibiteurs de protéases s'apparente à des PKPI de classe A, 2 à des PKPI de classe B et 4 à des PKPI de classe C (Figure 36b)(Ishikawa et al., 1994). Les séquences des deux autres inhibiteurs de protéases contiennent des domaines Proteinase Inhibitor II (PINII) et appartiennent à la famille des inhibiteurs de protéases de pomme de terre de type II (Figure 40b).

**Tableau 3. Autres protéines identifiées comme associées à l'amidon de pomme de terre.** Ces protéines ont été identifiées lors l'analyse en shotgun du protéome de l'amidon de pomme de terre, à partir de la digestion eFASP.

Identifiant PGSC	Nom de la protéine	MW prédit (kDa)	Nombre de peptides uniques	Couverture (%)
PGSC0003DMP400016823	Kunitz-type proteinase inhibitor	20,1	5	35
PGSC0003DMP400017933	Serine protease inhibitor 7	24	7	35
PGSC0003DMP400017950	Cysteine protease inhibitor 1	24,8	5	37
PGSC0003DMP400017953	Kunitz-type tuber invertase inhibitor	24,5	16	67
PGSC0003DMP400017939	Cysteine protease inhibitor 1	21,8	10	26
PGSC0003DMP400016824	Aspartic protease inhibitor 5	23,9	2	22
PGSC0003DMP400008029	Proteinase inhibitor type-2 P303.51	16,6	11	63
PGSC0003DMP400017942	Cysteine protease inhibitor 1	25,1	2	15
PGSC0003DMP400017952	Cysteine protease inhibitor 9	13,8	6	66
PGSC0003DMP400016822	Aspartic protease inhibitor 8	24,1	3	52
PGSC0003DMP400021964	Thioredoxin	19,5	4	38
PGSC0003DMP400028845	Stem 28 kDa glycoprotein	29,5	5	26
PGSC0003DMP400045625	Glutathione peroxidase	26,1	7	33
PGSC0003DMP400009292	Peptidyl-prolyl <i>cis-trans</i> isomerase	26,5	8	29
PGSC0003DMP400011774	Ci21A protein	12,4	4	30
PGSC0003DMP400051081	Beta-tubulin	50,5	2	16

Etonnamment, cette analyse a aussi révélé la présence de protéines impliquées dans des fonctions d'oxydoréduction, tel qu'un homologue de la thiorédoxine M d'Arabidopsis (TRX, PGSC0003DMP400021964) ainsi qu'une Glutathion peroxydase (GPX, PGSC0003DMP400045625), similaire à GPX1 d'Arabidopsis (Tableau 3). Ces 2 enzymes interviennent dans la régulation rédox chez les végétaux, avec la thiorédoxine M participant à la régulation des processus métaboliques liés à la photosynthèse et GPX1 agissant en réponse au stress cellulaire chez la plante (Buchanan, 1984; Noctor et al., 2012). Une glycoprotéine de 28 kDa (PGSC0003DMP400028845) et une Peptidyl-prolyl *cis-trans* isomérase

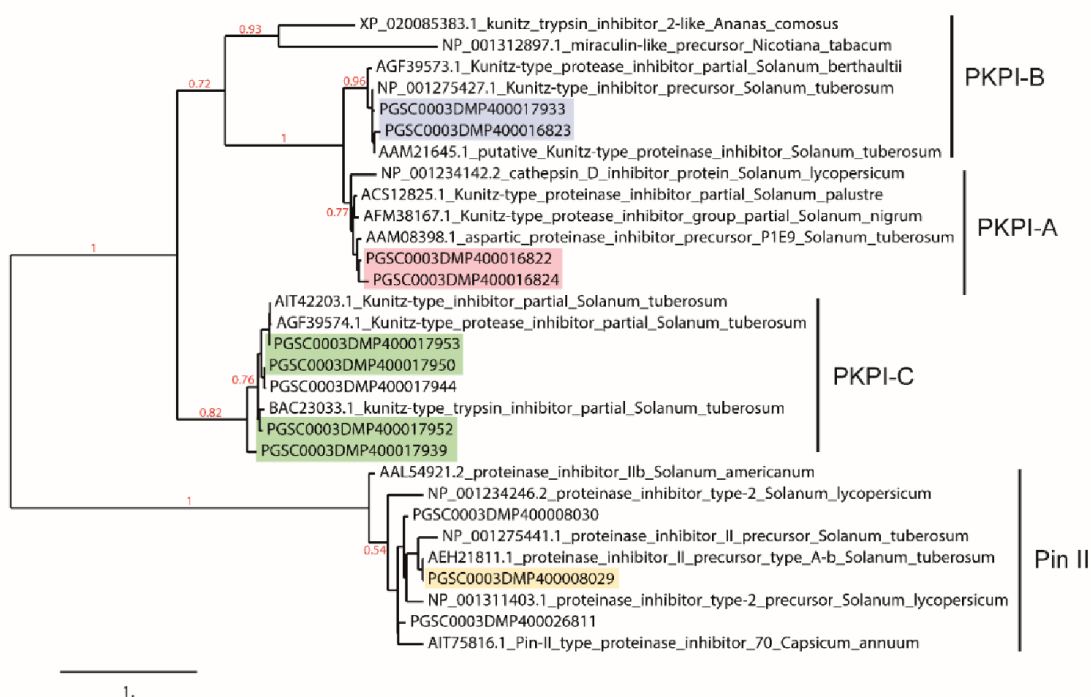


(PGSC0003DMP400009292) appartenant à la famille des cyclophillines ont également été identifiées. Enfin, la protéine Ci21A et la  $\beta$ -tubuline ont été identifiées avec respectivement 4 et 2 peptides uniques (Tableau 3). La présence de ces protéines au sein du grain est plutôt étonnante, du fait qu'elles n'ont pas de rôle connu dans le métabolisme de l'amidon. Il est possible qu'ils s'agissent de protéines adsorbées à la surface des grains lors des étapes d'extraction de l'amidon. Si c'est le cas, ces protéines ne sont présente qu'en surface et peuvent être retiré par traitement préalable des grains avant extraction.

(a)

PGSC0003DMP400008029	100												
PGSC0003DMP400008030	85,71	100											
PGSC0003DMP400026811	81,81	91,21	100										
PGSC0003DMP400017939	12,98	15,3	14,79	100									
PGSC0003DMP400017953	14,93	16,09	15,12	78,06	100								
PGSC0003DMP400017952	5,69	13,00	12,19	82,11	88,61	100							
PGSC0003DMP400017950	15,58	16,09	15,12	78,06	92,69	93,49	100						
PGSC0003DMP400017944	14,93	15,6	14,63	77,55	91,78	92,68	99,09	100					
PGSC0003DMP400016823	11,03	11,29	10,75	25,8	33,87	24,39	34,4	34,4	100				
PGSC0003DMP400017933	11,03	12,68	12,19	28,06	32,87	31,7	33,03	33,48	95,16	100			
PGSC0003DMP400016824	11,68	13,17	13,17	27,04	31,19	29,26	31,65	32,56	73,11	73,85	100		
PGSC0003DMP400016822	11,68	13,17	13,17	26,02	30,59	27,64	31,05	31,96	72,04	74,42	91,28	100	
	PGSC0003DMP400008029	PGSC0003DMP400008030	PGSC0003DMP400026811	PGSC0003DMP400017939	PGSC0003DMP400017953	PGSC0003DMP400017952	PGSC0003DMP400017950	PGSC0003DMP400017944	PGSC0003DMP400016823	PGSC0003DMP400017933	PGSC0003DMP400016824	PGSC0003DMP400016822	

(b)

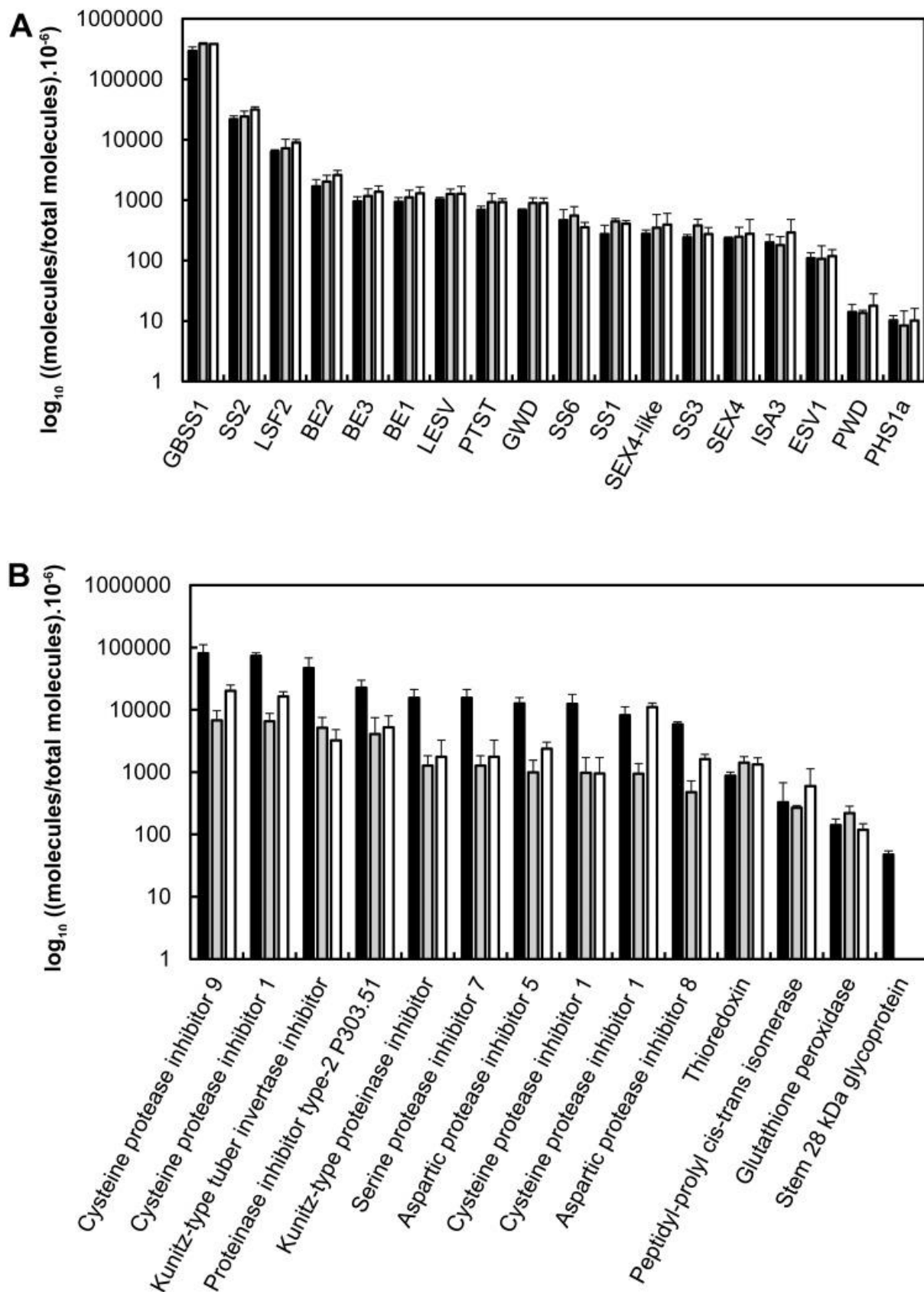


1.

**Figure 40. Identité de séquence et arbre phylogénétique des inhibiteurs de protéases de pomme de terre.** (a) Identité de séquence et des inhibiteurs de protéases. Les scores d'identité (exprimés en %) ont été obtenus à partir de l'alignement des séquences réalisés sur clustalω (<https://www.ebi.ac.uk/Tools/msa/clustalo/>). Les pourcentages d'identité ont été calculés avec SIAS (<http://imed.med.ucm.es/Tools/sias.html>). Pour le calcul d'identité de séquence, la longueur des séquences a été normalisée à partir de la séquence la plus courte. (b) Arbre phylogénétique des séquences des inhibiteurs de protéases. L'arbre a été construit à partir de PhyML, dont l'algorithme se base sur le maximum de vraisemblance. Les intervalles de confiance ont été définis par un Test du rapport de vraisemblance (aLRT). Les identifiants des séquences protéiques utilisées pour la conception de cet arbre sont indiqués sur la figure.

### **3. Quantification des protéines et traitement des grains d'amidon au SDS et à la thermolysine.**

Pour pouvoir étudier la force d'interaction des protéines identifiées avec le grain d'amidon, les échantillons ont été soumis à des traitements de surface au SDS, ou à la thermolysine avant d'extraire les protéines et de les analyser par spectrométrie de masse après digestion eFASP (Figure 41). Les protéines ont été quantifiées sans marquage à partir du score iBAQ, un score de quantification absolue basée sur l'intensité des pics, du programme MaxQuant (Figure 41. Annexe 2). Comme attendu, la quantification sans marquage des protéines associées au grain d'amidon de pomme de terre confirme que GBSS est la protéine majoritaire dans le grain d'amidon, représentant respectivement 87 et 69 % du protéome des grains traités et non traités au SDS (Figure 41.A). Cette disparité entre les deux conditions s'explique par la forte diminution du contenu en inhibiteurs de protéases (75 à 92 %) après traitement de la surface des grains alors que les quantités des protéines du métabolisme de l'amidon, incluant GBSS, ne présentent pas de changement significatif. La thiorédoxine, la cyclophiline et la glutathion peroxydase ont un profil similaire à celui des enzymes du métabolisme de l'amidon, suggérant que leur interaction avec le grain est de même nature.



**Figure 41. Quantification des protéines après traitement de surface au SDS ou à la thermolysine.** Les protéines ont été isolées à partir de 300 mg d'amidon et directement digérées par la trypsine avant analyse par spectrométrie de masse. (A) Protéines du métabolisme de

l'amidon. L'abondance de chaque protéine a été estimée grâce au programme MaxQuant et le score molécules/total molécules a été utilisé pour la quantification. Les données sont représentées en échelle  $\log_{10}$ . Barres noires : protéines issues d'amidon non-traité. Barres grises : protéines issues d'amidon préalablement traité au SDS. Barres blanches : protéines issues d'amidon préalablement traité à la thermolysine. (B) Autres protéines identifiées dans cette étude. Les résultats ont été acquis et sont représentés de la même manière qu'en (A).

## **II. Fluctuations de la stœchiométrie des protéines en fonction de la morphologie du grain.**

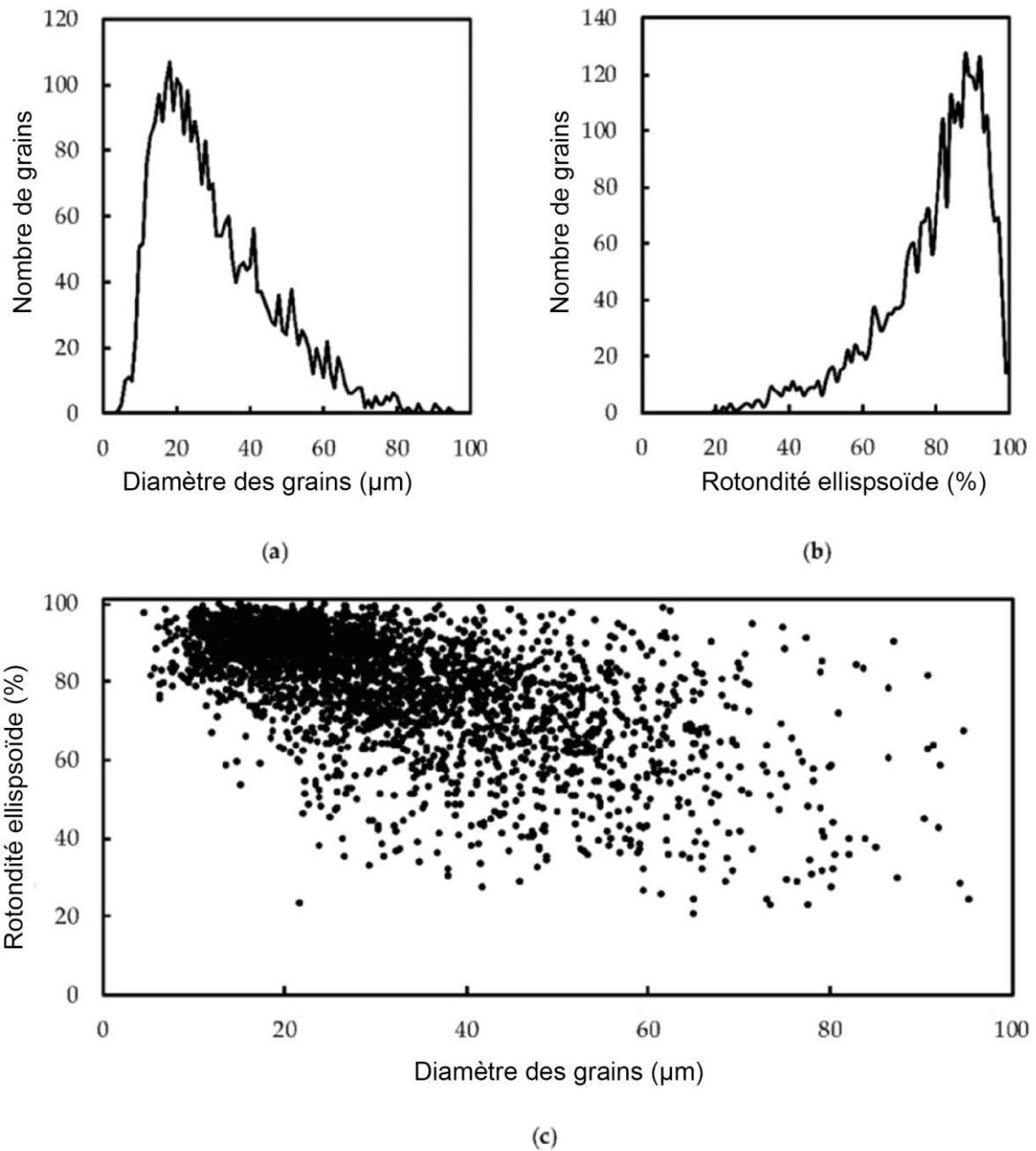
Suite à l'établissement du protéome complet et de la stœchiométrie des protéines dans les grains d'amidon de pomme de terre (cv. Monalisa), nous nous sommes interrogés sur les fluctuations de la structure et de la composition de l'amidon en fonction de la morphologie des grains. Jérémy Verbeke et moi-même avons réalisé un fractionnement à l'aide de filtres en nylon de différentes tailles de mailles. Je me suis occupé de l'analyse en granulomorphométrie de chaque fraction, ainsi que des dosages en protéines et de la distribution de longueurs de chaînes. La structure des polysaccharides et le protéome de l'amidon ont ensuite été caractérisés pour chaque fraction. Jérémy Verbeke et Nicolas Szydlowski se sont occupés de l'analyse du contenu en phosphate. J'ai réalisé la préparation des échantillons pour analyse en spectrométrie de masse, analyses qui ont été faites par Stéphanie Flament et Fabrice Bray de la plateforme de spectrométrie de masse. Ces travaux ont été réalisés en collaboration avec le CERMAV de Grenoble. Jean Luc Puteaux du CERMAV de Grenoble a réalisé l'imagerie des grains d'amidon par microscopie électronique à balayage ainsi que les analyses de diffraction aux rayons X. Les données ont été interprétées par Nicolas Szydlowski et moi-même. Ces travaux ont donné lieu à une publication dans le journal *Plants* en 2019.

Publication 2: Intra-sample heterogeneity of potato starch reveals fluctuation of starch-binding proteins according to granule morphology. (Annexe 3)

### **1. Analyse morphologique des grains d'amidon de pomme de terre (cv. Monalisa) et fractionnement**

L'amidon de *Solanum tuberosum* cv. Monalisa a été analysé à l'aide d'un granulomorphomètre pour déterminer la distribution de diamètres et de rotondité des grains (Figure 42.). L'appareil prend des photographies de manière régulière des grains d'amidons passant par une fenêtre de détection. Les images ainsi obtenues sont ensuite analysées automatiquement par le logiciel présent dans l'appareil pour détourner chaque particule et en déterminer son diamètre moyen et sa rotondité ellipsoïde. La rotondité ellipsoïde

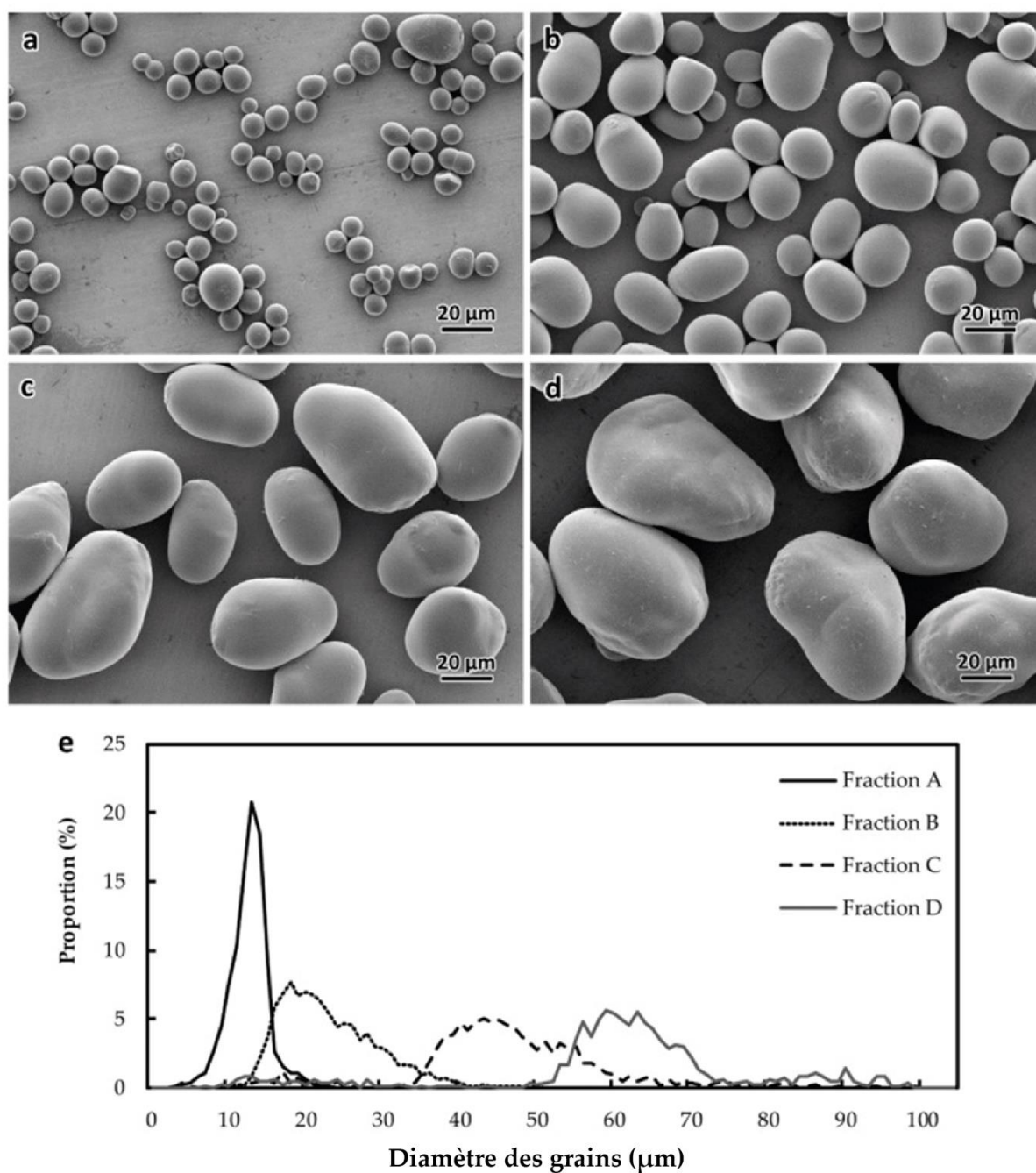
correspond au rapport entre le rayon moyen des cercles circonscrits d'une particule et le rayon du cercle inscrit dans la particule. Le diamètre moyen des grains a été déterminé à partir de l'aire de la particule. Une forte hétérogénéité dans la taille des grains a été observée, avec des diamètres allant de 4 à 100  $\mu\text{m}$ . La distribution des tailles de grains du cv. Monalisa est unimodale avec la majorité des grains présentant un diamètre compris entre 15 et 40  $\mu\text{m}$  (Figure 42a.). La plupart des grains d'amidon ont une forme très ronde avec une valeur de rotondité ellipsoïde supérieure à 60 % (Figure 42b.). De plus, les grains dont la taille est inférieure à 40  $\mu\text{m}$  ont pour la plupart une rotondité particulièrement élevée, dépassant 80 % alors que les grains de plus de 40  $\mu\text{m}$  montrent une plus grande diversité de formes (Figure 42c.).



**Figure 42. Analyse de la taille et de la forme des grains d'amidon de pomme de terre du cv. Monalisa.** Environ 3 000 grains d'amidon ont été utilisés pour l'analyse granulomorphométrique. (a) Distribution de tailles. Le paramètre retenu pour l'estimation de la taille est l'« ISO Area Diameter ». (b) Distribution de forme. Le paramètre retenu est l'« Ellipsoid roundness ». (c) Représentation en nuage de points de la rotondité ellipsoïde en fonction du diamètre des grains.

L'amidon a été fractionné séquentiellement à l'aide de filtres de nylon (11, 31 e 50 μm de largeur de maille) afin d'étudier les disparités de structure et de composition de l'amidon en fonction de leur morphologie. Chacune des fractions (A, B, C et D) a ensuite été analysée par

granulomorphométrie et par microscopie électronique à balayage (Figure 43). Le diamètre moyen était respectivement de  $14,2 \pm 3,7 \mu\text{m}$ ,  $24,5 \pm 6,5 \mu\text{m}$ ,  $47,7 \pm 12,8 \mu\text{m}$  et  $61,8 \pm 17,4 \mu\text{m}$  pour les fractions A, B, C et D. Comme le suggérait l'analyse de l'échantillon non fractionné, les grains des fractions A et B présentent une rotondité ellipsoïde élevée, aux alentours de 80 % (Figure 43a, 43b et Tableau 4). Les grains des fractions C et D montrent une roondeur ellipsoïde plus faible, aux alentours de 70 % (Figure 43c, 43d, et Tableau 4).



**Figure 43. Morphologie des grains d'amidon de pomme de terre du cv. Monalisa fractionnés en fonction de leur taille.** (a-d) imagerie par microscopie électronique à balayage des fractions A (a), B (b), C (c), et D (d). (e) Distribution des diamètres de grains dans chaque

fraction. L'analyse a été réalisée avec 3 035 grains pour la fraction A, 3 436 grains pour la fraction B, 1 198 grains pour la fraction C et 784 grains pour la fraction D.

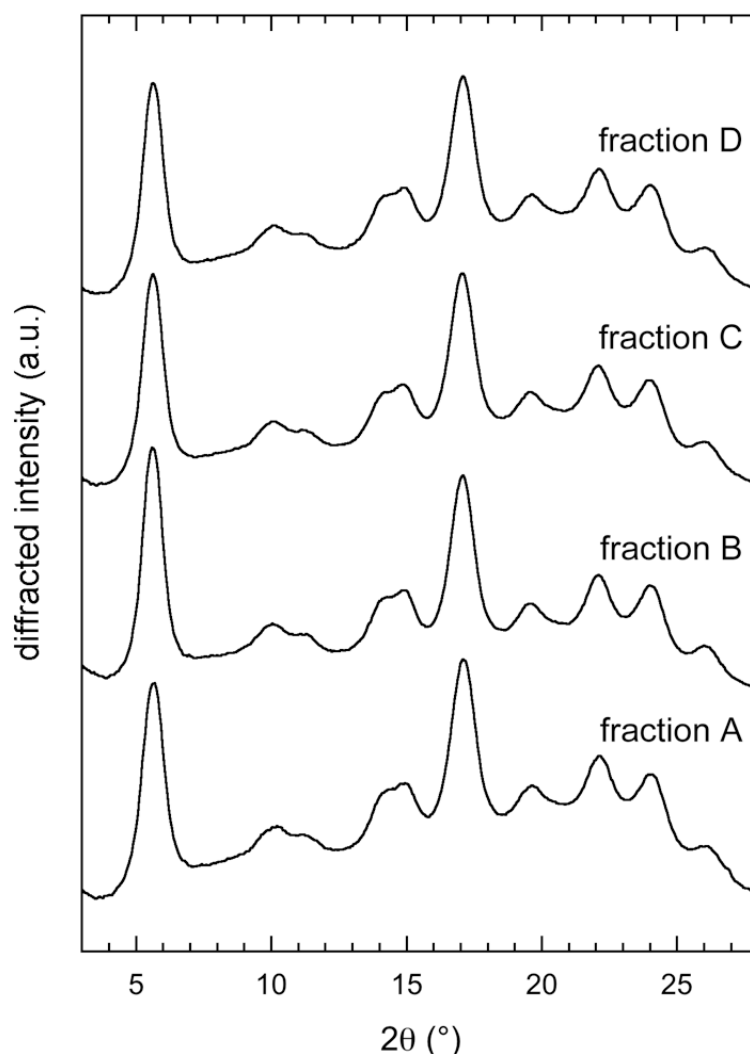
**Tableau 4. Tailles et formes, contenus en protéines, contenus en amylose et longueurs d'ondes au maximum d'absorbance ( $\lambda_{max}$ ) du complexe amylopectine-I<sub>2</sub> dans les différentes fractions.** Ces données ont été obtenues par granulomorphométrie, dosage BCA, kit K-Amyl de Megazyme® et coloration à l'iode, respectivement. L'analyse statistique a été réalisée à l'aide de tests de Student en comparant la fraction à l'échantillon non fractionné. \* correspond à  $P < 0,05$  et \*\* à  $P < 0,001$ .

Fraction	Diamètre moyen ( $\mu\text{m}$ )	Rotondité ellipsoïde (%)	Contenu en protéine ( $\mu\text{g}\cdot\text{mg}^{-1}$ )	Contenu en Amylose (%)	$\lambda_{max}$ du complexe Amylopectine-I <sub>2</sub> (nm)
A	14,2 ± 3,7 **	85,9 ± 12,5 **	2,6 ± 0,2 *	22,6 ± 5,1	548 ± 2
B	24,5 ± 6,5 **	82,5 ± 13,6 **	2,8 ± 0,1 **	20,7 ± 2,1	546 ± 4
C	47,7 ± 12,8 **	72,0 ± 14,8 **	2,1 ± 0,4 *	20,8 ± 2,9	552 ± 2
D	61,8 ± 17,4 **	73,2 ± 15,0 **	2,4 ± 0,1 *	20,6 ± 1,9	549 ± 2
Non fractionné	30,7 ± 15,9	79,9 ± 14,6	2,5 ± 0,2	21,5 ± 1,0	554 ± 4

## 2. Structure et composition de l'amidon dans les différentes fractions

Dans une précédente étude, des différences de profils de diffraction aux rayons X, de contenus en amylose et de distributions de longueurs de chaînes ont été observées entre 3 fractions de grains d'amidon de pomme de terre de diamètre moyen variable et obtenues par tamisage (Wang et al., 2016). L'analyse de diffraction aux rayons-X dans notre cas n'a pas montré de différence de cristallinité en fonction de la taille des grains (Figure 44). Afin d'étudier ces variations plus en détails, nous avons déterminé le contenu en amylose et en phosphate, ainsi que la structure de l'amylopectine dans les fractions A, B, C et D. D'une part, nous n'avons observé aucune différence significative dans les contenus en amylose, la cristallinité et la longueur d'onde au maximum d'absorbance ( $\lambda_{max}$ ) du complexe amylopectine-I<sub>2</sub> (Tableau 3 et Annexe 5). En revanche, les contenus en protéines sont significativement modifiés dans les différentes fractions, avec une concentration maximale de 2,8  $\mu\text{g}\cdot\text{mg}^{-1}$  d'amidon dans la Fraction B (Tableau 3). Toutefois, ces changements restent minimes et ne sont pas corrélées avec le diamètre moyen des grains de la fraction.

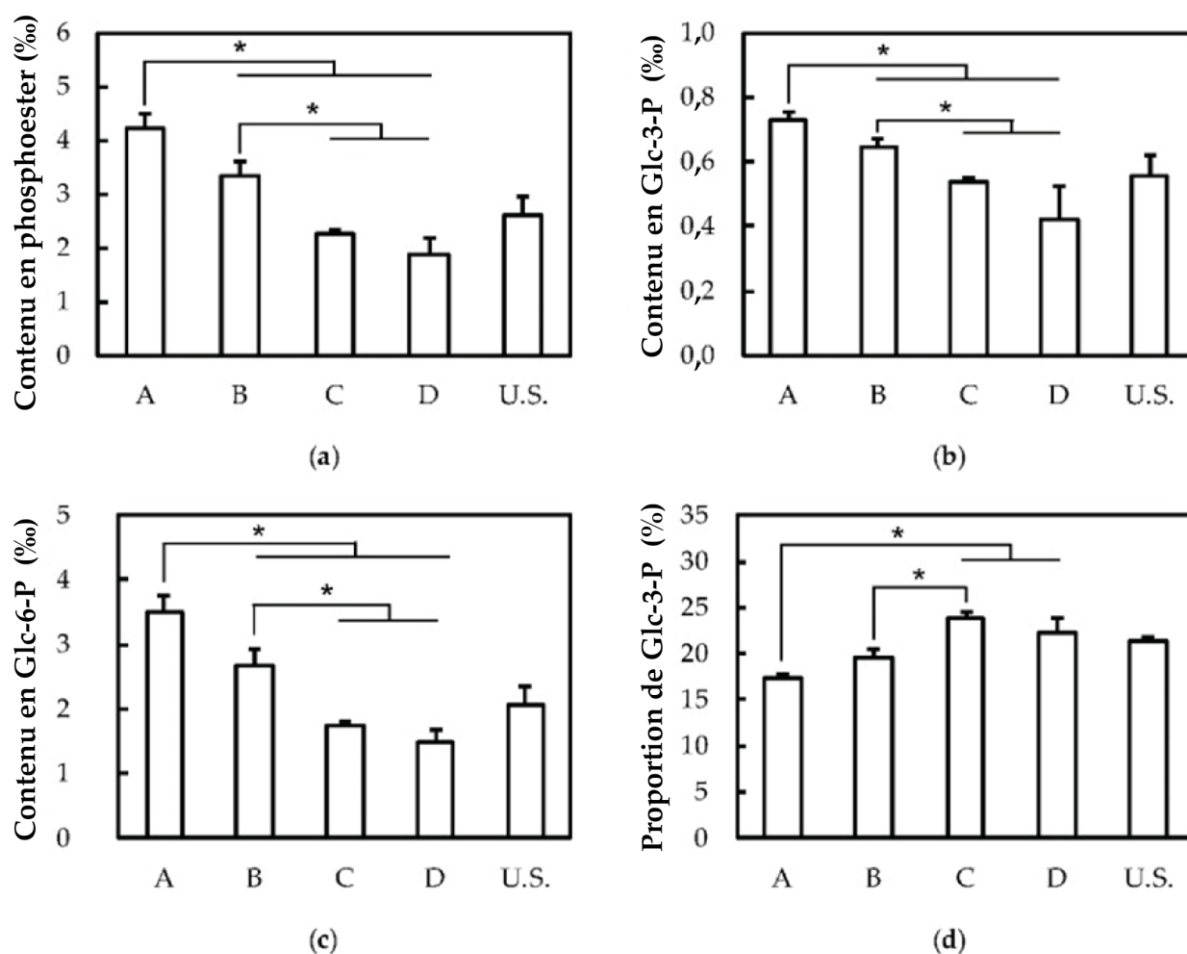




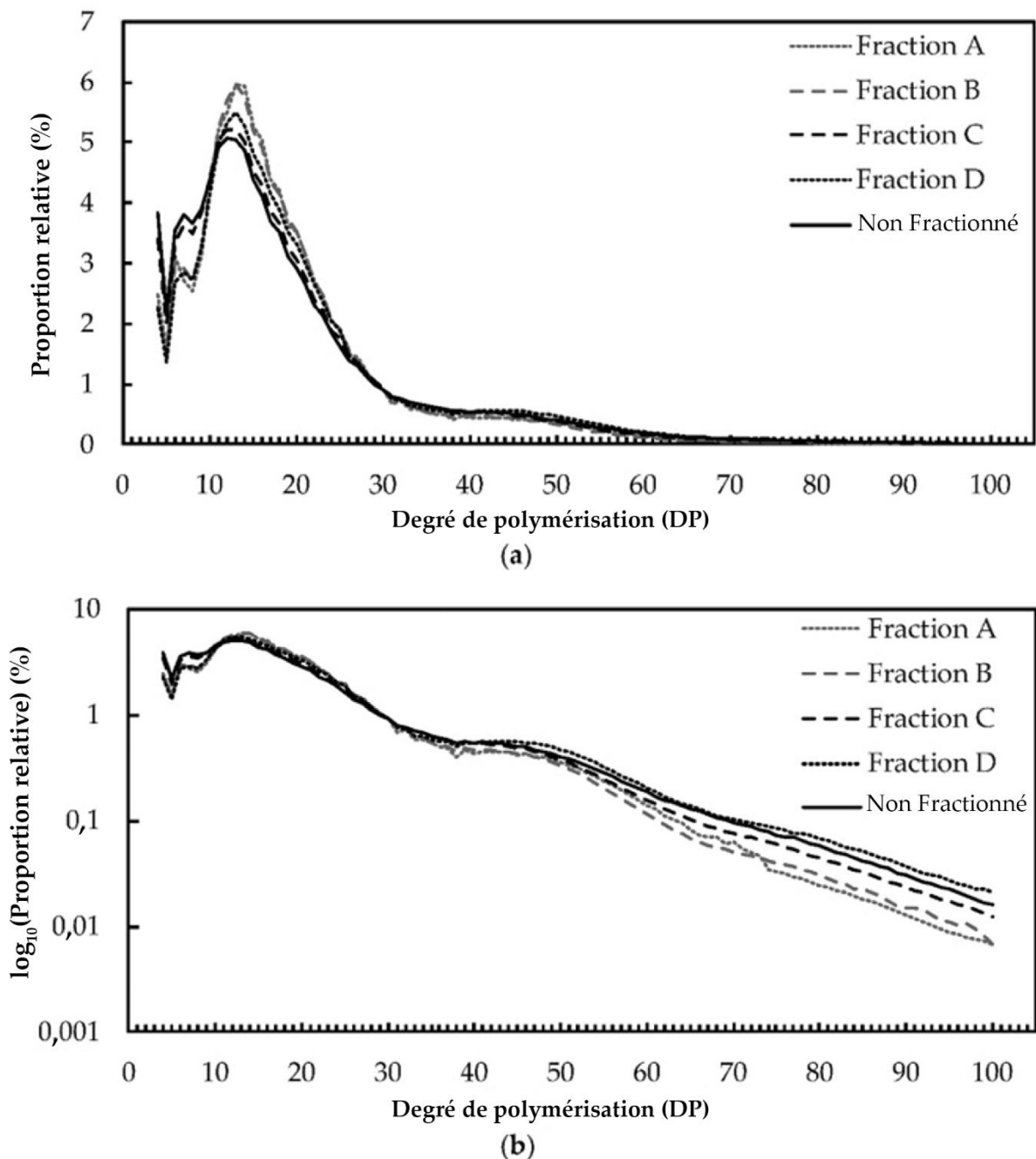
**Figure 44. Profils de diffraction aux rayons X des grains d'amidon de pomme de terre cv. Monlisa fractionnés en fonction de leur taille.** Les analyses de diffraction aux rayons-X ont été réalisées au CERMAV de Grenoble par Jean-Luc Puteaux. Les grains de la fraction A avait un diamètre de  $14,2 \pm 3,7 \mu\text{m}$ , ceux de la fraction B de  $24,5 \pm 6,5 \mu\text{m}$ ; les grains de la fraction C et D avaient un diamètres de  $47,7 \pm 12,8$  et  $61,8 \pm 17,4 \mu\text{m}$ , respectivement.

Le contenu en phosphate est directement corrélé à la taille des grains puisque son contenu diminue de manière significative lorsque la taille des grains augmente (Figure 45a). En particulier, la quantité de phosphate est diminuée d'un facteur 2 entre la fraction A et la fraction D avec des proportions respectives de 4,2 % et 1,9 % (Figure 45a). Les contenus en phosphates en position C3 et C6 suivent le même profil (Figure 45b et 45c). De manière à étudier l'hétérogénéité structurale de l'amylopectine dans les différentes fractions, l'amidon a été débranché (hydrolyse des liaisons  $\alpha$ -1.6 à l'aide d'une isoamylase et de pullulase) et sa distribution de longueurs de chaînes a été déterminée par électrophorèse capillaire assisté par fluorescence (FACE). Une faible diminution des glucanes de DP 10 à DP 24 ainsi qu'une

augmentation des longues chaînes de DP45 à DP 100 peuvent être observées avec l'augmentation du diamètre moyen des grain (Figure 46).



**Figure 45. Contenus en phosphate dans les différentes fractions.** 1 mg d'amidon des fractions A, B, C et D d'amidon non fractionné a été hydrolysé avant marquage à l'APTS et analyse par électrophorèse capillaire. (a) Contenu en glucose-phosphate total. (b) Contenu en glucose-3-phosphate. (c) Contenu en glucose-6-phosphate. (d) Proportion de glucose-3-phosphate par rapport au glucose-phosphate total. Les barres d'erreurs représentent l'écart type calculé à partir de 3 réplicats. Les astérisques indiquent une différence significative en se basant sur un test ANOVA dont la p-value  $\leq 0.01$ .



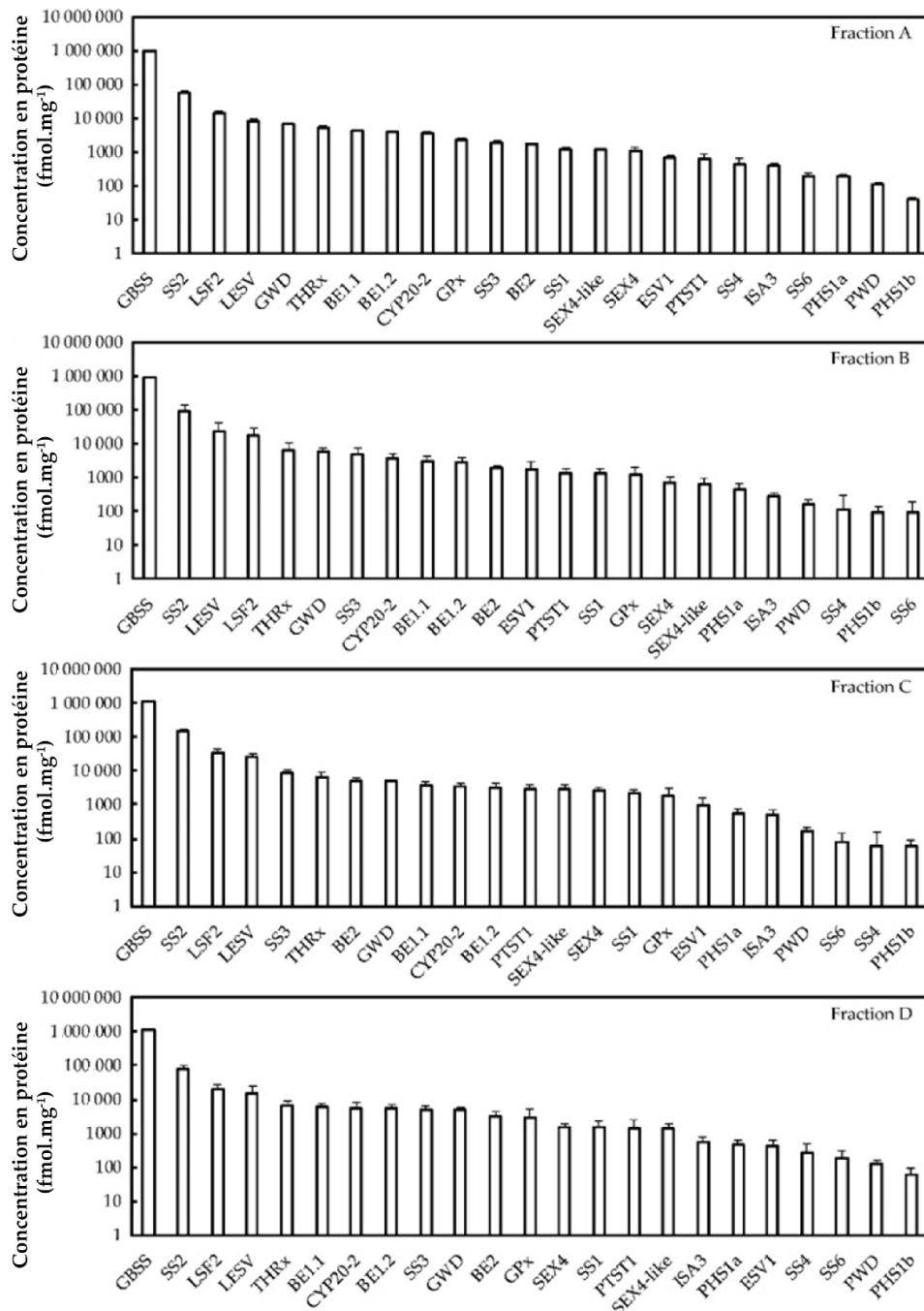
**Figure 46. Distribution de longueur de chaîne de l'amylopectine dans les différentes fractions.** 2 milligrammes d'amidon de chaque fraction de d'amidon non fractionné ont été dérivatisés avec de l'APTS après digestion enzymatique à l'isoamylase et la pullulanase. Les pics allant du DP 4 au DP 100 ont été intégrés et leur proportion relative a été calculé pour établir ces figures. (a) distribution de la longueur des chaînes représenté en échelle linéaire, mettant en valeur les DP courts. (b) distribution de longueur de chaîne représenté en échelle  $\log_{10}$ , mettant en valeur les DP longs. La proportion des différentes chaînes de glucanes a été moyennée sur la base d'analyses de 2 fractionnements réalisés indépendamment.

### 3. Protéomique quantitative comparative des différentes fractions

Chaque fraction a été analysée par spectrométrie de masse en méthode shotgun (sans séparation préalable en gel SDS-PAGE). Les grains d'amidon ont été traités au SDS 2 % (m/v) de manière à retirer les protéines de surface, potentiellement aspécifiques avant gélatinisation des grains et isolation des protéines. Les protéines ont été digérées par la méthode eFASP et les peptides résultants analysés à l'aide d'un spectromètre de masse en tandem de type quadrupole-orbitrap couplé à une chromatographie liquide (nanoLC) avec comme source d'ion un électronébuliseur (ESI). La quantification des protéines a été effectuée avec les programmes MaxQuant et Perseus à partir du score iBAQ, avec l'utilisation d'une banque de donnée composée uniquement des protéines associées au grain d'amidon de pomme de terre et précédemment identifiées (Tableau 3, Annexe 4 et 5). Cette banque de donnée restreinte permet de : i, réduire drastiquement le temps d'interrogation pour la quantification ; ii, d'éviter les biais de quantification liés à la présence de fragments de gènes dans la banque de donnée PGSC.

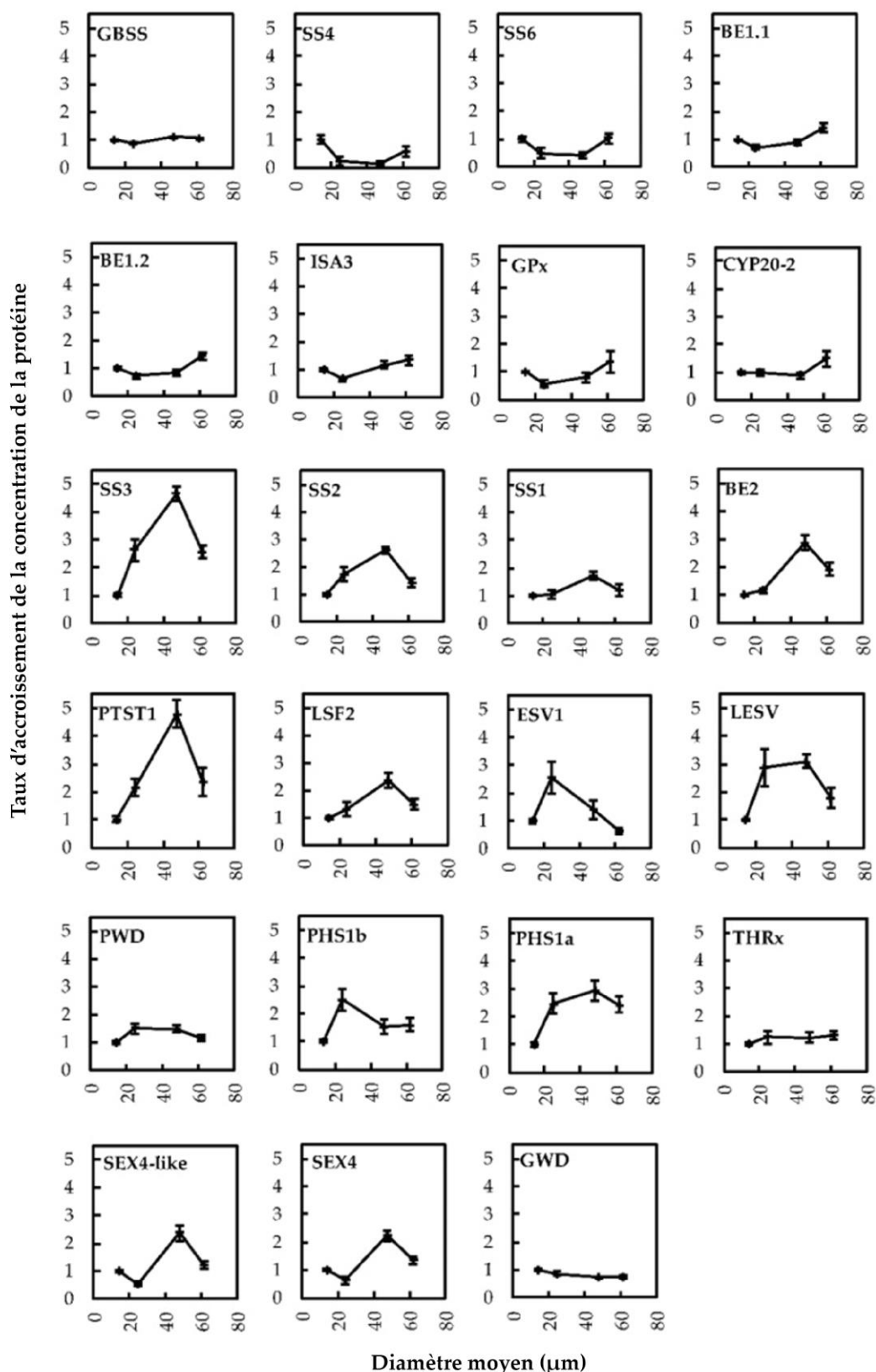
Le score d'abondance iBAQ (pour « intensity based absolute quantification ») de chaque protéine a été converti en fmol par mg d'amidon, en se basant sur les résultats des dosages protéiques au BCA (Tableau 4, Figure 47, Annexe 5). Quel que soit la fraction observée, GBSS et SS2 restent les deux protéines majoritaires, directement suivies par LSF2 et LESV qui sont deux fois moins abondantes que SS2 (Figure 47). La quantité relative de GWD diminue continuellement avec l'augmentation du diamètre des grains. Ceci est corrélé avec la diminution du contenu en glucose-6-phosphate précédemment décrite. En effet, GWD est responsable de la phosphorylation de l'amidon en position C6 (Figure 45, 48) (Ritte et al., 2006). L'abondance relative des autres protéines présente aussi des variations significatives entre les 4 fractions (Figure 47). Ceci est le cas pour les protéines les moins abondantes telles que ISA3, SS4, SS6, PHS1a et PHS1b. De manière à mieux illustrer ces changements dans l'abondance des protéines, la concentration de chaque protéine de la fraction A a été fixée à 1, et la variation de quantité de chacune des protéines a été calculée pour les fractions B, C et D (Figure 48). Les protéines SS3 et PTST1 montrent les modifications les plus importantes avec des valeurs supérieures à 4, alors que d'autres protéines comme GBSS et SS4 ne présentent que très peu de variations avec des valeurs inférieures à 0,2 (Figure 48). Dans l'ensemble, les protéines du métabolisme de l'amidon peuvent être classées en 4 groupes au regard de leurs profils de variation. SS4 et 6, BE1.1 et 1.2, ISA 3, la glutathion peroxydase GPx et la cyclophilline CYP20-2 varie de manière similaire avec une diminution de la concentration en protéines dans la fraction B suivi d'une augmentation dans les fractions C et D (Figure 48). SS1, 2 et 3, BE2,

PTST1, LSF2, ESV1, LESV, PWD, PHS1a et b ont quant à elle un profil opposé, avec une augmentation dans les fractions B et C suivi d'une diminution dans la fraction D (Figure 48). GWD se démarque des autres protéines avec un profil spécifique, sa concentration diminuant constamment avec l'augmentation du diamètre des grains. Enfin, les quantités de GBSS et THR<sub>x</sub> varient très peu entre les fractions (Figure 48). L'analyse de corrélation confirme les comportements similaires des protéines de chaque groupe. C'est par exemple le cas de PTST1 et LSF2 qui présentent un coefficient de Pearson de 1.0 (Tableau 5). SS4 et SS6, qui sont fortement corrélées entre elles, et sont anticorrélés avec PWD et LESV, illustrant ainsi l'opposition observée entre leurs motifs correspondants (Figure 48, Tableau 5).



**Figure 47. Protéomique quantitative sans marquage de chacune des fractions d'amidon.**

Les protéines ont été isolées à partir de 300 mg d'amidon des fractions A, B, C et D. Elles ont été digérées en peptides et analysées par spectrométrie de masse en tandem. L'interrogation des résultats a été faite avec les programmes MaxQuant et Perseus. Le score utilisé pour le calcul de la concentration de chaque protéine dans le grain est « Protein Concentration », exprimé en nM. Ce score est donné par Perseus en se basant sur le score iBAQ donné par MaxQuant. Ces concentrations sont ensuite exprimées en fmol normalisées sur la quantité d'amidon (mg) en utilisant les concentrations en protéines totales de chaque fraction (Tableau 3). Les barres d'erreurs représentent l'écart type calculé à partir de 3 répliques expérimentaux et 3 répliques techniques pour chacun d'entre eux. Les analyses statistiques ont été réalisées par ANVOA sur Perseus à partir du score LFQ donné par MaxQuant (Annexe 5).



**Figure 48. Taux d'accroissement de la concentration des protéines en fonction du diamètre moyen des grains d'amidon de chaque fraction.** Le taux d'accroissement a été calculé en normalisant le contenu en protéine dans la fraction A à 1. Les barres d'erreur représentent l'erreur type calculée à partir de 3 répliques expérimentaux, et 3 répliques techniques pour chaque.

**Tableau 5. Analyse de corrélation par coefficient de Pearson des variations concentrations en protéines liées aux grains d'amidon.** Les concentration moyennes pour chaque paramètre, en fmol.mg<sup>-1</sup> ont été utilisées pour l'analyse des corrélations. Les corrélations positives avec un score  $\rho > 0,95$  et  $p < 0,05$ ,  $\rho > 0,98$  and  $p < 0,02$ , ou  $\rho > 0,99$  et  $p < 0,01$  sont surlignées en rose. Les corrélations négatives jugés comme significative avec les mêmes seuils que décrits précédemment sont surlignées en jaune.

	GBSS	SS2	LESV	LSF2	THRx	GWD	SS3	CYP20.2	BE1.1	BE1.2	BE2	ESV1	PTST1	SS1	GPx	SEX4	SEX4.like	PHS1a	ISA3	PWD	SS4	PHS1b	SS6
GBSS	1,00	0,33	-0,08	0,61	0,14	-0,59	0,44	0,26	0,60	0,51	0,78	-0,79	0,55	0,69	0,63	0,88	0,84	0,23	0,94	-0,21	0,00	-0,64	0,20
SS2	0,33	1,00	0,90	0,95	0,51	-0,69	0,98	-0,43	-0,37	-0,45	0,84	0,32	0,97	0,90	-0,43	0,72	0,74	0,85	0,09	0,84	-0,90	0,32	-0,85
LESV	-0,08	0,90	1,00	0,73	0,64	-0,59	0,86	-0,39	-0,54	-0,57	0,56	0,65	0,78	0,62	-0,63	0,35	0,37	0,89	-0,25	0,99	-0,99	0,70	-0,94
LSF2	0,61	0,95	0,73	1,00	0,50	-0,79	0,97	-0,25	-0,10	-0,19	0,97	0,00	1,00	0,99	-0,13	0,90	0,90	0,80	0,40	0,64	-0,76	0,06	-0,64
THRx	0,14	0,51	0,64	0,50	1,00	-0,87	0,63	0,45	0,23	0,23	0,50	0,15	0,53	0,37	0,09	0,23	0,16	0,89	0,24	0,58	-0,73	0,64	-0,35
GWD	-0,59	-0,69	-0,59	-0,79	-0,87	1,00	0,81	-0,36	-0,36	-0,30	-0,84	0,17	-0,78	-0,72	-0,26	-0,67	-0,60	-0,89	-0,59	-0,48	0,68	-0,25	0,31
SS3	0,44	0,98	0,86	0,97	0,63	-0,81	1,00	-0,25	-0,20	-0,27	0,90	0,19	0,99	0,92	-0,26	0,77	0,77	0,91	0,25	0,79	-0,89	0,30	-0,76
CYP20.2	0,26	-0,43	-0,39	-0,25	0,45	-0,36	0,25	1,00	0,90	0,93	-0,06	-0,57	-0,28	-0,29	0,84	-0,13	-0,25	0,03	0,57	-0,46	0,29	-0,05	0,68
BE1.1	0,60	-0,37	-0,54	-0,10	0,23	-0,36	0,20	0,90	1,00	0,99	0,15	-0,87	-0,16	-0,07	0,99	0,18	0,08	-0,08	0,84	-0,63	0,42	-0,45	0,77
BE1.2	0,51	-0,45	-0,57	-0,19	0,23	-0,30	0,27	0,93	0,99	1,00	0,05	-0,83	-0,25	-0,17	0,98	0,08	-0,03	-0,12	0,77	-0,65	0,46	-0,40	0,81
BE2	0,78	0,84	0,56	0,97	0,50	-0,84	0,90	-0,06	0,15	0,05	1,00	-0,24	0,95	0,97	0,11	0,95	0,93	0,74	0,62	0,45	-0,62	-0,10	-0,43
ESV1	-0,79	0,32	0,65	0,00	0,15	0,17	0,19	-0,57	-0,87	-0,83	-0,24	1,00	0,08	-0,11	-0,92	-0,41	-0,35	0,29	-0,89	0,75	-0,57	0,82	-0,76
PTST1	0,55	0,97	0,78	1,00	0,53	-0,78	0,99	-0,28	-0,16	-0,25	0,95	0,08	1,00	0,97	-0,20	0,86	0,86	0,83	0,33	0,70	-0,81	0,14	-0,70
SS1	0,69	0,90	0,62	0,99	0,37	-0,72	0,92	-0,29	-0,07	-0,17	0,97	-0,11	0,97	1,00	-0,08	0,95	0,96	0,70	0,46	0,53	-0,66	-0,09	-0,57
GPx	0,63	-0,43	-0,63	-0,13	0,09	-0,26	0,26	0,84	0,99	0,98	0,11	-0,92	-0,20	-0,08	1,00	0,20	0,10	-0,20	0,84	-0,72	0,52	-0,57	0,83
SEX4	0,88	0,72	0,35	0,90	0,23	-0,67	0,77	-0,13	0,18	0,08	0,95	-0,41	0,86	0,95	0,20	1,00	0,99	0,51	0,69	0,23	-0,40	-0,38	-0,29
SEX4.like	0,84	0,74	0,37	0,90	0,16	-0,60	0,77	-0,25	0,08	-0,03	0,93	-0,35	0,86	0,96	0,10	0,99	1,00	0,48	0,61	0,26	-0,40	-0,38	-0,35
PHS1a	0,23	0,85	0,89	0,80	0,89	-0,89	0,91	0,03	-0,08	-0,12	0,74	0,29	0,83	0,70	-0,20	0,51	0,48	1,00	0,17	0,82	-0,94	0,59	-0,68
ISA3	0,94	0,09	-0,25	0,40	0,24	-0,59	0,25	0,57	0,84	0,77	0,62	-0,89	0,33	0,46	0,84	0,69	0,61	0,17	1,00	-0,38	0,14	-0,59	0,44
PWD	-0,21	0,84	0,99	0,64	0,58	-0,48	0,79	-0,46	-0,63	-0,65	0,45	0,75	0,70	0,53	-0,72	0,23	0,26	0,82	-0,38	1,00	-0,97	0,76	-0,96
SS4	0,00	-0,90	-0,99	-0,76	-0,73	0,68	0,89	0,29	0,42	0,46	-0,62	-0,57	-0,81	-0,66	0,52	-0,40	-0,40	-0,94	0,14	-0,97	1,00	-0,69	0,89
PHS1b	-0,64	0,32	0,70	0,06	0,64	-0,25	0,30	-0,05	-0,45	-0,40	-0,10	0,82	0,14	-0,09	-0,57	-0,38	-0,38	0,59	-0,59	0,76	-0,69	1,00	-0,60
SS6	0,20	-0,85	-0,94	-0,64	-0,35	0,31	0,76	0,68	0,77	0,81	-0,43	-0,76	-0,70	-0,57	0,83	-0,29	-0,35	-0,68	0,44	-0,96	0,89	-0,60	1,00



**Tableau 6. Analyses des corrélations par coefficient de Pearson entre les variations de concentration des protéines liés au grain d'amidon de pomme de terre et les variations en proportion relative de chaque degré de polymérisation issus de l'analyse de distribution de longueur de chaîne.** Les concentrations moyennes en fmol.mg<sup>-1</sup> et les proportions relatives de chaque DP en % ont été utilisées pour l'analyse de corrélation. Les corrélations positives avec un score  $\rho > 0.80$  et  $p < 0.2$ ,  $\rho > 0.98$  et  $p < 0.02$ , or  $\rho > 0.99$  et  $p < 0.01$  sont surlignées en rouge. Les corrélations négatives jugées significatives avec les mêmes seuils sont surlignées en jaune.

	GBSS	SS2	LESV	LSF2	THR3	GWD	SS3	CYP20.2	BE1.1	BE1.2	BE2	ESV1	PTST1	SS1	GPX	SEX4	SEX4.like	PHS1a	ISA3	PWD	SS4	PHS1b	SS6
DP4	0.58	0.81	0.50	0.86	0.00	-0.39	0.76	-0.59	-0.30	-0.40	0.81	-0.03	0.85	0.92	-0.27	0.87	0.92	0.42	0.27	0.44	-0.49	-0.25	-0.59
DP5	0.47	0.79	0.51	0.81	-0.09	-0.28	0.72	-0.70	-0.43	-0.52	0.73	0.07	0.80	0.86	-0.38	0.80	0.86	0.37	0.14	0.47	-0.48	-0.22	-0.64
DP6	0.12	0.59	0.40	0.52	-0.38	0.12	0.45	-0.93	-0.70	-0.77	0.38	0.30	0.52	0.58	-0.63	0.48	0.58	0.09	-0.22	0.42	-0.33	-0.17	-0.65
DP7	0.50	0.97	0.76	0.97	0.35	-0.64	0.95	-0.46	-0.30	-0.39	0.90	0.13	0.98	0.97	-0.32	0.85	0.88	0.73	0.24	0.69	-0.77	0.07	-0.75
DP8	0.56	0.95	0.71	0.98	0.34	-0.65	0.94	-0.44	-0.25	-0.34	0.92	0.06	0.98	0.98	-0.26	0.88	0.91	0.71	0.30	0.64	-0.73	0.01	-0.70
DP9	0.54	0.97	0.75	0.99	0.39	-0.68	0.96	-0.41	-0.25	-0.34	0.92	0.09	0.99	0.98	-0.27	0.87	0.89	0.75	0.28	0.67	-0.76	0.06	-0.72
DP10	0.25	0.97	0.85	0.89	0.30	-0.49	0.91	-0.64	-0.54	-0.62	0.75	0.39	0.91	0.87	-0.56	0.68	0.73	0.70	-0.04	0.82	-0.83	0.24	-0.89
DP11	-0.84	-0.37	-0.13	-0.60	-0.62	0.87	-0.53	-0.59	-0.73	-0.66	-0.77	0.63	-0.56	-0.60	-0.67	-0.71	-0.62	-0.55	-0.90	0.00	0.25	0.20	-0.14
DP12	-0.99	-0.45	-0.06	-0.71	-0.24	0.68	-0.56	-0.22	-0.54	-0.44	-0.86	0.70	-0.66	-0.78	-0.55	-0.93	-0.88	-0.37	-0.91	0.07	0.14	0.53	-0.07
DP13	-0.85	-0.75	-0.46	-0.92	-0.52	0.87	-0.84	-0.08	-0.29	-0.20	-0.99	0.37	-0.89	-0.93	-0.26	-0.95	-0.92	-0.70	-0.72	-0.33	0.53	0.17	0.30
DP14	-0.77	-0.76	-0.53	-0.91	-0.66	0.94	-0.87	-0.17	-0.31	-0.23	-0.97	0.30	-0.89	-0.89	-0.25	-0.88	-0.84	-0.80	-0.69	-0.41	0.61	0.02	0.33
DP15	-0.79	-0.80	-0.54	-0.94	-0.58	0.90	-0.89	-0.06	-0.24	-0.15	-0.99	0.28	-0.92	-0.94	-0.20	-0.93	-0.89	-0.77	-0.67	-0.42	0.61	0.07	0.38
DP16	-0.78	-0.79	-0.55	-0.93	-0.62	0.92	-0.88	-0.11	-0.27	-0.18	-0.98	0.28	-0.91	-0.92	-0.22	-0.91	-0.87	-0.79	-0.67	-0.42	0.62	0.03	0.36
DP17	-0.79	-0.83	-0.54	-0.96	-0.49	0.84	-0.89	0.05	-0.16	-0.07	-1.00	0.26	-0.94	-0.97	-0.13	-0.96	-0.94	-0.73	-0.63	-0.43	0.60	0.12	0.42
DP18	-0.79	-0.81	-0.54	-0.95	-0.56	0.88	-0.89	-0.03	-0.22	-0.12	-1.00	0.28	-0.93	-0.95	-0.18	-0.94	-0.91	-0.76	-0.66	-0.42	0.61	0.08	0.39
DP19	-0.88	-0.73	-0.36	-0.90	-0.27	0.70	-0.78	0.11	-0.19	-0.09	-0.96	0.41	-0.87	-0.95	-0.20	-1.00	-0.99	-0.53	-0.70	-0.25	0.42	0.35	0.29
DP20	-0.80	-0.83	-0.53	-0.96	-0.46	0.82	-0.89	0.07	-0.15	-0.05	-1.00	0.27	-0.94	-0.97	-0.13	-0.97	-0.95	-0.71	-0.63	-0.42	0.59	0.14	0.42
DP21	-0.80	-0.83	-0.53	-0.96	-0.45	0.81	-0.88	0.08	-0.14	-0.05	-1.00	0.27	-0.94	-0.98	-0.12	-0.97	-0.95	-0.70	-0.63	-0.41	0.58	0.15	0.42
DP22	-0.63	-0.93	-0.67	-0.99	-0.37	0.70	-0.94	0.35	0.15	0.25	-0.95	0.02	-0.98	-1.00	0.16	-0.92	-0.93	-0.71	-0.39	-0.59	0.70	0.04	0.64
DP23	-0.79	-0.83	-0.52	-0.96	-0.37	0.75	-0.88	0.17	-0.09	0.02	-0.99	0.26	-0.94	-0.99	-0.07	-0.98	-0.98	-0.66	-0.60	-0.41	0.57	0.19	0.44
DP24	-0.63	-0.83	-0.52	-0.90	-0.07	0.47	-0.80	0.52	0.23	0.34	-0.86	0.08	-0.88	-0.95	0.20	-0.91	-0.95	-0.48	-0.34	-0.44	0.52	0.24	0.57
DP25	-0.79	-0.83	-0.51	-0.96	-0.32	0.72	-0.87	0.21	-0.05	0.06	-0.98	0.24	-0.93	-0.99	-0.04	-0.98	-0.98	-0.63	-0.57	-0.40	0.55	0.22	0.45
DP26	-0.99	-0.24	0.16	-0.54	-0.15	0.58	-0.36	-0.36	-0.69	-0.61	-0.72	0.84	-0.47	-0.61	-0.71	-0.82	-0.77	-0.19	-0.97	0.29	-0.07	0.65	-0.29
DP27	-0.55	-0.83	-0.72	-0.90	-0.82	0.97	-0.92	-0.14	-0.16	-0.09	-0.90	0.04	-0.90	-0.83	-0.07	-0.74	-0.69	-0.94	-0.48	-0.62	0.80	-0.28	0.50
DP28	-0.14	-0.95	-0.98	-0.85	-0.70	0.73	-0.95	0.30	0.38	0.43	-0.72	-0.47	-0.89	-0.76	0.47	-0.52	-0.53	-0.94	0.03	-0.94	0.99	-0.58	0.88
DP29	0.10	0.77	0.88	0.69	0.93	-0.85	0.83	0.08	-0.10	-0.12	0.62	0.38	0.73	0.57	-0.23	0.36	0.32	0.99	0.07	0.83	-0.93	0.70	-0.67
DP30	0.87	0.60	0.18	0.79	-0.01	-0.47	0.62	-0.24	0.14	0.03	0.85	-0.47	0.74	0.88	0.19	0.97	0.98	0.29	0.66	0.08	-0.22	-0.56	-0.20
DP31	0.62	0.80	0.65	0.89	0.80	-0.98	0.90	0.19	0.24	0.17	0.92	-0.13	0.88	0.83	0.16	0.77	0.72	0.90	0.56	0.54	-0.74	0.20	-0.43
DP32	0.81	0.82	0.51	0.96	0.41	-0.79	0.87	-0.11	0.14	0.03	0.99	-0.28	0.93	0.98	0.12	0.98	0.97	0.67	0.63	0.39	-0.56	-0.19	-0.41
DP33	0.50	0.95	0.82	0.97	0.69	-0.88	0.99	-0.14	-0.08	-0.16	0.93	0.11	0.98	0.92	-0.15	0.78	0.77	0.93	0.34	0.74	-0.87	0.28	-0.68
DP34	0.75	0.87	0.60	0.98	0.52	-0.85	0.92	-0.08	0.11	0.01	1.00	-0.19	0.96	0.98	0.07	0.94	0.92	0.77	0.58	0.49	-0.66	-0.05	-0.48
DP35	0.81	0.80	0.52	0.94	0.54	-0.88	0.88	0.04	0.24	0.14	1.00	-0.30	0.92	0.94	0.20	0.94	0.91	0.74	0.67	0.40	-0.59	-0.11	-0.37
DP36	0.76	0.82	0.58	0.95	0.62	-0.91	0.91	0.06	0.21	0.13	0.99	-0.24	0.93	0.93	0.16	0.91	0.87	0.80	0.63	0.46	-0.65	-0.01	-0.41
DP37	0.80	0.76	0.51	0.91	0.63	-0.93	0.86	0.16	0.32	0.23	0.97	-0.32	0.89	0.90	0.27	0.90	0.85	0.77	0.71	0.38	-0.59	-0.06	-0.31
DP38	0.68	0.79	0.61	0.90	0.75	-0.97	0.89	0.18	0.27	0.19	0.94	-0.19	0.89	0.86	0.19	0.82	0.77	0.87	0.61	0.50	-0.70	0.12	-0.39
DP39	0.75	0.73	0.52	0.88	0.71	-0.96	0.85	0.24	0.36	0.28	0.95	-0.30	0.86	0.85	0.29	0.84	0.79	0.81	0.70	0.40	-0.61	0.02	-0.30
DP40	0.70	0.67	0.51	0.81	0.79	-0.99	0.80	0.36	0.42	0.36	0.89	-0.29	0.79	0.76	0.34	0.76	0.69	0.83	0.69	0.38	-0.60	0.10	-0.24
DP41	0.73	0.61	0.43	0.77	0.78	-0.98	0.75	0.42	0.50	0.43	0.86	-0.36	0.75	0.73	0.42	0.74	0.67	0.78	0.74	0.31	-0.54	0.05	-0.16
DP42	0.73	0.43	0.27	0.63	0.76	-0.94	0.60	0.59	0.65	0.60	0.76	-0.48	0.60	0.59	0.58	0.65	0.56	0.68	0.81	0.14	-0.39	-0.01	0.03
DP43	0.74	0.41	0.24	0.61	0.74	-0.92	0.57	0.61	0.68	0.63	0.75	-0.51	0.58	0.57	0.61	0.64	0.55	0.65	0.82	0.11	-0.35	-0.04	0.06
DP44	0.71	0.29	0.14	0.50	0.73	-0.88	0.47	0.71	0.76	0.72	0.66	-0.56	0.47	0.46	0.68	0.56	0.46	0.58	0.83	0.02	-0.26	-0.05	0.17
DP45	0.65	0.17	0.05	0.38	0.72	-0.83	0.36	0.80	0.81	0.79	0.55	-0.57	0.35	0.34	0.74	0.45	0.34	0.51	0.81	-0.06	-0.18	-0.04	0.27
DP46	0.67	0.12	-0.01	0.35	0.67	-0.79	0.31	0.81	0.85	0.82	0.53	-0.63	0.31	0.32	0.78	0.45	0.34	0.45	0.84	-0.14	-0.11	-0.11	0.33
DP47	0.65	0.00	-0.13	0.24	0.59	-0.71	0.19	0.87	0.91	0.89	0.44	-0.69	0.19	0.22	0.85	0.37	0.26	0.34	0.85	-0.25	0.01	-0.17	0.44
DP48	0.69	-0.01	-0.16	0.24	0.55	-0.69	0.18	0.85	0.92	0.89	0.45	-0.73	0.20	0.23	0.87	0.40	0.29	0.31	0.88	-0.28	0.04	-0.22	0.46
DP49	0.64	-0.10	-0.23	0.15	0.53	-0.63	0.09	0.90	0.94	0.93	0.36	-0.73	0.10	0.14	0.89	0.32	0.21	0.25	0.85	-0.34	0.10	-0.22	0.53
DP50	0.65	-0.14	-0.29	0.12	0.47	-0.59	0.05	0.90	0.96	0.95	0.34	-0.77	0.07	0.11	0.92	0.31	0.20	0.19	0.86	-0.40	0.16	-0.27	0.58

Tableau 6. Suite et fin

	GBSS	SS2	LESV	LSF2	THRX	GWD	SS3	CYP20.2	BE1.1	BE1.2	BE2	ESV1	PTST1	SS1	GPX	SEX4	SEX4.like	PHS1a	ISA3	PWD	SS4	PHS1b	SS6
DP51	0.66	-0.13	-0.28	0.13	0.47	-0.60	0.06	0.89	0.96	0.94	0.35	-0.77	0.08	0.13	0.92	0.33	0.21	0.20	0.87	-0.39	0.15	-0.28	0.57
DP52	0.61	-0.23	-0.37	0.03	0.41	-0.52	-0.04	0.92	0.98	0.97	0.26	-0.79	-0.02	0.03	0.95	0.25	0.13	0.11	0.84	-0.47	0.25	-0.31	0.65
DP53	0.69	-0.15	-0.32	0.13	0.42	-0.57	0.04	0.87	0.97	0.95	0.35	-0.82	0.07	0.14	0.94	0.35	0.24	0.15	0.89	-0.44	0.20	-0.34	0.60
DP54	0.69	-0.15	-0.33	0.13	0.41	-0.57	0.04	0.87	0.97	0.95	0.35	-0.82	0.07	0.14	0.94	0.35	0.24	0.15	0.90	-0.44	0.20	-0.35	0.60
DP55	0.71	-0.14	-0.33	0.14	0.40	-0.57	0.05	0.86	0.97	0.95	0.37	-0.83	0.08	0.15	0.94	0.37	0.26	0.14	0.91	-0.45	0.21	-0.37	0.60
DP56	0.73	-0.15	-0.36	0.13	0.36	-0.55	0.03	0.85	0.97	0.95	0.37	-0.85	0.07	0.15	0.95	0.38	0.27	0.11	0.92	-0.47	0.24	-0.41	0.62
DP57	0.75	-0.14	-0.35	0.15	0.35	-0.55	0.04	0.83	0.97	0.94	0.38	-0.87	0.09	0.18	0.95	0.41	0.30	0.11	0.93	-0.47	0.24	-0.43	0.61
DP58	0.75	-0.08	-0.29	0.20	0.42	-0.61	0.10	0.83	0.95	0.92	0.42	-0.83	0.14	0.21	0.93	0.42	0.32	0.19	0.93	-0.41	0.16	-0.37	0.56
DP59	0.77	-0.09	-0.32	0.20	0.37	-0.58	0.09	0.81	0.96	0.92	0.43	-0.86	0.14	0.22	0.94	0.44	0.34	0.15	0.94	-0.44	0.20	-0.42	0.57
DP60	0.78	-0.07	-0.29	0.22	0.39	-0.60	0.12	0.81	0.95	0.91	0.45	-0.85	0.16	0.24	0.93	0.46	0.36	0.18	0.94	-0.41	0.17	-0.41	0.55
DP61	0.78	-0.06	-0.28	0.23	0.40	-0.61	0.12	0.80	0.95	0.91	0.45	-0.84	0.17	0.25	0.92	0.46	0.36	0.19	0.94	-0.41	0.16	-0.40	0.55
DP62	0.78	0.00	-0.22	0.28	0.46	-0.66	0.19	0.80	0.93	0.89	0.50	-0.81	0.22	0.29	0.89	0.49	0.39	0.25	0.94	-0.34	0.09	-0.35	0.49
DP63	0.78	0.00	-0.22	0.28	0.45	-0.66	0.18	0.80	0.93	0.89	0.50	-0.81	0.22	0.29	0.90	0.49	0.39	0.25	0.94	-0.34	0.09	-0.36	0.49
DP64	0.80	0.01	-0.22	0.29	0.44	-0.66	0.19	0.78	0.92	0.88	0.51	-0.82	0.23	0.31	0.89	0.51	0.41	0.25	0.95	-0.35	0.10	-0.38	0.49
DP65	0.78	0.05	-0.16	0.32	0.50	-0.71	0.24	0.78	0.91	0.87	0.53	-0.78	0.27	0.33	0.87	0.51	0.41	0.31	0.93	-0.29	0.04	-0.31	0.44
DP66	0.78	0.10	-0.11	0.36	0.54	-0.74	0.28	0.77	0.89	0.85	0.56	-0.75	0.31	0.36	0.84	0.53	0.43	0.36	0.93	-0.24	-0.01	-0.28	0.40
DP67	0.79	0.09	-0.12	0.36	0.52	-0.73	0.28	0.76	0.89	0.85	0.57	-0.76	0.31	0.37	0.85	0.54	0.44	0.35	0.93	-0.25	0.00	-0.30	0.40
DP68	0.79	0.02	-0.20	0.30	0.46	-0.68	0.21	0.78	0.92	0.88	0.52	-0.81	0.25	0.31	0.89	0.51	0.41	0.27	0.94	-0.33	0.08	-0.36	0.47
DP69	0.75	0.10	-0.09	0.35	0.57	-0.75	0.28	0.79	0.88	0.84	0.55	-0.72	0.31	0.35	0.83	0.51	0.40	0.38	0.91	-0.22	-0.04	-0.23	0.38
DP70	0.76	0.01	-0.19	0.28	0.49	-0.68	0.19	0.81	0.92	0.89	0.50	-0.79	0.23	0.29	0.89	0.48	0.37	0.28	0.93	-0.32	0.07	-0.32	0.48
DP71	0.75	0.08	-0.11	0.33	0.56	-0.74	0.26	0.80	0.89	0.86	0.54	-0.73	0.29	0.33	0.84	0.49	0.38	0.36	0.91	-0.23	-0.02	-0.24	0.40
DP72	0.73	0.10	-0.07	0.35	0.60	-0.76	0.29	0.80	0.88	0.84	0.55	-0.70	0.30	0.34	0.82	0.49	0.38	0.40	0.89	-0.20	-0.05	-0.20	0.37
DP73	0.72	0.13	-0.03	0.37	0.63	-0.78	0.31	0.79	0.86	0.83	0.56	-0.67	0.33	0.35	0.80	0.49	0.38	0.43	0.88	-0.16	-0.09	-0.16	0.34
DP74	0.65	0.23	0.11	0.43	0.75	-0.86	0.41	0.76	0.78	0.75	0.59	-0.54	0.40	0.39	0.70	0.48	0.37	0.56	0.80	-0.01	-0.23	-0.01	0.21
DP75	0.64	0.22	0.11	0.42	0.75	-0.86	0.41	0.77	0.78	0.75	0.59	-0.54	0.39	0.38	0.70	0.47	0.36	0.56	0.80	-0.01	-0.24	0.00	0.21
DP76	0.64	0.21	0.10	0.41	0.74	-0.85	0.39	0.78	0.79	0.76	0.58	-0.54	0.38	0.37	0.71	0.46	0.36	0.55	0.80	-0.02	-0.22	-0.01	0.23
DP77	0.64	0.22	0.11	0.42	0.75	-0.85	0.40	0.77	0.78	0.75	0.58	-0.54	0.39	0.38	0.70	0.47	0.36	0.56	0.80	-0.01	-0.23	0.00	0.22
DP78	0.63	0.19	0.08	0.39	0.74	-0.84	0.38	0.79	0.80	0.77	0.56	-0.55	0.36	0.35	0.72	0.44	0.34	0.54	0.80	-0.03	-0.21	-0.01	0.24
DP79	0.61	0.16	0.07	0.36	0.74	-0.82	0.35	0.81	0.81	0.78	0.53	-0.55	0.33	0.32	0.72	0.41	0.30	0.52	0.79	-0.05	-0.19	0.00	0.26
DP80	0.63	0.19	0.09	0.39	0.75	-0.84	0.37	0.79	0.80	0.77	0.55	-0.54	0.36	0.34	0.71	0.44	0.33	0.54	0.79	-0.03	-0.21	0.00	0.24
DP81	0.62	0.16	0.06	0.36	0.73	-0.82	0.35	0.81	0.81	0.79	0.53	-0.56	0.33	0.32	0.73	0.42	0.31	0.51	0.80	-0.06	-0.18	-0.01	0.27
DP82	0.63	0.17	0.07	0.37	0.73	-0.83	0.36	0.80	0.81	0.78	0.54	-0.56	0.34	0.33	0.73	0.43	0.32	0.52	0.80	-0.05	-0.19	-0.01	0.26
DP83	0.62	0.16	0.06	0.37	0.73	-0.82	0.35	0.81	0.81	0.78	0.54	-0.55	0.34	0.32	0.73	0.42	0.31	0.52	0.79	-0.05	-0.19	-0.01	0.27
DP84	0.64	0.16	0.04	0.37	0.72	-0.82	0.34	0.81	0.82	0.79	0.54	-0.58	0.33	0.33	0.74	0.44	0.33	0.50	0.81	-0.07	-0.17	-0.04	0.28
DP85	0.60	0.15	0.06	0.35	0.74	-0.82	0.34	0.82	0.81	0.78	0.52	-0.54	0.32	0.31	0.73	0.40	0.29	0.52	0.78	-0.05	-0.19	0.01	0.27
DP86	0.61	0.14	0.05	0.35	0.73	-0.81	0.33	0.82	0.82	0.79	0.52	-0.56	0.32	0.30	0.74	0.40	0.29	0.51	0.79	-0.07	-0.18	-0.01	0.28
DP87	0.63	0.16	0.05	0.36	0.73	-0.82	0.35	0.81	0.81	0.79	0.54	-0.56	0.33	0.32	0.74	0.43	0.32	0.51	0.80	-0.06	-0.18	-0.02	0.27
DP88	0.63	0.14	0.03	0.35	0.71	-0.81	0.33	0.82	0.83	0.80	0.52	-0.58	0.31	0.31	0.75	0.42	0.31	0.49	0.80	-0.08	-0.16	-0.03	0.29
DP89	0.65	0.13	0.02	0.35	0.70	-0.80	0.32	0.82	0.84	0.81	0.53	-0.60	0.32	0.31	0.76	0.43	0.32	0.48	0.82	-0.10	-0.14	-0.06	0.31
DP90	0.65	0.15	0.03	0.37	0.71	-0.81	0.34	0.81	0.83	0.80	0.54	-0.59	0.33	0.33	0.75	0.44	0.33	0.49	0.82	-0.09	-0.16	-0.05	0.29
DP91	0.62	0.15	0.06	0.36	0.73	-0.82	0.34	0.81	0.81	0.79	0.53	-0.55	0.33	0.31	0.73	0.41	0.30	0.51	0.79	-0.06	-0.18	-0.01	0.27
DP92	0.62	0.19	0.10	0.39	0.75	-0.84	0.38	0.79	0.79	0.76	0.55	-0.53	0.36	0.34	0.71	0.43	0.32	0.55	0.78	-0.02	-0.22	0.01	0.23
DP93	0.63	0.17	0.06	0.37	0.73	-0.83	0.35	0.80	0.81	0.78	0.54	-0.56	0.34	0.33	0.73	0.43	0.32	0.52	0.80	-0.06	-0.19	-0.02	0.27
DP94	0.65	0.16	0.04	0.37	0.71	-0.82	0.35	0.80	0.82	0.79	0.55	-0.58	0.34	0.33	0.74	0.44	0.33	0.50	0.82	-0.07	-0.17	-0.04	0.28
DP95	0.63	0.18	0.08	0.39	0.74	-0.84	0.37	0.79	0.80	0.77	0.56	-0.55	0.36	0.34	0.72	0.44	0.33	0.53	0.80	-0.04	-0.20	-0.01	0.25
DP96	0.63	0.15	0.04	0.36	0.72	-0.81	0.34	0.81	0.82	0.80	0.53	-0.58	0.32	0.32	0.75	0.43	0.32	0.50	0.81	-0.08	-0.17	-0.03	0.29
DP97	0.61	0.18	0.09	0.38	0.75	-0.84	0.37	0.80	0.79	0.77	0.54	-0.53	0.35	0.33	0.71	0.42	0.31	0.54	0.78	-0.03	-0.21	0.02	0.24
DP98	0.62	0.18	0.08	0.38	0.75	-0.83	0.37	0.80	0.80	0.77	0.54	-0.54	0.35	0.33	0.71	0.43	0.32	0.54	0.79	-0.03	-0.21	0.01	0.25
DP99	0.63	0.11	0.00	0.33	0.69	-0.79	0.30	0.83	0.84	0.82	0.51	-0.60	0.29	0.29	0.77	0.41	0.30	0.46	0.81	-0.12	-0.13	-0.06	0.32
DP100	0.69	0.12	-0.02	0.35	0.66	-0.79	0.31	0.81	0.85	0.82	0.54	-0.64	0.31	0.33	0.79	0.4	0.35	0.45	0.85	-0.14	-0.11	-0.12	0.34

**Tableau 7. Analyses des corrélations par coefficient de Pearson entre les variations de concentrations des protéines liées au grain d'amidon, les variations de contenu en phosphate et la morphologie des grains.** Les concentrations moyennes en fmol.mg<sup>-1</sup>, les proportions relatives de chaque phosphoester en ‰, les diamètres (µm) et rotondités ellipsoïdes moyens ont été utilisées pour l'analyse de corrélation. Les corrélations positives avec un score  $\rho > 0.90$  et  $p < 0.1$ ,  $\rho > 0.98$  et  $p < 0.02$ , or  $\rho > 0.99$  et  $p < 0.01$  sont surlignées en rouge. Les corrélations négatives jugées significatives avec les mêmes seuils sont surlignées en jaune.

	GBSS	SS2	LESV	LSF2	THR <sub>x</sub>	GWD	SS3	CYP20.2	BE1.1	BE1.2	BE2	ESV1	PTST1	SS1	GP <sub>x</sub>	SEX4	SEX4.like	PHS1a	ISA3	PWD	SS4	PHS1b	SS6
Total-P (‰)	-0,61	-0,56	-0,46	-0,69	-0,87	0,99	-0,71	-0,50	-0,50	-0,45	-0,78	0,28	-0,68	-0,63	-0,40	-0,62	-0,53	-0,82	-0,66	-0,35	0,57	-0,19	0,16
C3-P (‰)	-0,61	-0,35	-0,26	-0,52	-0,83	0,92	-0,53	-0,69	-0,68	-0,64	-0,65	0,42	-0,50	-0,46	-0,58	-0,51	-0,41	-0,68	-0,74	-0,14	0,38	-0,11	-0,07
C6-P (‰)	-0,61	-0,59	-0,49	-0,71	-0,87	0,99	-0,73	-0,47	-0,47	-0,42	-0,80	0,26	-0,70	-0,65	-0,37	-0,63	-0,55	-0,83	-0,65	-0,37	0,59	-0,20	0,19
Diameter (µm)	0,69	0,42	0,28	0,60	0,79	-0,94	0,58	0,62	0,65	0,61	0,73	-0,45	0,57	0,55	0,57	0,60	0,51	0,69	0,78	0,16	-0,40	0,04	0,03
Ellipsoid roundness	-0,75	-0,68	-0,48	-0,83	-0,74	0,98	-0,80	-0,33	-0,42	-0,35	-0,91	0,33	-0,81	-0,80	-0,35	-0,80	-0,74	-0,80	-0,72	-0,36	0,58	-0,03	0,23

Des analyses de corrélations ont aussi été réalisées entre les concentrations de chaque protéines prise individuellement et les autres paramètres ayant été mesurés dans cette étude (Tableau 5 ; 6). Les concentrations de SS1, 2 et 3 montrent ainsi une corrélation négative avec l'abondance relative des chaînes d'amylopectine de DP 13 à DP 28, alors qu'elles sont corrélées positivement avec celles de DP 4 à 10 et de DP 31 à 40 (Tableau 6). Du fait de la corrélation positive des concentrations de SS1, 2 et 3 avec celles de LSF2, BE2, PTST1, SEX4, or SEX4-like, ces dernières montrent aussi une corrélation similaire avec les longueurs de chaînes de l'amylopectine décrites précédemment (Tableau 6). D'autre part, BE1.1, BE1.2, CYP20.2, GPX, et ISA3 corrélaient positivement avec la proportion des chaînes les plus longues de l'amylopectine (supérieures à DP 42). Enfin, l'abondance de GWD ne corréla avec aucune autre protéine associée à l'amidon de pomme de terre (Tableau 5). Toutefois, elle est positivement corrélée avec la quantité de phosphate dans l'amidon, plus particulièrement avec le phosphate en position C6, ainsi qu'avec la morphologie des grains (Tableau 7).

### **III. Dynamique d'association des protéines avec le grain d'amidon**

Après avoir démontré que les contenus en protéines fluctuent en fonction de la morphologie des grains, nous avons cherché à comprendre si cette observation résulte de l'état de croissance des grains, ou s'il existe un déterminisme du protéome sur leur morphologie. Dans cette optique, nous avons mis en œuvre des expériences de marquage à l'azote lourd ( $^{15}\text{N}$ ). Nous avons apporté du  $^{15}\text{N}$  exogène à des plants de pommes de terre à différents instants de la tubérisation, puis nous avons observé le taux d'incorporation au point final, c'est-à-dire au moment de la récolte après 6 semaines de croissance (Figure 43). De cette façon, nous pouvons savoir si l'ensemble des protéines du grain d'amidon s'associent au même moment ou non. Ces travaux ont été initiés au début de l'année 2019. La culture des plantes et le marquage au  $^{15}\text{N}$  a été réalisé par François Sevestre, Maud Facon, Walaa Derbass et moi-même. Walaa Derbass, Benoît Poullain et moi-même nous sommes occupés de l'extraction de l'amidon, des protéines, et de la préparation des échantillons à l'analyse en spectrométrie de masse. L'analyse en spectrométrie de masse a été réalisée par Fabrice Bray et Stéphanie Flament. Ces travaux sont toujours en cours, il est encore nécessaire d'étudier le taux d'incorporation au niveau de la fraction soluble et au niveau de l'interface grain-stroma pour tenter de modéliser la dynamique d'association des protéines au grain.

## 1. Marquage et sélection des peptides

Les plantes utilisées pour ces expériences ont été cultivées dans 5 conditions différentes. Les deux premières conditions sont des contrôles. L'une consiste à cultiver les plants de pomme de terre sur un terreau pauvre en substances azotées. Dès le début, et tout au long de la tubérisation, les plantes sont arrosées avec un milieu nutritif contenant une source d'azote non marquée. Le deuxième contrôle est identique au précédent mis à part que la source d'azote utilisée est entièrement marquée à l'azote lourd  $^{15}\text{N}$ . Dans les 3 autres conditions, les plantes ont été supplémentées en  $^{15}\text{N}$  à une période définie. D'une part, les plantes ont été enrichies seulement au début de la tubérisation. D'autre part, les plantes ont été enrichies au milieu de la période de croissance des tubercules. Dans le 3ème cas, la supplémentation en  $^{15}\text{N}$  a eu lieu à la fin de la tubérisation. Chaque tubercule a été plongé dans l'azote liquide juste après la récolte et conservé à  $-80\text{ }^\circ\text{C}$ . Après extraction et purification de l'amidon, les protéines associées aux grains ont été extraites et digérées par la trypsine avant leur analyse par spectrométrie de masse. Les protéines des plantes contrôles sans marquage ont été identifiées à l'aide de Thermo® Proteome Discoverer. Pour chaque protéine identifiée dans une plante, trois peptides ont été sélectionnés à partir du fichier "Peptide Spectrum Matches" (PSM), le fichier contenant toutes les occurrences de spectre où le logiciel Thermo® Proteome Discoverer a pu identifier un peptide. Cette sélection s'est faite de manière empirique sur la base de la lisibilité des spectres de masse (bonne visibilité du massif de pic, faible quantité de pic parasite, forte intensité du signal du pic majoritaire...) à partir du logiciel Xcalibur (Tableau 8).

**Tableau 8. Peptides sélectionnés pour calculer le taux d'incorporation du  $^{15}\text{N}$ .** Trois peptides ont été choisis à partir du fichier PSMs de Thermo® Proteome Discoverer PSMs file pour chaque protéine identifiée. Le contenu en azote (N) a été déterminé à partir de la séquence du peptide et à l'aide de l'outil en ligne PROTEOMICS TOOLKIT (<http://db.systemsbio.net:8080/proteomicsToolkit/IsotopeServlet.html>). Un "m" minuscule dans la séquence représente une méthionine oxydée. m/z : ratio masse sur charge. RT : temps de rétention.

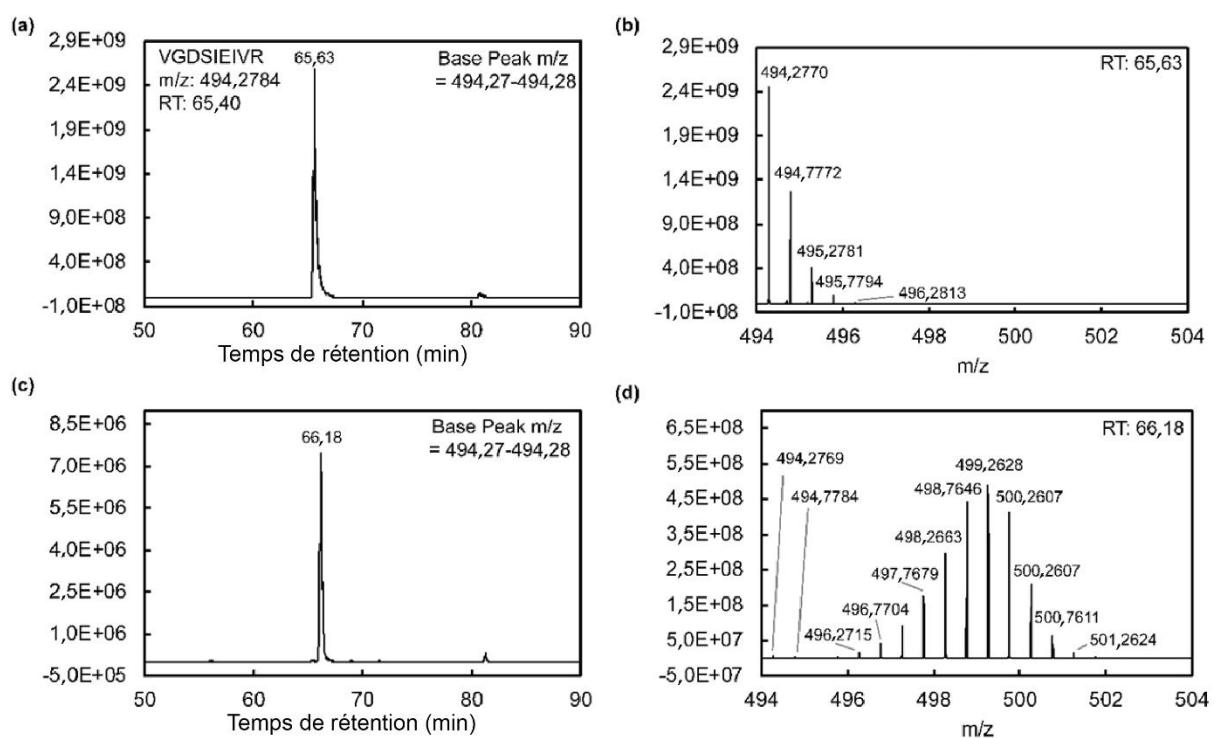
Protéine	m/z [Da]	RT [min]	Séquence du peptide	Charge	nombre d'azote
GBSS	494,2784	65,40	VGDSIEIVR	2	12
GBSS	558,3246	68,41	GVELDSVLRK	2	14
GBSS	410,2236	40,70	VmTISPR	2	10
SS2	785,4019	57,70	SSITASSQISSTVSSK	2	18
SS2	836,3839	74,27	ATGENSGEAASADESNDALQVTIEK	3	29
SS2	766,8831	74,31	NTPPAISQDFVESK	2	17
GWD	828,3912	70,88	AAETPFSASSSDGLTSK	2	18
GWD	541,8086	63,87	VLEPPALSK	2	11

GWD	629,8352	59,09	VLSDDINQGVAK	2	15
LESV	656,3723	78,35	QGATGNVISIVPR	2	18
LESV	535,7961	56,04	ILQVSPEER	2	13
LESV	433,7528	45,42	PPSPVIEK	2	9
SS3	628,3539	87,51	LAEENLLQGIR	2	16
SS3	907,9976	106,48	LYEILQVDVEPQQLK	2	19
SS3	748,8407	69,43	FLEDEDEINGSTK	2	15
BE1	531,2666	81,11	EFADDVLPR	2	12
BE1	822,4009	115,05	YEGPLEEFAQGYLEK	2	16
BE1	789,8646	43,84	IYEAHVGMSSSEPR	2	19
BE2	861,9057	72,61	TLNTSEETIIDESDR	2	19
BE2	783,3823	113,63	VALDSDDPLFGGFR	2	18
BE2	795,8761	45,17	IYESHIGMSSPEPK	2	17
SS1	734,3937	62,45	AHALDTGETVNVLK	2	18
SS1	568,2905	68,24	YANAVDL DVR	2	14
SS1	597,2857	38,32	SVSDGSSIAENK	2	14
CYP-20.2	1004,0111	104,47	IVIGLYGDDVPQTAENFR	2	23
CYP-20.2	704,8790	101,88	VYFDISIGNPVGK	2	15
CYP-20.2	366,6748	28,77	DSAFHR	2	11
SEX4	545,3114	81,31	SATADILTGLK	2	12
SEX4	624,8114	42,09	TYIHcTAGLGR	2	16
SEX4	407,7307	33,83	LNNAINR	2	13
SEX4-like	416,8761	42,48	TYIHcTAGIGR	3	16
SEX4-like	717,3229	56,44	LTGDDPDLTSDER	2	16
SEX4-like	530,5822	51,19	RLTGDDPDLTSDER	3	20
LSF2	482,7221	73,76	DFDPDSL R	2	11
LSF2	583,8067	62,15	SIQGATYDLAK	2	13
LSF2	386,7372	29,69	KLIQDR	2	11
ISA3	842,3771	62,04	AFTADETSSLDQDQR	2	20
ISA3	483,2566	54,10	IAGSADLYR	2	12
ISA3	419,7480	34,92	LSHNVLR	2	13

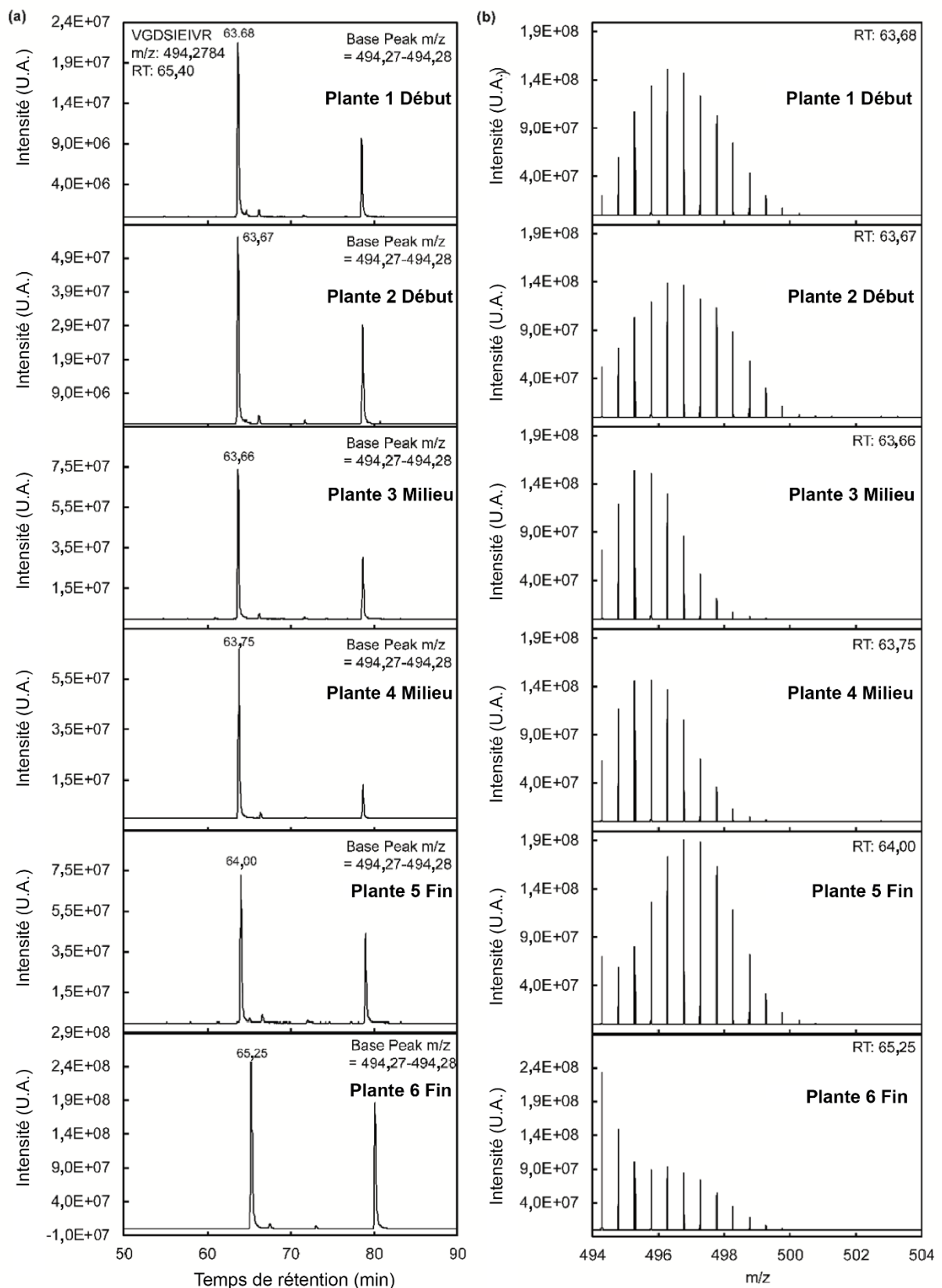
## 2. Identification des massifs de pics correspondants aux peptides d'intérêt

Les massifs de pics ont été repérés grâce aux ratios masse sur charge ( $m/z$ ) et aux temps de retentions (RT) de chaque peptide à l'aide du programme Thermo® Xcalibur. Dans les prochaines figures présentées, le peptide VGDSIEIVR, un peptide unique de la protéine GBSS, a été utilisé comme exemple, celui-ci possédant les caractéristiques idéales pour l'interprétation des spectres avec peu ou pas de pics parasites, et une très bonne intensité (Figure 49, Figure 50). La comparaison entre les contrôles non-marqué et marqué au  $^{15}\text{N}$  a montré un taux d'incorporation de 78,6 %. Des décalages de plusieurs minutes ont été observés entre les RTs des analyses par rapport au contrôle non marqué. Toutefois, grâce à la très grande abondance de GBSS (environ 80% des protéines présentes dans le grain d'amidon), Les pics ont facilement

pu être retrouvés sur le chromatogramme (Figures 49a, 49c et 50a). Pour les protéines minoritaires, l'utilisation du décalage constaté avec les peptides de GBSS a permis de retrouver les pics correspondants dans les différentes analyses. L'abondance relative de chacun des pics du contrôle sans marquage (Figure 49b) a été utilisée pour soustraire l'intensité du pic correspondant à l'abondance naturelle des isotopes lourds présents dans la plante à l'intensité totale observé sur les spectres des expériences de marquages. Le résultat permet d'estimer l'intensité des pics dû à l'incorporation de  $^{15}\text{N}$  exogène. Par la suite, le taux d'incorporation moyen est calculé à partir de l'intensité relative de chaque pic et divisé par le nombre d'azote présent dans la séquence du peptide.



**Figure 49. Chromatogrammes et spectres de masse des expériences contrôles.** Le peptide VGDSIEIVR utilisé pour cette figure a été identifié avec Thermo® Proteome Discoverer 2.4 dans les expériences contrôles  $^{14}\text{N}$  au RT = 65,40 min, m/z 494,2784, charge (Z) = 2. Les représentations des chromatogrammes et des spectres de masse ont été obtenues en exportant les valeurs de Thermo® XCalibur vers Microsoft® Excel. (a) Chromatogramme du contrôle  $^{14}\text{N}$  non marqué. (b) Spectre de masse du contrôle  $^{14}\text{N}$  non marqué. (c) Chromatogramme du contrôle  $^{15}\text{N}$  marqué. (d) Spectre de masse du contrôle  $^{15}\text{N}$  marqué.



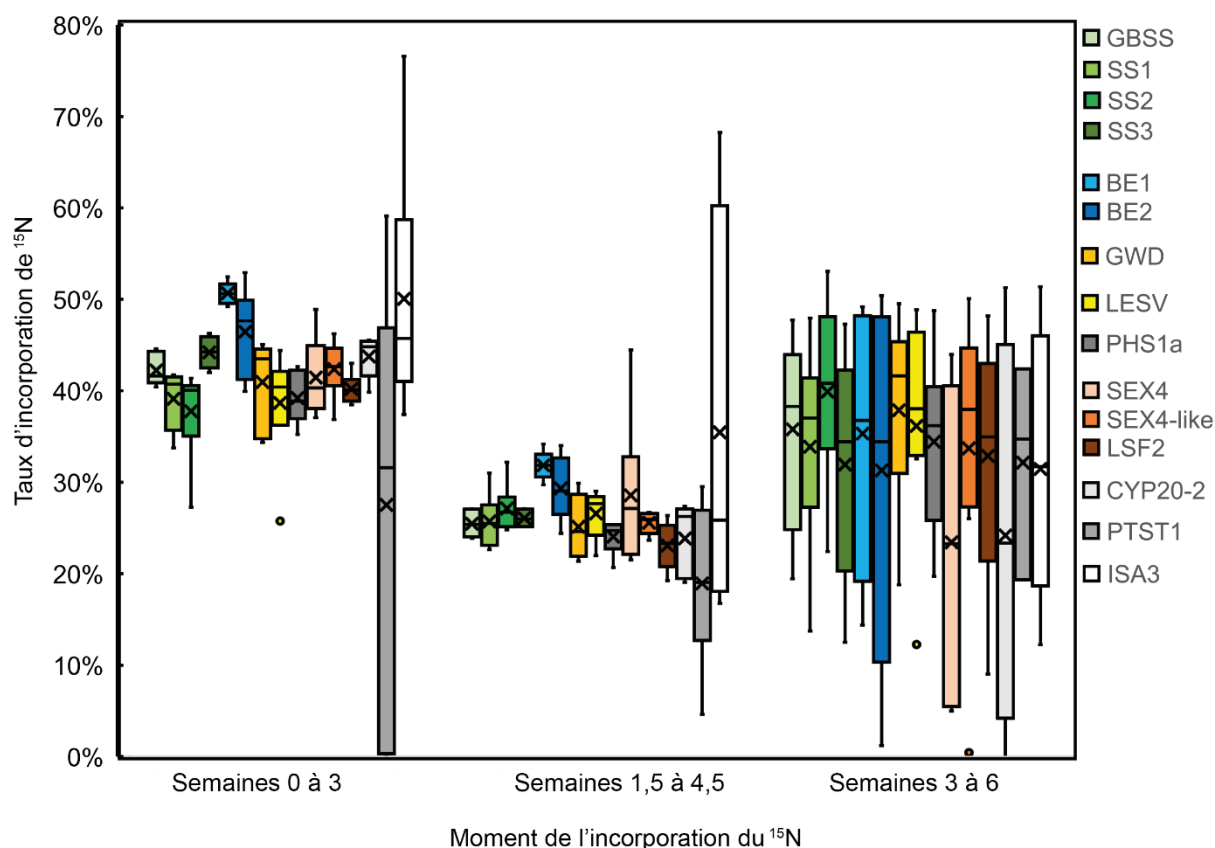
**Figure 50. Chromatogrammes et spectres de masses des expériences d'incorporation et de dilution du  $^{15}\text{N}$  pendant 6 semaines après l'apparition des stolons.** Le peptide VGDSIEIVR utilisé pour cette figure a été identifié avec Thermo® Proteome Discoverer 2.4 dans les expériences contrôles  $^{14}\text{N}$  au RT = 65,40 min, m/z 494,2784, charge (Z) = 2. La représentation des chromatogrammes et les spectres de masses a été faite en exportant les valeurs de Thermo®



XCalibur vers Microsoft® Excel. (a) Chromatogrammes des expériences d'incorporation et de chasse. (b) Spectres de masse des expériences d'incorporation et de chasse.

### **3. Taux d'incorporation pour chacune des protéines.**

Les taux d'incorporation ont été calculés pour l'ensemble des peptides de chaque expérience (Figure 51.). L'ensemble des protéines, excepté PTST1 et ISA3 pour lesquelles les fortes divergences entre les massifs de pics des différents peptides rendent l'interprétation difficile, présentent un taux d'incorporation plus élevé quand le marquage a été effectué au cours des 3 premières semaines de tubérisation. Ce taux d'incorporation oscille autour de 40 % pour la plupart des protéines. On note toutefois une différence significative entre les enzymes (ANOVA  $p$ -value= 0,0015). Les enzymes de branchement ont en effet un taux d'incorporation d'environ 50% lorsque le marquage est réalisé en début de tubérisation, l'incorporation pour BE1 est par ailleurs significativement supérieure (T-test  $p$ -value <0.05) à l'ensemble des autres enzymes (figure 51, Annexe 7). Au niveau des amidons synthase, on remarque aussi un taux d'incorporation significativement supérieur (T-test  $p$ -value <0.05) pour SS3 par rapport aux autres amidons synthases. Le taux d'incorporation de SS3 est aussi significativement plus élevé que celui de PHS1a et LSF2. Le taux d'incorporation est plus faible (25 à 30%) pour l'ensemble des protéines quand le marquage a lieu au milieu de la tubérisation (commençant 1,5 semaines après l'apparition des stolons, et s'arrêtant 1,5 semaines avant la récolte). Les enzymes de branchements possèdent toujours un taux d'incorporation plus élevé pendant cette période, mais SS3 ne montre cette fois pas de différence significative avec les autres amidons synthases. Enfin, pour les incorporations ayant eu lieu 3 semaines avant la récolte, le taux d'incorporation est soit très faible (15 %), soit de même niveau que lors du marquage en début de tubérisation, cela rend les résultats difficilement interprétables dans l'état. Il est probable que cette forte divergence soit le fait de condition de croissance différentes entre les différentes plantes récoltées. On remarque toutefois que les valeurs pour les enzymes de branchement restent plus faibles que celles constatées en début de tubérisation.



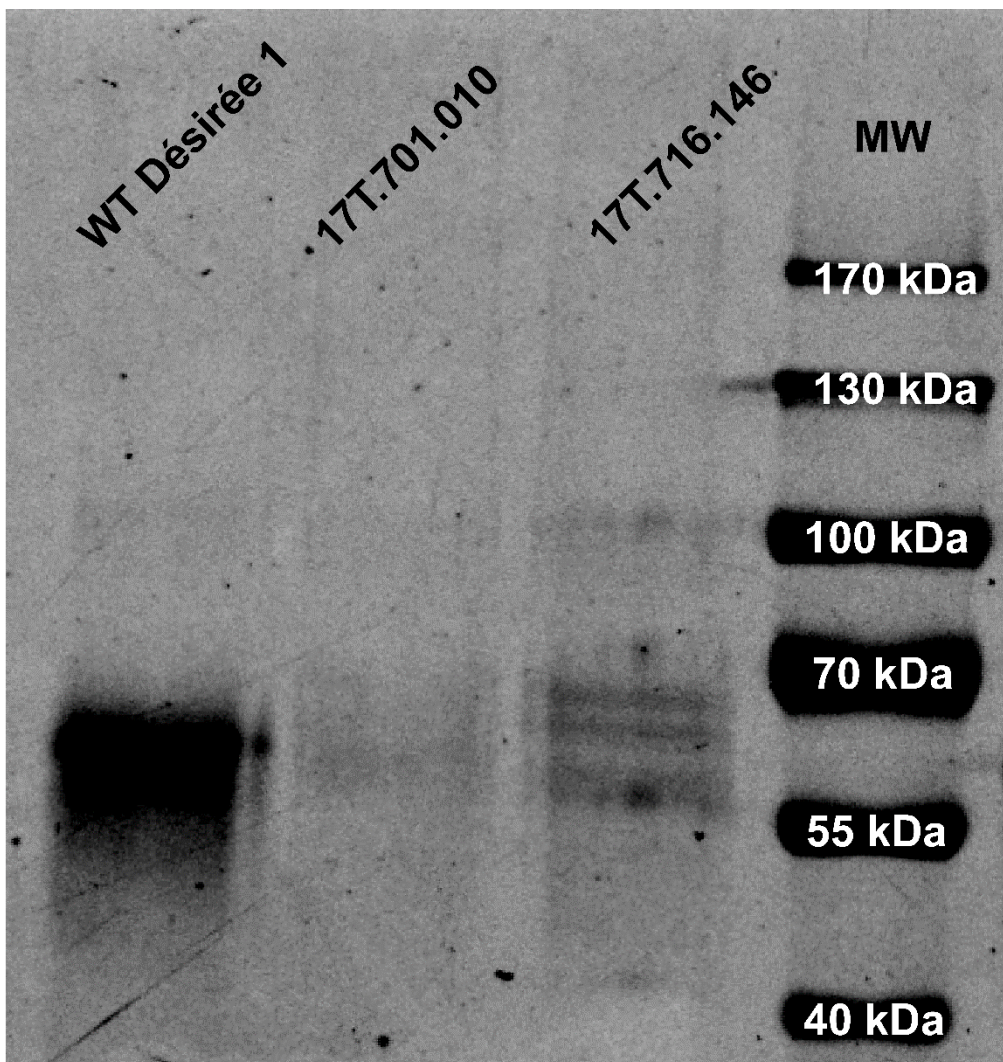
**Figure 51. Diagramme en boîte à moustache du taux d'incorporation basé sur la période de marquage pour l'ensemble des protéines identifiées dans le grain d'amidon.** Les taux d'incorporation ont été calculés à partir des intensités des pics obtenues avec le programme Thermo® Xcalibur, pour chacun des peptides dans l'ensemble des conditions. 2 plantes ont été analysées lors du marquage de la semaine 0 à 3, ainsi que lors du marquage des semaines 1,5 à 4,5. 4 plantes ont été analysées pour les semaines 3 à 6. Pour chaque plante, 3 peptides ont été analysés.

#### **IV. Impact d'une protéine sur l'ensemble du protéome, étude des mutants pour la protéine GBSS.**

Nous avons montré que la morphologie des grains d'amidon avait un impact sur le protéome de ceux-ci. De plus, de récentes études montrent que certaines protéines s'associant au grain forment des complexes (Tetlow et al., 2008). D'autres protéines permettent l'adressage au grain, c'est notamment le cas de PTST1 qui permet l'adressage de GBSS (Seung et al., 2015). GBSS est la protéine majoritaire au sein du grain d'amidon. Toutefois, la morphologie des grains d'amidon issus de mutants *gbss* (totalement dépourvus d'amylose) ne semble pas impactée. Nous avons donc cherché à savoir si le contenu de GBSS pouvait impacter les concentrations des autres protéines dans le grain.

## 1. Dosage des protéines dans l'amidon de pomme de terre des mutants *waxy*

Afin d'étudier l'impact de la modulation de GBSS sur l'ensemble du protéome, nous avons étudié deux lignées mutantes générées et fournies par l'IGEPP INRA de Ploudaniel (Veillet et al., 2019). La première lignée, 17T.716.146, est une lignée mutante obtenue par CRISPR-cas9 ne produisant pas d'amylose, et par conséquent dont l'activité de GBSS a totalement été inhibée. La deuxième lignée, 17T.701.010, est une lignée mutante dont l'amidon contient une quantité d'amylose très faible. Il a de plus été démontré par analyse des séquences que cette lignée exprime une GBSS intacte et active sur seulement un allèle parmi les 4 existants chez la pomme de terre (Veillet et al., 2019). La lignée sauvage cv. Désirée, correspondant au fond génétique des mutants, a été utilisée comme contrôle. Les gels réalisés sur les extraits de protéines, ainsi que les données obtenues dans l'étude précédente, montrent que GBSS n'est pas ou peu synthétisé dans ces mutants (figure 52.) (Veillet et al., 2019).



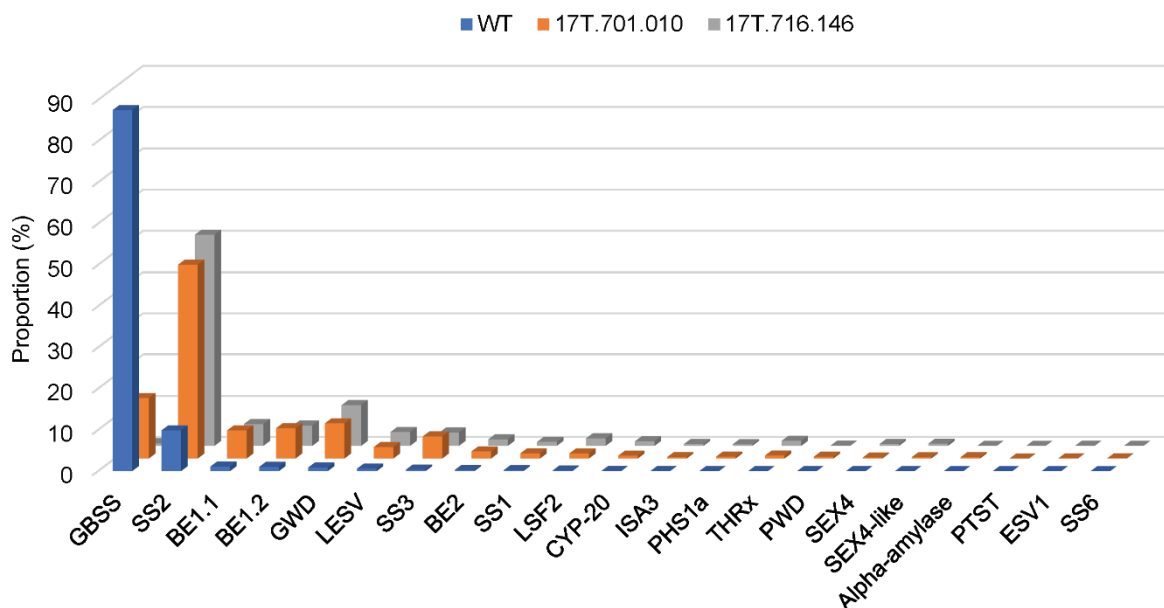
**Figure 52.** Gel SDS-PAGE des mutants *waxy* chez la pomme de terre. De gauche à droite : Puit 1 : WT, protéines de l'amidon de tubercule Désirée sauvage récolté au même moment que

les mutants 17T.701.010 et 17T.716.146. Puit 2 : Protéines provenant de l'amidon de tubercule du mutant 17T.701.010. Puits 3 : protéines provenant de l'amidon de tubercule du mutant 17T.716.146 Puit 4 : MW, échelle de poids moléculaire. Protéines

De plus, on remarque que d'autres bandes apparaissent sur le gel en l'absence de GBSS (Figure 52). Ces bandes supplémentaires sont plus intenses dans le mutant 17T.716.146 que dans le mutant 17T.701.010. L'intensité des bandes dans chaque cas semble montrer une plus forte concentration de protéines dans le sauvage que dans les mutants. Il semble toutefois y avoir plus de protéines dans le mutant total 17T.716.146 que dans le mutant 17T.701.010. Le dosage des protéines du grain est nécessaire pour avoir une idée précise de la quantité de protéine dans chacun des échantillons. L'analyse des bandes est actuellement en cours pour pouvoir les identifier. Toutefois, le protéome complet de l'amidon des mutants a d'ores et déjà été analysé.

## 2. Protéomique quantitative comparative des mutants *waxy*

Les extraits protéiques totaux des amidons mutants ont été analysés par protéomique quantitative. Ceci a révélé que la quantité de GBSS était fortement diminuée dans les deux lignées mutantes. Sa quantité est d'autant plus diminuée dans la lignée 17T.716.146, sans pour autant être totalement absente (Figure 53.). En effet, GBSS représente en proportion relative environ 87% des protéines de l'amidon WT, 14% des protéines dans le mutant 17T.701.010 et à peine 0,6% des protéines dans l'amidon du mutant 17T.716.146 (Figure 53).



**Figure 53. Histogramme des proportions relatives des protéines du grain d'amidon de pomme de terre Désirée sauvage et mutantes pour la protéine GBSS. La proportion de**

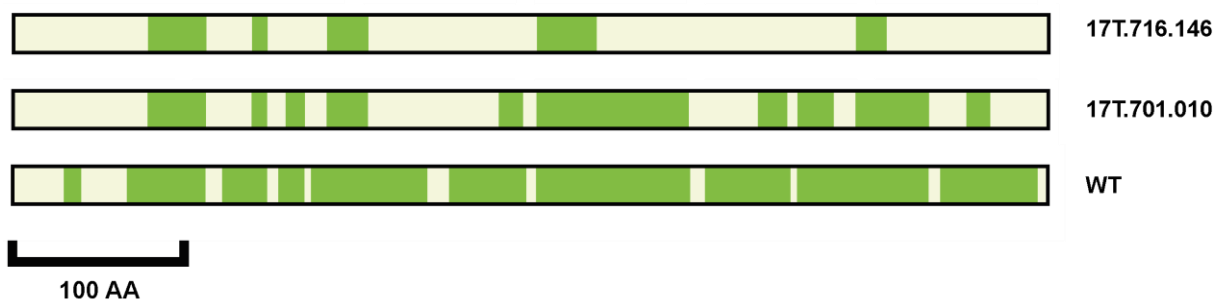
protéines a été estimé à partir du score iBAQ de chacune des protéines, ramené en pourcentage en divisant par la somme des scores iBAQ de l'ensemble des protéines pour chaque injection. WT : lignée sauvage du cultivar Désirée. 17T.701.010 : lignée mutante *gbss* pour 3 allèles du cultivar Désirée. 17T.716.146 : lignée mutante *gbss* pour les 4 allèles du cultivar Désirée.

Le nombre de peptide identifié pour GBSS reste dérisoire dans le mutant 17T.716.146, avec 15 « Peptide spectrum matches » (PSM) soit 100 fois moins que le Wild type (Tableau 9). Cette quantité est toutefois suffisante pour valider la présence de GBSS dans l'échantillon.

**Tableau 9. Couverture de séquence de GBSS et quantité de peptides trouvés dans les échantillons sauvages et mutants pour la protéine GBSS.** WT : lignée sauvage du cultivar Désirée. 17T.701.010 : lignée mutante *gbss* pour 3 allèles du cultivar Désirée. 17T.716.146 : lignée mutante *gbss* pour les 4 allèles du cultivar Désirée. Les peptides ont été identifié avec Thermo® Proteome Discoverer 2.4.0 en utilisant la banque de donnée maison créée à partir de l'identification des protéines du grain faite dans l'article de Helle et coll. 2018 (Annexe 4) (Helle et al., 2018).

échantillon	couverture (%)	PSMs	peptides uniques
WT	80	1539	50
17T.701.010	45	421	25
17T.716.146	20	15	9

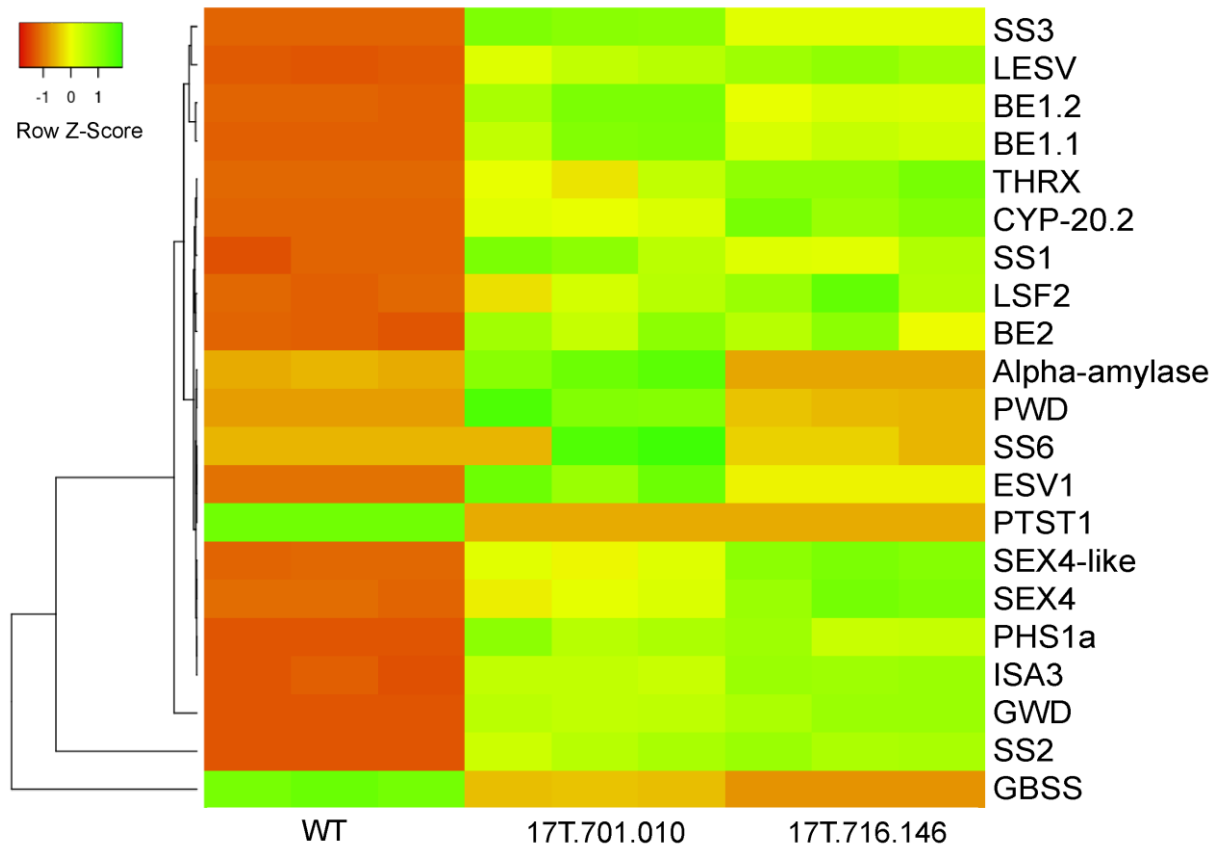
L'analyse de la couverture de séquence pour le mutant KO montre que les peptides identifiés sont dispersés sur l'ensemble de la séquence, toutefois, la couverture de séquence est trop faible pour pouvoir affirmer que la protéine GBSS identifiée n'est pas tronquée (Figure 54).



**Figure 54. Couverture de séquence de GBSS dans les différents échantillons.** Les rectangles verts représentent les séquences couvertes par les peptides identifiés avec Thermo® Proteome Discoverer 2.4.0 en utilisant la banque de donnée maison créée à partir de l'identification des protéines du grain faite dans l'article de Helle et coll. 2018 (Annexe 4) (Helle et al., 2018).

La protéine PTST1, qui est la protéine responsable de l'adressage de GBSS au grain chez Arabidopsis, est absente dans les analyses des mutants, ou en quantité tellement moindre qu'elle ne peut être détectée par spectrométrie de masse (quantité inférieure à l'attomole) (Annexe 6) (Figure 54). La proportion de l'ensemble des autres protéines est augmentée dans les mutants

par rapport au WT. On remarque toutefois que pour PWD, SS6, et ESV1 et l'Alpha-amylase, leurs proportions sont plus fortes dans le mutant partiel 17T.701.010 que dans le mutant 17T.716.146 et le WT (Figure 55).

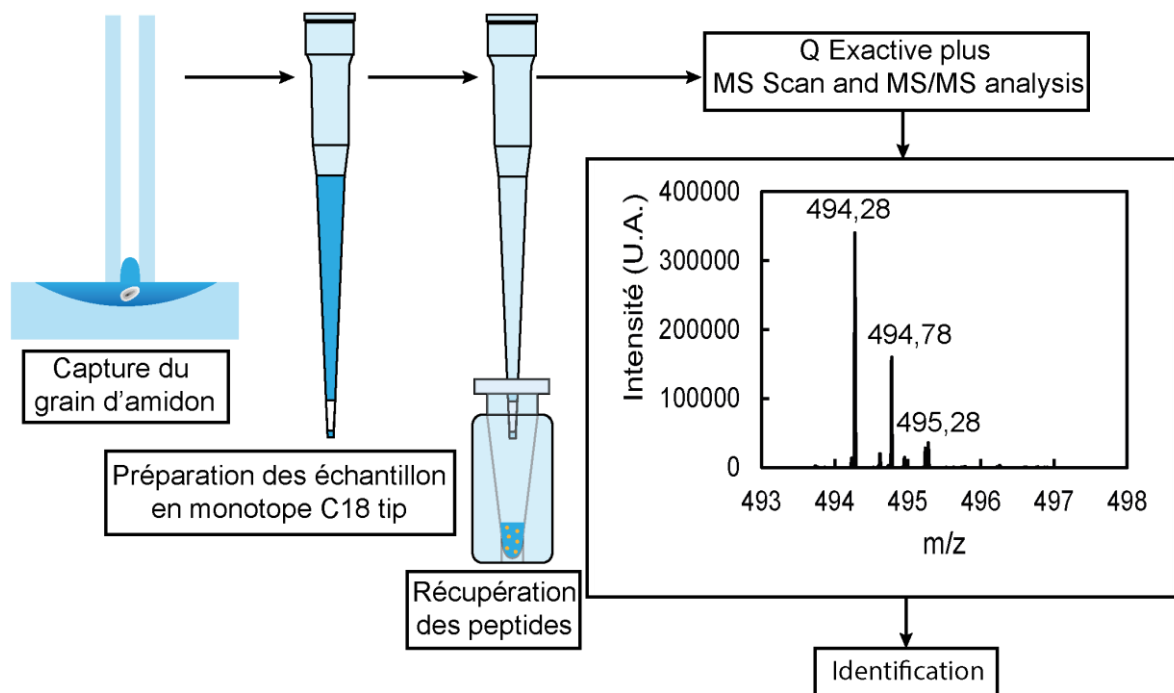


**Figure 55. Carte de chaleur de l'analyse protéomique comparative des mutants *waxy* et de la lignée sauvage *cv. Désirée*.** La proportion de protéines a été estimée à partir du score iBAQ de chacune des protéines, ramené en pourcentage en divisant par la somme des scores iBAQ de l'ensemble des protéines pour chaque injection. WT : lignée sauvage du cultivar Désirée. 17T.701.010 : lignée mutante *gbss* pour 3 allèles du cultivar Désirée. 17T.716.146 : lignée mutante *gbss* pour les 4 allèles du cultivar Désirée.

## V. Protéomique du grain unique d'amidon

Les analyses de la stœchiométrie des protéines liées au grain nous ont montré que celle-ci fluctue en fonction de la taille des grains. Toutefois, ces analyses ont été réalisées sur des fractions de grains possédant elles-mêmes une hétérogénéité morphologique. En effet, plusieurs tubercules de pomme de terre, composés de différents tissus, ont été utilisés pour extraire l'amidon en quantité suffisante. Par conséquent, les échantillons sont composés par des grains d'amidon d'origine cellulaire et de maturité variables qui pourraient conduire à des différences de composition ou de structure. Il est particulièrement intéressant de savoir si les protéomes de deux grains d'amidon de morphologie similaire peuvent présenter des différences. Comme

présenté lors de l'introduction, plusieurs mécanismes sont possibles pour l'initiation et la synthèse du grain, il est donc possible que plusieurs mécanismes puissent mener à un même grain. Pour répondre à cette question, j'ai développé une stratégie permettant d'étudier le protéome à l'échelle d'un grain. Pour cela, nous avons adapté la technique « in Stage Tip » (iST) de manière à pouvoir l'utiliser pour l'extraction et l'analyse des protéines du grain d'amidon (Figure 56) (Kulak et al., 2014).



**Figure 56.** Pipeline expérimental des analyses grain unique. Le grain d'amidon est capture dans une colonne de diamètre interne 100  $\mu\text{m}$  par capillarité. Ensuite, le grain d'amidon est élué dans un cône de pipette 100  $\mu\text{L}$  contenant un disque C18 apolaire. La gélatinisation du grain d'amidon, ainsi que la réduction, l'alkylation et la digestion sont réalisé à l'intérieur du cône. Les peptides ainsi obtenus sont dessalés avec une solution aqueuse, puis élué avec de l'Acétonitrile directement dans le tube d'analyse. L'Acétonitrile est évaporé puis remplacé par 2  $\mu\text{L}$  d'acide formique à 0,1% (v/v) avant injection en spectrométrie de masse. L'interrogation est réalisée sur la banque de donnée MASCOT greenplant.

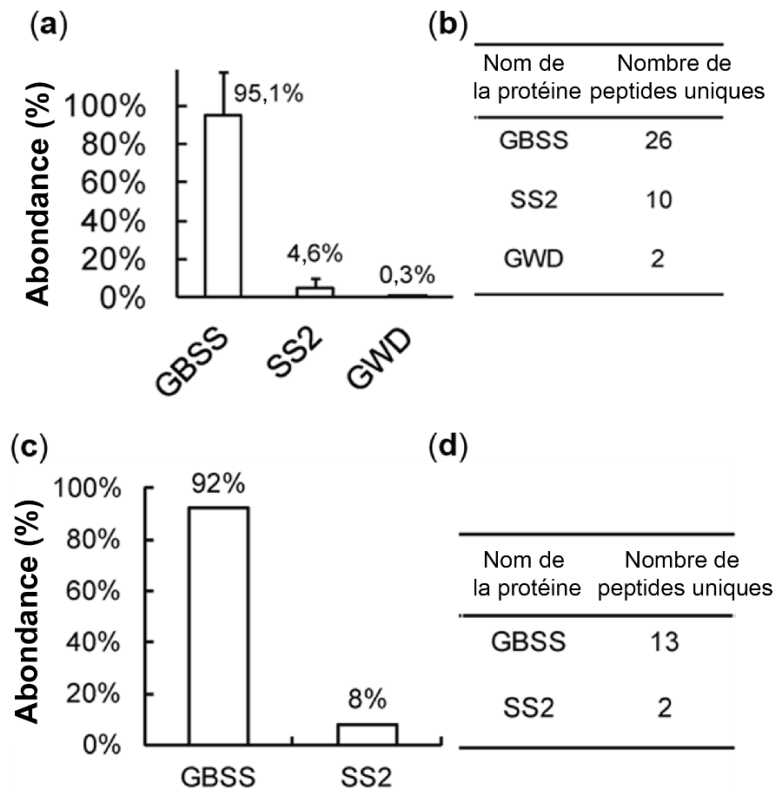
### 1. Analyse protéomique à partir de 30 ou d'un unique grain d'amidon

Les expériences présentées précédemment dans ce manuscrit ont toutes été réalisées à partir d'échantillons de 300 mg d'amidon de pomme de terre. Cette quantité d'amidon est largement suffisante pour pouvoir détecter l'ensemble des protéines du grain. Qui plus est, sur les protéines extraites de 300 mg, seulement environ  $\frac{1}{4}$  de l'extrait protéique total est digéré en peptide (50  $\mu\text{L}$  d'extrait de protéine pour 200  $\mu\text{L}$  en tout). Cela nous permet d'aboutir *in fine* à environ 20 mg de peptides après digestion. Les injection se faisant à 1  $\text{mg}\cdot\text{mL}^{-1}$  à partir de 10

$\mu\text{L}$ , seul 10 mg sont utilisé pour une analyse avec 3 réplicats techniques, une analyse requérant 2  $\mu\text{L}$ . La quantité réelle d'amidon utilisé pour une analyse en spectrométrie de masse est donc de 7,5 mg d'amidon sans réplicats techniques. Un grain d'amidon de 100  $\mu\text{m}$  représente en termes de masse environ 0,75  $\mu\text{g}$ , soit un facteur  $10^{-4}$ . Les calculs théoriques réalisés à partir du volume des grains, de la concentration en protéines dans le grain et de leur stœchiométrie indiquent que les protéines majoritaires peuvent être détectées par spectrométrie de masse à partir d'un unique grain d'un diamètre de 100  $\mu\text{m}$ . Selon ces mêmes calculs, l'analyse d'un grain d'amidon de 10  $\mu\text{m}$  de diamètre permettrait de détecter GBSS et SS2. Nous avons d'abord analysé le protéome extrait de 30 grains de tailles différentes puis celui d'un grain d'amidon d'environ 70  $\mu\text{m}$  de diamètre. Les analyses ont été réalisées par la méthode in stage tip (iST) de manière à réduire les étapes de transfert par rapport à l'eFASP, et par conséquent la perte de matériel durant les étapes de préparation (Kulak et al., 2014). Les grains d'amidon sont dénombrés dans une goutte de 20  $\mu\text{L}$  et sont capturés à l'intérieur d'un capillaire de silice de 100  $\mu\text{m}$  de diamètre interne. Les grains sont ensuite élués à l'aide d'une seringue dans un tube par 50  $\mu\text{L}$  de tampon de dénaturation qui va à la fois dénaturer, réduire et alkyler les protéines. Cette solution est ensuite diluée pour que le milieu soit adapté à l'action de la trypsine. Enfin, les peptides sont purifiés à l'aide d'un cône de pipette contenant un disque C18, puis analysés en spectrométrie de masse.

Ces analyses ont permis d'identifier les 3 protéines majoritaires : GBSS, SS2 et GWD) à partir de 30 grains et les 2 protéines majoritaires (GBSS et SS2) à partir d'un grain unique (Figure 57). De plus, 13 peptides uniques de GBSS ont pu être identifiés à partir d'un grain unique, ce qui correspond à une couverture de séquence de 30%.





**FIGURE 57.** Analyses du protéome de l'amidon de pomme de terre à partir de 30 grains et d'un unique grain. (a) Abondances relatives des protéines identifiées à partir de 30 grains. (b) nombre de peptides uniques identifiés à partir de 30 grains. (c) Abondance relative des protéines identifiées avec un seul grain. (d) Nombre de peptides uniques par protéines identifiées à partir d'un seul grain.

Comme décrit dans l'introduction ; les calculs théoriques indiquent la possibilité de détecter les 5 protéines majoritaires. D'autres analyses ont été lancées sur des grains d'amidon de morphologie similaire à celui analysé dans la figure 48, c'est à dire avec un diamètre de 70  $\mu\text{m}$ . Ces expériences n'ont mené qu'à l'identification de GBSS, ne permettant pas de comparer la stœchiométrie des protéines pour une taille similaire. Toutefois, l'identification de GBSS avec un seul grain peut être utile à d'autres fins.

## 2. Identification d'espèces à partir de la séquence de GBSS.

La couverture de séquence de GBSS à partir d'un grain unique serait suffisante pour permettre de déterminer si la séquence de GBSS provient d'une espèce ou d'une autre. Pour s'assurer de cela, une digestion *in-silico* de la protéine GBSS a été réalisée sur plusieurs espèces végétales (pomme de terre, riz, maïs, pois, orge, blé et patate douce). Cela a permis d'établir la liste des peptides uniques de GBSS de chaque espèce. La liste des peptides de GBSS révèle que

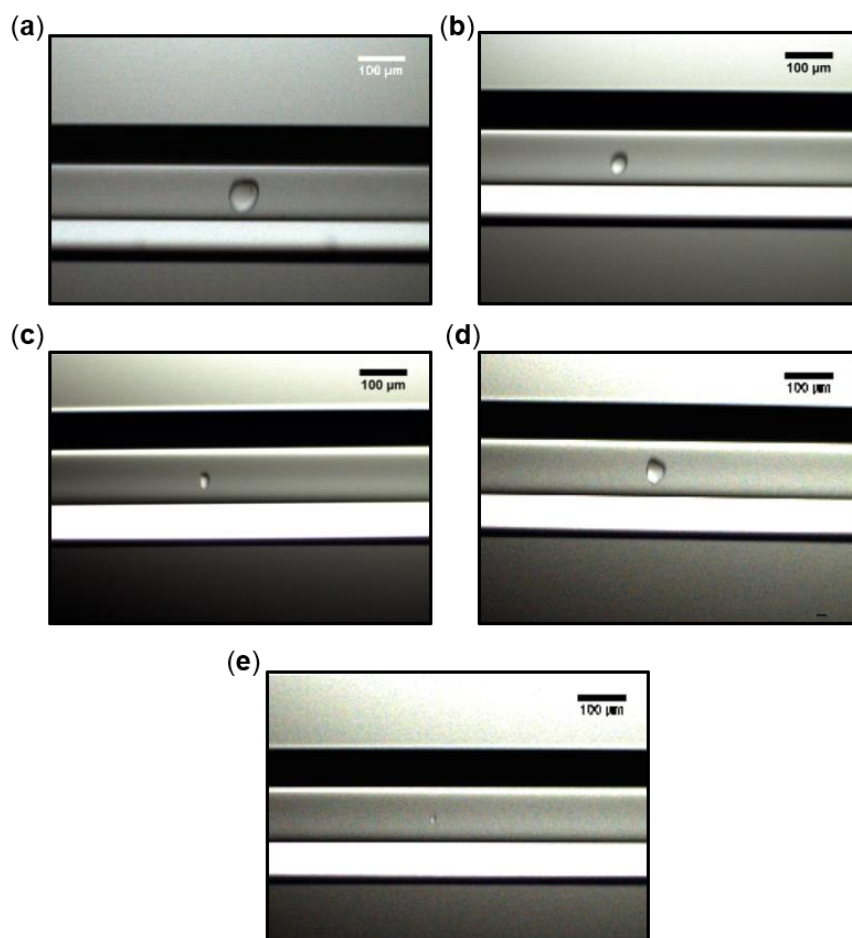
77 % des peptides sont spécifique de la pomme de terre parmi les 7 espèces interrogées (Tableau 10).

**Tableau 10. Liste des peptides de GBSS pouvant être obtenu de pomme de terre après digestion à la trypsine.** La digestion in silico a été réalisé à partir du site ExPASy. Les clivages manqués n'ont pas été pris en compte. Les cellules grisées représentent les peptides qui sont commun à au moins une des 6 autres espèces (riz, maïs, pois, orge, blé et patate douce).

Masse	Position	Séquence du peptide
3973,9469	215-249	VLNLNSSNYFSGPYGEDVLFANDWHTALIPCYLK
2723,4137	581-607	WETLLLGLGASGSEPGVEGEEIAPLAK
2713,2272	519-543	EGYTGFMGAFNVECDVDPADV LK
2294,0427	350-370	TCITGIVNGMDTQEWNPATDK
2287,1725	467-487	FNVPLAHMITAGADFMLVPSR
2007,9909	445-460	EFEQEIEQLEVLYPNK
1968,0324	322-339	VVTVSPYYAQELVSAVDK
1943,9537	276-291	FSFSDFLLNLPDEF R
1780,9149	501-518	YGTVPICASTGGLVDTV K
1734,8883	86-101	GMNLIFVGTEVGPWSK
1634,9224	102-119	TGGLGDVLGGLPPAL AAR
1607,8713	376-389	YDITTVMDAKPLL K
1601,7845	292-305	GSFDFIDGYEKPV K
1597,8658	551-565	ALAVYGTLAFAEMIK
1558,9202	429-443	FIGLDVQIVVLGTG K
1514,7606	488-500	FEPCGLIQLHAMR
1491,7525	263-275	VAFCIHNIA YQGR
1390,7399	202-214	FSLLCQAAL EAPK
1361,6922	167-177	VFVDHPMFLEK
1319,6477	135-146	DAWDTSVAVEV K
1310,7314	390-402	EALQAAVGLPV D K
1278,6324	191-201	AGLDYLDNEL R
1226,518	566-575	NCMSEELSW K
1191,5972	74-85	RPGCSATIVCG K
1162,6076	37-46	NHTLTHNGL R
1123,6469	418-428	GSDILVAAIHK
987,5469	147-155	VGDSIEIV R
987,5469	340-348	GVELDSVLR
985,6193	404-412	IPLIGFIGR
974,5629	28-36	STLSQIGLR
882,4792	314-321	AGILESHR
879,4418	20-27	SQTS LDTK
844,381	156-161	FFHCYK
803,4444	123-129	VMTISPR
788,4261	51-57	LDGLQSR
778,4457	256-262	GIY LNAK
771,3454	250-255	SMYQSR
759,4723	544-550	IVTTVAR

716,325	130-134	YDQYK
691,3596	309-313	INWMK
646,3406	413-417	LEEQK
630,3093	608-613	ENVATP
625,3191	371-375	YTDVK
577,3344	186-190	IYGPK

La capture de grains uniques d'amidon venant de différentes espèce a ensuite été réalisée. En plus d'un grain de pomme de terre ont été analysé : un grain de patate douce, un grain de blé, un grain de maïs et un grain de pois. (Figure 58.) Le protéome de chaque grain a été réalisé mais n'a malheureusement pas permis de discriminer l'espèce à laquelle chaque grain appartenait, soit car aucun peptide n'a pu être identifié, soit le peptide de GBSS identifié ne correspond pas à l'espèce attendu. Des développements sont encore nécessaires pour parvenir à notre objectif. Le but final étant d'être capable d'identifier l'espèce végétal à laquelle appartient un grain d'amidon retrouvé sur des échantillons archéologiques.



**FIGURE 58. Grains uniques de différentes espèces capturés en capillaire.** (a) Grain d'amidon de pomme de terre. (b) Grain d'amidon de patate douce. (c) grain d'amidon de blé (type A) (d) Grain d'amidon de pois (e) Grain d'amidon de maïs.

# Discussions et Perspectives

## I. Établissement du protéome complet de l'amidon de pomme de terre.

Lors de cette étude, nous avons extrait l'amidon de pommes de terre cv. Monalisa, puis extrait les protéines présente dans les grains d'amidon. Par SDS-PAGE, suivi d'une identification des bandes par spectrométrie de masse, nous avons identifié 36 protéines dont 20 sont connues pour agir de manière directe dans le métabolisme de l'amidon. Ceci constitue le protéome du grain d'amidon le plus complet à ce jour. Ce projet a pu être réalisé grâce au séquençage préalable du génome de *Solanum phureja* et aux analyses génomiques réalisées chez la pomme de terre (Xu et al., 2011; Van Harsselaar et al., 2017). La quantification sans marquage basé sur l'intensité des pics des protéines extraites et analysées par méthode shotgun bottom-up ont permis d'établir la stœchiométrie des protéines au sein de l'amidon de tubercule de pomme de terre. Enfin, différents traitements préalables à la thermolysine et au SDS ont été réalisé pour pouvoir distinguer les protéines en interactions faibles et en interaction forte avec le grain.

Nos données ont confirmé la présence de GBSS, SS2 et GWD, qui avaient déjà été précédemment identifiées en association avec le grain d'amidon de pomme de terre par SDS-PAGE et immunoblot, suivi d'une identification des bandes par spectrométrie de masse (Stensballe et al., 2008). Nous avons en outre identifié SS1, SS3, BE1, BE2 et PHS1a, qui avaient déjà été observées chez d'autres espèces (Grimaud et al., 2008). La présence de ces protéines dans le grain indique que les complexes entre les enzymes de branchement et les amidon-synthases, associés au grain d'amidon et déjà observés chez le blé, le maïs et l'orge, pourraient également exister chez la pomme de terre (Tetlow et al., 2008; Liu et al., 2009; Ahmed et al., 2015). Les limitations techniques propres à l'extraction protéique, qui implique la dénaturation des protéines, ne permettent pas d'identifier avec certitude ces complexes. L'utilisation d'agents « cross-linker », capables de former des liaisons covalentes entre des protéines proches, avant la gélatinisation des grains et la dénaturation des protéines pourraient permettre d'identifier les protéines en interaction dans le grain. Le formaldéhyde semble être un agent de choix pour cela. Celui-ci est en effet assez petit pour pouvoir pénétrer dans le grain et créer des liaisons covalentes entre deux protéines en interaction. L'analyse du phosphoprotéome de l'amidon de pomme de terre devrait par la suite nous permettre d'identifier les protéines phosphorylées au sein du grain, puisque la plupart des protéines organisées en complexe dans le grain d'amidon sont phosphorylées (Tetlow et al., 2004; Ahmed et al., 2015). Mis à part GBSS, il n'est pas encore clair si, *in planta*, les enzymes sont actives uniquement

sous leurs formes solubles ou quand elles sont physiquement liées au grain d'amidon, ou encore les deux.

Les travaux réalisés sur les protéines d'adressage chez *Arabidopsis* ont montré que GBSS est activement adressée au grain d'amidon par PTST1 grâce aux interactions qui s'établissent entre ces deux protéines notamment par les domaines coiled-coil (Seung et al., 2015). Nous avons observé PTST1 en association avec le grain, confirmant son interaction avec les polysaccharides de l'amidon chez la pomme de terre. La question des interactions de PTST1 avec GBSS et/ou d'autres protéines chez la pomme de terre restent cependant en suspens et nécessiterait des travaux complémentaires. D'autres protéines possèdent aussi des domaines faisceau d'hélices (ou coiled coil) pouvant mener à des interactions avec PTST1. Chez *Arabidopsis*, il a récemment été démontré que l'isoforme SS4 interagit avec PTST2, un homologue de PTST1 et un orthologue de la protéine FLO6 chez le riz, ou BGC1 chez le blé (Seung et al., 2017; Chia et al., 2019a; Chia et al., 2019b). La fonction supposée de PTST2 est d'amener un malto-oligosaccharide primordial à SS4 pour l'initiation de la synthèse du grain, en association avec une autre protéine d'adressage, PTST3 (Seung et al., 2017). Dans notre étude, Quelques peptides de SS4 ont été observés, mais aucun peptide correspondant à un orthologue de PTST2 ou PTST3 n'a pu être détecté. Toutefois l'interaction entre PTST2 et SS4 est transitoire et pourrait se dérouler uniquement lors de l'étape d'initiation de la synthèse du grain, rendant peu probable la séquestration de ces protéines dans l'amidon. Par ailleurs, PTST2 est impliquée dans le contrôle de l'initiation des grains d'amidon de type B chez le blé (Chia et al., 2019b). Ceci suggère que le complexe PTST2/SS4 pourrait également intervenir dans le métabolisme de l'amidon de réserve (les travaux sur PTST1, 2 et 3 ont été réalisés sur les feuilles d'*Arabidopsis* qui accumulent de l'amidon transitoire). Le rôle de ce complexe reste à étudier chez le tubercule de pomme de terre où aucune étude, aussi bien sur le rôle de SS4 que sur le rôle d'orthologues à PTST2, n'a été réalisée.

Dans ce travail, nous avons également observé la présence d'enzymes de dégradation de l'amidon dans le grain. Celles-ci comprennent les dikinases GWD et PWD, les phosphatases SEX4 et LSF2 ainsi que l'enzyme de débranchement ISA3. Qui plus est, les protéines LESV et ESV1, des régulateurs de la dégradation d'amidon transitoire chez *Arabidopsis*, ont aussi été identifiées (Feike et al., 2016). L'interaction de GWD et PWD avec le grain a déjà été démontrée à plusieurs reprises (Yu et al., 2001a; Baunsgaard et al., 2005; Kötting et al., 2005; Stensballe et al., 2008). De plus, nos travaux ont confirmé l'association de SEX4 et LSF2 avec le grain d'amidon de pomme de terre (Santelia et al., 2011). Nous avons pour la première fois

observé ISA3 en association avec le grain d'amidon. Tout comme ISA1, ISA3 hydrolyse les liaisons  $\alpha$ -1,6 de l'amylopectine. Toutefois, ISA3 est impliquée dans la dégradation et non dans la synthèse de l'amylopectine chez le riz et chez *Arabidopsis* (Wattebled et al., 2005; Yun et al., 2011). La fonction de ISA3 dans le métabolisme de l'amidon de réserve n'est pas encore totalement connue. L'inhibition de l'ensemble des isoamylases connues chez la pomme de terre n'altère pas le métabolisme de l'amidon transitoire (Ferreira et al., 2017). L'accumulation d'amidon dans les tissus de réserves est toutefois fortement réduite (Ferreira et al., 2017). La morphologie des grains est aussi altérée avec une augmentation de la proportion des petits grains, ce qui avait déjà été observé dans des lignées antisens réduisant l'expression de ISA1 et 2 (Bustos et al., 2004; Ferreira et al., 2017). De ce fait, si la fonction du complexe ISA1/ISA2 dans la synthèse du grain d'amidon est bien comprise, la fonction de ISA3 dans le métabolisme de l'amidon de réserve nécessite des études complémentaires, comme la caractérisation de lignées mutantes knock-out (Hussain et al., 2003). Le fait que la quantité de ISA3 ne diminue pas lors des traitements au SDS 2% et à la thermolysine démontre une interaction forte de ISA3 avec le grain d'amidon. La présence de ISA3 au sein des grains pourrait s'expliquer par une interaction avec une protéine d'adressage comme PTST2/FLO6 qui, chez le riz, interagit avec ISA1 (Peng et al., 2014). L'absence de domaine coiled-coil chez ISA3 de pomme de terre ne va toutefois pas dans ce sens et le ciblage d'ISA3 au grain d'amidon pourrait dépendre d'un autre processus.

La présence de PHS2 au sein du grain est surprenante car cette isoforme est historiquement décrite comme cytosolique (Fettke et al., 2004). PHS2 a toutefois été identifiée dans le chloroplaste d'*Arabidopsis* par protéomique, il est donc possible que cette enzyme soit présente à la fois dans le plaste et dans le cytosol (Helm et al., 2014). On retrouve deux isoformes prédites de PHS2 (PHS2a et PHS2b) dans le génome de la pomme de terre (Van Harsselaar et al., 2017). L'observation de PHS2a dans le grain semble montrer que cette isoforme n'est pas seulement présente dans le cytosol. Cependant, PHS2a n'a pas été retrouvée en quantité suffisante pour permettre sa quantification dans notre étude (Helle et al., 2018; Helle et al., 2019). Il est donc possible que la présence de PHS2 soit due à son interaction avec l'amidon et qu'elle ait été co-purifiée avec les grains au cours des étapes d'extraction. Dans ce cas, il serait attendu que ces protéines soient absente ou en quantité encore plus limitée lors des traitements de surface au SDS2% ou à la thermolysine.

La découverte de SS5 et SS6 dans cette étude est l'une des plus intéressantes. Ces deux isoformes avaient été annotées dans le génome de plusieurs espèces, mais elles n'avaient pas

encore été caractérisées (Deschamps et al., 2008; Yu et al., 2012; Liu et al., 2015; Van Harsselaar et al., 2017). L'expression de *SS5* et *SS6* avait été observée lors d'analyses transcriptomiques chez le maïs et la pomme de terre (Yu et al., 2012; Liu et al., 2015). Notre étude montre pour la première fois leur interaction avec le grain d'amidon, suggérant fortement qu'elles soient fonctionnelles et impliquées dans le métabolisme de l'amidon. Nos analyses phylogénétiques révèlent que les gènes correspondants ont une origine monophylétique et que ces isoformes font partie du groupe B des amidon-synthases, auquel appartiennent également *SS3* et *SS4* (Leterrier et al., 2008). Cette découverte a ouvert la voie à la caractérisation fonctionnelle des gènes de *SS5* et de *SS6*, ce qui pourrait déboucher sur la découverte de nouvelles fonctions dans la synthèse de l'amidon. *SS5* a récemment été caractérisée chez *Arabidopsis*, et la caractérisation de *SS6* est en cours chez le manioc et la pomme de terre (Abt et al., 2020). La ressemblance des séquences avec celles de *SS4*, ainsi que les récents travaux sur *SS5* d'*Arabidopsis*, tendent à montrer que ces protéines joueraient aussi un rôle dans l'initiation de la synthèse du grain d'amidon.

Notre étude a révélé la présence d'inhibiteurs de protéases appartenant à la famille des PKPI (Potato Kunitz-type protease inhibitors) ou PINII (proteinase inhibitor type II). Les PKPI sont divisés en 3 groupes : A, B et C. Dans le groupe A, on retrouve des inhibiteurs de protéases à aspartate et des chymotrypsines alors que le groupe B inclut les inhibiteurs de protéases à sérine, des trypsines ainsi que des chemotrypsines (Heibges et al., 2003). Le groupe C comprend des inhibiteurs de protéases à cystéines ainsi que des inhibiteurs d'invertases (Heibges et al., 2003; Poltronieri et al., 2012). Il existe toutefois un fort polymorphisme fonctionnel entre les PKPIs (Heibges et al., 2003). Les inhibiteurs de protéases de type II sont des inhibiteurs de protéases à sérine contenant 8 cystéines formant des ponts disulfures et dont l'activité serait principalement liée à la défense contre les pathogènes (Turra and Lorito, 2011). Contrairement aux autres protéines, les inhibiteurs de protéases voient leur quantité diminuer lors des traitements au SDS 2% ou à la thermolysine. Ceci suggère une interaction plus faible de ces protéines avec le grain. Ces protéines peuvent soit être présentes spécifiquement en surface du grain et protéger les enzymes de la protéolyse pendant la synthèse du grain, ou avoir été co-purifiées avec le grain d'amidon lors des étapes d'extraction. La nature de l'interaction entre ces protéines et le grain d'amidon reste à caractériser pour étudier leurs fonctions potentielles dans le métabolisme de l'amidon.

Contrairement aux inhibiteurs de protéases, l'ensemble des protéines du métabolisme de l'amidon reste attaché au grain d'amidon après traitement au SDS 2% ou à la thermolysine. Ces



résultats indiquent que ces protéines sont séquestrées dans la matrice du grain et ne sont pas seulement présentes en surface. Ceci est aussi le cas pour la glutathion-péroxydase GPx et la thiorédoxine TRx, qui est un homologue de la thiorédoxine M d'Arabidopsis. Cette dernière est impliquée dans les mécanismes de régulation redox (Buchanan, 1984). La thiorédoxine M4 participe notamment à la régulation redox du métabolisme de l'amidon chez Arabidopsis en activant SS1 et BAM1 (Valerio et al., 2011; Skryhan et al., 2015). Les glutathion-péroxydases agissent quant à elles dans le maintien de l'homéostasie redox (contrôle de la concentration en espèces réactives de l'oxygène (ROS) au sein de la cellule) lors de la réponse au stress (Noctor et al., 2012). Ainsi, ces deux protéines, TRx et GPx, pourraient être impliquées dans la régulation redox de l'activité des enzymes du métabolisme de l'amidon, en agissant au niveau des ponts disulfures ou en facilitant la phosphorylation ou la déphosphorylation des enzymes. Leur présence au sein du grain suggère que l'homéostasie redox joue un rôle clé dans la régulation du métabolisme de l'amidon et ouvre de nouvelles perspectives pour étudier ces phénomènes chez des plantes les plus cultivées.

L'ensemble de ces protéines mise en évidence lors des expériences de protéomiques ouvre les champs de recherches sur de nouvelles fonctions dans la synthèse du grain d'amidon, ce qui pourrait déboucher sur la production d'amidons aux propriétés intéressantes. Ces protéines peuvent par extension servir de marqueurs de sélection potentiels pour une propriété physique donnée du grain d'amidon. L'étude comparative du contenu en protéine en fonction de la morphologie des grains a permis d'apporter plus de réponse sur ces questionnements.

## **II. Fluctuation de la stoechiométrie basée sur la morphologie des grains.**

Après avoir établi le protéome complet de l'amidon de pomme de terre et déterminé la stoechiométrie des protéines, nous nous sommes penchés sur la stabilité de ce protéome en fonction de la morphologie des grains. Pour cela, nous avons fractionné les grains en fonction de leur taille et, pour chaque fraction, étudié leurs caractéristiques structurales par granulomorphométrie et microscopie électronique à balayage pour la taille et la forme des grains, diffraction aux rayons-X pour leur cristallinité, et par électrophorèse capillaire assisté par fluorescence pour leur contenu en phosphate, ainsi que la distribution des longueurs de chaînes de l'amylopectine. De plus, nous avons comparé le contenu en protéine de chaque fraction par spectrométrie de masse. Bien que les analyses sur l'hétérogénéité du contenu en protéines des grains d'amidon soit nombreuses, elles concernent pour la plupart des études

comparatives entre amidons d'origines botaniques ou tissulaires différentes (Jane et al., 1994). A contrario, l'hétérogénéité au sein d'un même échantillon a été peu étudiée. Une précédente étude sur l'amidon de pomme de terre a montré que les plus petits grains d'amidon arborent une morphologie facettée avec un type cristallin A, similaire aux grains d'amidon de riz, alors que les gros grains ont une forme plus ovoïde et un type cristallin B (Wang et al., 2016). Lors de notre étude, nous n'avons pas observé de différence de cristallinité entre les petits et les gros grains d'amidon. Quelle que soit la fraction observée, les profils de diffraction des rayons X étaient toujours du type cristallin B. De plus, les petits grains d'amidon étaient significativement plus ronds que les grains d'amidon de plus grande taille. Ceci est confirmé par l'analyse par microscopie électronique à balayage des différentes fractions qui révèle des grains lisses et sphériques. Cette différence entre notre étude et l'étude précédente peut s'expliquer par l'utilisation de cultivars différents, ou par l'origine de l'échantillon étudié (origine commerciale pour Wang et coll., préparation à partir de tubercules frais dans notre cas). Cependant, l'origine botanique de l'amidon utilisé dans l'article de Wang et collaborateurs n'était pas indiquée (Wang et al., 2016). Nous avons toutefois observé des résultats similaires pour les variations de distribution de longueurs de chaînes de l'amylopectine en fonction du diamètre du grain. En effet, la proportion relative des glucanes de DP 10 à DP 24 diminue lorsque la taille du grain augmente et cela s'accompagne d'une augmentation des chaînes les plus longues allant de DP 45 à DP 100. Cette différence peut s'expliquer par plusieurs paramètres. Tout d'abord, le métabolisme de l'amidon peut différer selon le type cellulaire, conduisant à des différences de structure des molécules d'amylopectine. On sait par exemple que chez *Arabidopsis* il existe une forte disparité entre le métabolisme de l'amidon des cellules de garde et celui des cellules du mésophylle (Santelia and Lunn, 2017). Une autre explication possible peut résider dans une régulation différentielle des enzymes de la biosynthèse de l'amidon. On peut supposer que la taille des grains d'amidon observée au moment de l'échantillonnage est relative à leur état de maturité et à l'âge du tissu dont ils proviennent. Des modulations d'expression des enzymes de biosynthèse pourraient exister en fonction du tissu considéré et au cours de la tubérisation. Ceci pourrait conduire à des différences structurales de l'amylopectine à l'instant de l'échantillonnage en fonction de la taille des grains. Bien que la dynamique de synthèse et de dégradation des grains d'amidon ait été étudiée dans les feuilles d'*Arabidopsis*, elle reste à établir pour les organes de réserve tels que le tubercule de pomme de terre (Crumpton-Taylor et al., 2012). Nous avons par ailleurs montré que les contenus en phosphate (aussi bien en position C3 qu'en position C6) diminuent quand le diamètre moyen des grains augmente. Ceci

corrobore les conclusions d'une précédente étude relatant que les petits grains d'amidon sont plus riches en phosphate que les plus gros (Jane and Shen, 1993).

Nous avons déterminé et quantifié le protéome des différentes populations de grains d'amidon. Grâce aux précédentes analyses ayant permis d'établir le protéome du grain d'amidon de pomme de terre, nous avons pu utiliser la quantification sans-marquage pour évaluer les changements de concentrations des protéines (Helle et al., 2018). Bien qu'il ne soit pas encore clair si les protéines du grain sont actives lorsqu'elles sont séquestrées dans le grain, ce type d'analyse donne une image de la quantité de chaque protéine au moment de la récolte. Notre étude a montré que la stœchiométrie des protéines au sein d'un grain d'amidon de pomme de terre changeait de manière significative entre les différentes fractions. Prise individuellement, la concentration d'une protéine peut varier de moins de 0,2 fois à plus de 4 fois en fonction de la fraction observée. Les concentrations les plus stables sont celles de GBSS, l'amidon synthase responsable de la synthèse de l'amylose, et de la thioredoxine THR<sub>x</sub>, dont le rôle au sein du grain d'amidon reste encore à démontrer. A contrario, les concentrations qui varient le plus avec la taille des grains sont celles de PTST1, la protéine d'adressage de GBSS au grain d'amidon, et de SS3, une des amidons synthases responsable de la synthèse de l'amylopectine. L'analyse des profils d'évolution des concentrations de chaque protéine en fonction du diamètre moyen des grains dans chaque fraction permet d'établir quatre groupes au sein desquels les protéines se comportent de façon semblable. Un premier groupe de protéines, comprenant SS4, SS6, BE1.1, BE1.2, ISA3, GP<sub>x</sub> et CYP20-2 montre une concentration qui décroît puis augmente avec l'augmentation du diamètre. SS1, 2 et 3, BE2, PTST1, LSF2, ESV1, LESV, PWD et PHS1a et b ont un profil opposé. Ces observations pourraient illustrer la présence de complexes protéiques au sein du grain tels que ceux décrits chez le maïs, le blé, l'orge et le riz (Tetlow et al., 2008; Liu et al., 2009; Ahmed et al., 2015; Hayashi et al., 2018). Ceci est particulièrement marquant pour SS1 et BE2 qui sont positivement corrélées entre elles, avec une quantité de BE2 double par rapport à celle de SS1, rappelant le complexe SS1-BE2-BE2 déjà observé chez plusieurs espèces (Tetlow et al., 2008; Liu et al., 2009; Ahmed et al., 2015). Cela peut indiquer que SS1 n'est présente dans le grain de pomme de terre que sous la forme d'un complexe avec deux BE2. Le complexe SS2-BE2-BE2 serait en revanche transitoire pour l'adressage de SS2 au grain puisque celle-ci est présente en bien plus grande quantité que BE2 dans le grain d'amidon, ou alors ce complexe n'est pas présent chez la pomme de terre.

Bien que la concentration en GWD dans l'amidon varie peu comparativement aux autres protéines, elle diminue significativement et de manière constante avec l'augmentation de la

taille des grains. Comme attendu, une corrélation positive forte existe entre le contenu en phosphate et la concentration de GWD. Cela peut amener à penser qu'il existerait un lien entre le contenu en phosphate, directement orchestré par GWD, et la morphologie des grains. Toutefois, les analyses de corrélation ne permettent pas de conclure sur le lien de cause à effet entre deux paramètres corrélés. Il est important de prendre en compte cela dans le cas du métabolisme de l'amidon, où de multiples relations existent entre les protéines elles-mêmes ou entre ces protéines et la structure du grain. Si l'on reprend l'exemple de GWD, sa forte corrélation avec la morphologie du grain peut aussi bien être le résultat d'un impact de la phosphoesterification de l'amidon sur la morphologie du grain, qu'une expression différentielle dans le temps de GWD pendant la croissance du grain, voire les deux.

Des études transcriptomiques en cellule unique seraient nécessaires pour étudier le niveau d'expression des gènes dans différentes cellules d'un même tissu à des stades de croissance différents. Nous avons pu remarquer que les cellules proches des tissus vasculaires du tubercule de pomme de terre, plus jeunes que les cellules en périphérie, accumulent des grains d'amidon de plus petite taille. Une autre expérience nécessaire pour mieux comprendre quand s'associent les protéines au grain consiste à marquer les protéines avec des isotopes stables par incorporation et chasse de nutriments enrichis en  $^{15}\text{N}$ .

### **III. Dynamique d'association des protéines au grain d'amidon**

Les analyses par marquage au  $^{15}\text{N}$  apportent un premier élément de réponse sur la façon dont les protéines sont associées au grain d'amidon. La nutrition des plantes avec une source d'azote enrichie en  $^{15}\text{N}$  permet, lors de la traduction de l'ARNm en protéine, d'insérer dans la séquence de la protéine des unités d'azote lourd. Cette incorporation, chez les êtres vivants complexes, ne se fait toutefois pas de manière instantanée du fait de l'existence de tissus de réserve contenant des acides aminés non marqués qui seront consommés en même temps que les acides aminés nouvellement synthétisés (Fan et al., 2016). Dans le cas des marquages en début et en milieu de tubérisation, l'évolution des massifs de pics illustre un effet de dilution, avec un massif étiré vers le pic monoisotopique. Au contraire, les échantillons marqués en fin de tubérisation montrent un profil d'incorporation, avec des massifs de pics bimodaux, correspondant à la superposition des massifs marqués et non marqués. Ceci indique que les protéines ne sont pas seulement séquestrées dans le grain à un instant  $t$ , mais qu'elles sont synthétisées et associées de façon continue. On remarque toutefois que les protéines marquées en début de tubérisation présentent un taux d'incorporation plus important que les protéines marquées en milieu de tubérisation. Les premières ont pourtant bénéficié de 3 semaines de

chasse après le marquage, là où celles marquées en fin de tubérisation n'ont subi que 1,5 semaines de chasse. Ceci suggère que les protéines retrouvées dans le grain sont synthétisées de manière plus importante dès le début de la tubérisation, de 0 à 3 semaines, par rapport au milieu de la tubérisation de 1,5 à 4,5 semaines.

De fortes différences sont retrouvées lors du marquage en fin de tubérisation, rendant de fait difficile l'interprétation de ces résultats. Néanmoins, l'incorporation maximale dans ces conditions reste inférieure à celle observée chez les plantes marquées en début de tubérisation. Les premières ont subi 3 semaines de chasse alors que les dernières n'ont pas subi de chasse. L'incorporation reste donc plus importante au début de la tubérisation qu'à la fin. A noter toutefois que la synthèse à un même moment n'implique pas forcément un adressage au même instant de ces protéines. Il est possible que l'ensemble des protéines soient synthétisées à un instant  $t$ , mais que leur adressage au grain ait lieu plus tard, et à des moments différents pour chaque protéine. La présence d'inhibiteurs de protéases à la surface des grains d'amidon pourrait indiquer que les protéines du métabolisme de l'amidon ont besoin d'une protection contre la dégradation et qu'elles doivent donc persister dans le temps (Helle et al., 2018). Les études sur GBSS ont montré que celle-ci est instable sous forme soluble, et qu'elle n'est stable que lors de son adressage au grain (Dry et al., 1992). Plusieurs enzymes, tel que SS3, GWD, SEX4 et SEX4-like, n'ont pas montré de différence significative dans leur taux d'incorporation, quel que soit l'instant du marquage. Cela laisse penser que celle-ci ont eu le même comportement et ont été adressé au même moment au grain. Concernant les enzymes de branchements, celles-ci ont un taux d'incorporation plus élevé lors du début de la tubérisation que les autres protéines étudiées. Au contraire, les amidons synthases SS1 et SS2, LESV, PHS1a et LSF2 ont un taux d'incorporation plus faible que les autres lorsque le marquage a eu lieu au début de la tubérisation. L'analyse transcriptomique du tubercule de pomme de terre à différentes étapes de sa croissance, ainsi que des analyses protéomiques après marquage sur les protéines solubles du tubercule de pomme de terre, nous permettrait de confirmer ou d'infirmer si les protéines synthétisées au début de la tubérisation sont directement adressées au grain ou si celles-ci persistent un certain temps avant leur adressage. Les études précédentes que nous avons réalisées sur le protéome de l'amidon de pomme de terre ont montré que la stœchiométrie des protéines fluctue en fonction de la morphologie des grains (Helle et al., 2019). Nous avons émis l'hypothèse que la stœchiométrie des protéines pourrait déterminer la morphologie du grain, ou qu'elle pourrait être représentative d'un état de croissance. L'état actuel des résultats ne permet pas de répondre clairement à cette question. Il est nécessaire d'étudier l'incorporation

du  $^{15}\text{N}$  dans le stroma du plaste et à l'interface stroma/grain d'amidon. Il est de plus nécessaire d'établir le turnover des protéines dans le tubercule, en récoltant ces derniers à différents stades de croissance dans un milieu enrichi en  $^{15}\text{N}$ . Ces expérimentations donneront une vision plus claire des étapes de synthèse de l'amidon.

#### **IV. Impact d'une protéine sur l'ensemble du protéome, le cas des mutants *gbss***

Afin d'étudier l'impact de l'absence de GBSS sur le protéome, nous avons analysé les protéines extraites de grains d'amidon sauvages ou issus de plantes mutées au locus *GBSS* par SDS-PAGE et protéomique quantitative shotgun. L'un des mutants utilisés (17T.701.010) possède encore un allèle sauvage et l'autre (17T.716.146) est un mutant knock-out dans lequel les 4 allèles sont inactivés (Veillet et al., 2019). Sur gel SDS-PAGE, la bande correspondant à GBSS disparaît dans les deux mutants. La quantité de polypeptides résiduels de GBSS dans le mutant 17T.701.010 se trouve très probablement sous le seuil de détection de la révélation au bleu de Coomassie. Ceci conduit à la disparition de la bande correspondante dans ce mutant, bien qu'un allèle sauvage soit toujours présent. Pour les deux mutants, l'absence de GBSS sur SDS-PAGE s'accompagne de l'apparition d'autres bandes. L'intensité de ces bandes est plus forte dans la lignée 17T.716.146 que dans la lignée 17T.701.010. Un dosage a été réalisé mais les résultats du dosage ne concordent pas avec ceux obtenus lors du gel d'électrophorèse. En effet, la concentration en protéine est similaire entre l'amidon du mutant 17T.701.010 et celui du sauvage, alors que le gel indique très clairement une plus forte présence de protéine dans l'amidon sauvage. Ceci peut être dû à un mauvais dessalage de l'échantillon avant analyse. Le dosage utilisé est un dosage BCA, qui a l'avantage d'être compatible avec le SDS, mais qui reste incompatible aux tampons contenant du DTT. Qui plus est, le BCA réagit avec l'extrémité réductrice des malto-oligosaccharides. Dans les premières analyses, cela n'avait pas posé de problème du fait du grand volume de surnageant récupéré, qui limitait la présence de malto-oligosaccharides, et de la concentration sur Amicon 50 kDa, qui devait en éliminer une grande partie encore présente. Cette méthode n'est faisable qu'avec une grande quantité d'amidon (300 mg). Hors, pour les expériences sur les mutants *waxy*, la quantité d'amidon disponible était beaucoup plus limitée et ne permettait pas de réaliser un dosage, la quantité d'amidon alloué pour celui-ci étant limitée à 30 mg. Des développements sont actuellement en cours pour trouver un moyen de doser les protéines sur une faible quantité d'amidon (10 mg). !ce dosage permettra de convertir les proportions relatives de chaque protéines obtenus par spectrométrie de masse

en quantité absolue, et ainsi d'avoir une idée plus claire de l'impact de l'absence de GBSS au sein du grain sur l'ensemble du protéome.

Nous avons réalisé les premières analyses protéomiques (l'identification du protéome et l'étude de sa variabilité au sein d'un échantillon en fonction de la morphologie des grains) sur de l'amidon extrait du cultivar Monalisa dont les plantes ont été cultivées en pleine terre. Pour des raisons pratiques, le marquage au  $^{15}\text{N}$  et l'étude des mutants *gbss* ont été effectuées à l'aide du cultivar Désirée. En effet, une lignée axénique de ce cultivar nous a été fournie par l'IGEPP INRA de Ploudaniel, permettant la culture des plants en conditions contrôlées. De plus, le cultivar Désirée est réputé pour sa bonne propension à la transformation et les mutants CRISPR/Cas9 utilisés dans cette étude ont été générés dans ce fond génétique. Lors de l'étude protéomique de l'amidon du cultivar Désirée, nous avons identifié les mêmes protéines que celles précédemment découvertes chez Monalisa (Helle et al., 2018). Toutefois, la stœchiométrie des protéines est différente entre les deux cultivars, indiquant que des variations sont possibles en fonction du cultivar et/ou des conditions de culture. Une étude réalisée sur différents cultivars ayant été cultivés et récoltés dans les mêmes conditions permettrait d'éclaircir ce point.

Les analyses par protéomique comparative des mutants *gbss* montrent que, au-delà de son rôle dans la synthèse de l'amylose, GBSS a aussi un impact sur la stœchiométrie des protéines de l'amidon. Cela peut être un effet soit direct (en conséquence de l'absence de GBSS) soit indirect (en conséquence de l'absence d'amylose). Dans l'amidon des mutants *gbss*, la proportion de la quasi-totalité des protéines identifiées augmente par rapport au sauvage. Cette augmentation est exacerbée dans la lignée 17T.716.146 par rapport à la lignée 17T.701.010. On remarque que, pour la protéine GBSS, une faible proportion de peptides pour la protéine GBSS ont tout de même été identifiés. Il est possible que dans le mutant 17T.716.146, l'un des allèles mutés exprime tout de même une forme inactive de GBSS, Cela expliquerai la présence de GBSS dans ces mutants malgré l'absence d'amylose constatée, mais parait peu probable du fait que la mutation induite entraîne un décalage du cadre ouvert de lecture (Veillet et al., 2019). Une autre possibilité, plus plausible, est la présence d'un ou quelques grains Wild Type dans l'échantillon par contamination lors des étapes de préparation, les échantillons WT et 17T.716.146 ayant été préparé en même temps. En comparant le nombre de peptides uniques identifiés sur un grain unique et dans le mutant 17T.716.146, on remarque que le nombre de peptides uniques identifiés pour GBSS sur un seul grain WT est supérieur à ceux identifié à partir de 300 mg d'amidon *waxy*.

Bien que la proportion relative de la plupart des protéines soit augmentée dans les mutants 17T.701.010 et 17T.716.146, une exception est en revanche observée pour la protéine PTST1. Celle-ci est absente du grain dans les 2 mutants alors qu'elle est présente dans l'amidon sauvage. Chez *Arabidopsis*, il a été montré que PTST1 permet l'adressage de GBSS au grain d'amidon (Seung et al., 2015). Son absence du grain lorsque la quantité de GBSS est fortement diminuée concorde avec une activité d'adressage de PTST1. De plus, les auteurs ont émis l'hypothèse que PTST1 n'est présente que de manière transitoire dans le grain chez *Arabidopsis thaliana* (Seung et al., 2015). Le fait que cette protéine ne soit présente qu'en très faible quantité et que sa quantité ne soit pas corrélée à celle de GBSS dans les analyses de populations de grains fractionnés indique un mécanisme similaire chez la pomme de terre (Helle et al., 2019). Les protéines PTST1 retrouvées dans le grain correspondrait soit à une séquestration dans le grain suite à des événements stochastiques, soit à des protéines encore en association avec GBSS.

Il est possible que GBSS, de manière directe ou indirecte, régule l'adressage des autres protéines au grain d'amidon, ou alors empêche la séquestration de celles-ci dans la matrice du grain. La vérification de cette hypothèse passe par l'étude de mutants exprimant une forme de GBSS intègre mais inactive. La mise en œuvre de l'édition de bases ciblée par CRISPR/Cas9 chez la pomme de terre permet de produire de tels mutants. Veillet et collaborateurs ont récemment généré plusieurs lignées exprimant une GBSS mutée dans l'un des motifs catalytiques des amidon-synthases (Veillet et al., 2019). Ces lignées n'accumulent plus d'amylose et se révèlent des outils de choix pour tester notre hypothèse.

## **V. Protéomique du grain unique**

De manière à étudier les protéines liées au grain d'amidon dans des échantillons miniaturisés, nous avons isolé un ou 30 grains d'amidon à l'aide d'un capillaire de silice et adapté le protocole d'extraction et de digestion des protéines pour limiter les pertes de matériel. Les peptides ont ensuite été analysés par spectrométrie de masse pour identifier et quantifier les protéines. La quantification des 3 protéines majoritaires à partir de 30 grains révèle une stoechiométrie similaire à celle observée avec des quantités d'amidon plus importantes. GBSS est la protéine majoritaire, suivie de SS2 puis de GWD. L'analyse des protéines extraites à partir d'un grain unique d'une taille de 70  $\mu\text{m}$  a permis l'identification des deux protéines majoritaires, GBSS et SS2. Dans cette analyse, GBSS a été identifiée avec une couverture de séquence de 30%. Le signal peut aussi être virtuellement amplifié pour les peptides de GBSS, par l'usage de peptides AQUA, ou par l'usage de marqueurs iTRAQ en mélangeant ensemble



un échantillon d'un unique grain et un échantillon contenant un plus grand nombre de grains. En l'état actuel, il semble toutefois peu probable d'identifier plus de 5 protéines avec un seul grain, et d'identifier plus de protéines que GBSS et SS2 à partir d'un grain de 20  $\mu\text{m}$  du fait de la limite de détection de la spectrométrie de masse. Il est toutefois envisageable d'étudier le ratio entre GBSS et SS2 au sein de grains d'amidon uniques. La compréhension des mécanismes de la synthèse d'un grain nécessite de travailler à l'échelle d'un grain. La synthèse de l'amidon n'est probablement pas un évènement unique et répété et plusieurs voies de synthèse peuvent potentiellement coexister au sein d'une même espèce, voire d'un même tissu. Les études réalisées sur les amidon-synthases 3, 4 et 5 ont déjà montré l'existence de plusieurs voies pour l'initiation de la synthèse (Figure 50)(Szydlowski et al., 2009; Abt et al., 2020). Comprendre les mécanismes régissant la synthèse d'un grain permettrait d'en faire la modélisation mathématique et de définir les variables qui permettraient la resynthèse d'amidon *in vitro*. Différentes approches ont déjà été implémentées pour synthétiser de l'amidon *in vitro*, soit pour produire un polymère imitant l'amidon natif, soit pour produire un amidon dérivé chimiquement (O'Neill and Field, 2015). Ces approches visent à comprendre les mécanismes d'adressage des glucanes au site de synthèse, le fonctionnement des enzymes du métabolisme de l'amidon et le maintien de la structure de ce dernier. Ces recherches ont un impact direct en recherche fondamentale, permettant de mieux comprendre les mécanismes de synthèse *in planta*, mais aussi en recherche industrielle, permettant à long terme la synthèse d'amidon « à la carte », avec une morphologie et une cristallinité prédéfinie. La biologie synthétique nécessite de travailler à l'échelle de la cellule ou de l'objet unique. Tenter de reproduire des systèmes biologiques complexes *in vitro* par ajustements empiriques des variables, est laborieux et chronophage. Etablir des modèles mathématiques se basant sur ou extrapolés à partir de mesures biologiques permet d'ajuster précisément ces variables. La synthèse de l'amidon, comme montré au long de ce mémoire, est un mécanisme complexe nécessitant un ensemble d'activités enzymatiques, agissant en synergie tout au long de la synthèse. Ces évènements peuvent aussi être séquentiels comme par exemple les produits des amidon-synthases servant de substrats à d'autres amidon-synthases et/ou des enzymes de branchement. Mais cette synthèse est aussi marquée d'évènements localisés et ponctuels, comme la formation du hile par exemple et le processus de nucléation primordial des premières molécules d'amylopectine. Ces évènements sont aussi dépendants de l'environnement interne du plaste qui n'est pas homogène. Ainsi la localisation du complexe d'initiation dans le plaste doit répondre à certaines exigences pour permettre d'initier la synthèse du polysaccharide de réserve (Gámez-Arjona et al., 2014; Seung et al., 2017; Seung et al., 2018). Cependant il reste encore

aujourd'hui des zones d'ombre dans notre compréhension du processus de synthèse, notamment au niveau de l'initiation qui semble requérir l'existence d'un glucane primordial d'un DP>10. Les hypothèses actuelles de la synthèse de ce glucane suggèrent l'implication des amidon-phosphorylases et d'évènements stochastiques (Seung et al., 2017). Par ailleurs, la façon dont les protéines s'associent au grain et quand elles le font restent encore à élucider de même que leur répartition entre le stroma et la matrice d'amidon. Imiter une telle machinerie nécessite une connaissance parfaite de l'ensemble des paramètres, tout en tenant compte des biais insérés par l'environnement propre du grain d'amidon lors de sa synthèse. Parmi ces biais, on retrouve le type cellulaire (la synthèse peut différer d'une cellule à une autre), la prise de décision cellulaire (quand et comment la cellule décide d'initier la synthèse d'un grain) et le bruit cellulaire (tenir compte de l'ensemble des évènements stochastiques qui vont interférer avec le « schéma classique » de la synthèse du grain). La synthèse *in vitro* d'amidon reste à ce jour un objectif distant à atteindre. Les analyses menées sur le grain unique, couplées aux études cinétiques à l'interface solide-liquide de l'ensemble des enzymes du métabolisme, aux études de dynamiques d'association des protéines au grain d'amidon, ainsi qu'à la caractérisation du hile de l'amidon, permettront *in fine* d'établir les modèles mathématiques nécessaires à cet objectif. Arriver à synthétiser *in vitro* de l'amidon permettra aussi, par la compréhension du métabolisme que cela apportera, de mieux choisir les cibles pour l'optimisation des plantes pour la production d'un amidon spécifique.

A moyen terme, l'utilisation des analyses protéomiques sur grain unique peuvent servir pour des objectifs transverses. Une couverture de séquence suffisamment importante de GBSS permettrait des analyses de grains d'amidon anciens originaires d'échantillons archéologiques et pourrait être utilisée pour l'identification de l'espèce auquel le grain appartient à partir de la séquence de GBSS. En effet, les variations dans les séquences des protéines représentent un critère avec un fort degré de confiance pour l'identification d'espèces, du fait de la quasi absence de variations intra-espèce (López et al., 2002; Buckley et al., 2009). De ce fait, utiliser les analyses en spectrométrie de masse pour discriminer les peptides spécifiques d'une espèce au sein du grain est une méthode robuste pour l'identification de l'espèce auquel le grain appartient.

# Bibliographie

- Abel, G.J., Springer, F., Willmitzer, L., and Kossmann, J.** (1996). Cloning and functional analysis of a cDNA encoding a novel 139 kDa starch synthase from potato (*Solanum tuberosum* L.). *The Plant Journal* **10**, 981-991.
- Abt, M., Pfister, B., Sharma, M., Eicke, S., Bürgy, L., Neale, I., Seung, D., and Zeeman, S.C.** (2020). STARCH SYNTHASE 5, a Noncanonical Starch Synthase-Like Protein, Promotes Starch Granule Initiation in *Arabidopsis*. *Plant Cell*.
- Ahmed, Z., Tetlow, I.J., Ahmed, R., Morell, M.K., and Emes, M.J.** (2015). Protein-protein interactions among enzymes of starch biosynthesis in high-amylose barley genotypes reveal differential roles of heteromeric enzyme complexes in the synthesis of A and B granules. *Plant Science* **233**, 95-106.
- Ainseba-Chirani, N., Dembahri, Z., Tokarski, C., Rolando, C., and Benmouna, M.** (2011). Newly designed polyacrylamide/dextran gels for electrophoresis protein separation: synthesis and characterization. *Polymer international* **60**, 1024-1029.
- Baba, T., Nishihara, M., Mizuno, K., Kawasaki, T., Shimada, H., Kobayashi, E., Ohnishi, S., Tanaka, K., and Arai, Y.** (1993). Identification, cDNA cloning, and gene expression of soluble starch synthase in rice (*Oryza sativa* L.) immature seeds. *Plant physiology* **103**, 565-573.
- Ball, S., Guan, H.-P., James, M., Myers, A., Keeling, P., Mouille, G., Buléon, A., Colonna, P., and Preiss, J.** (1996). From glycogen to amylopectin: a model for the biogenesis of the plant starch granule. *Cell* **86**, 349-352.
- Ball, S.G.** (1998). Regulation of Starch Biosynthesis. In *The Molecular Biology of Chloroplasts and Mitochondria in Chlamydomonas*, J.D. Rochaix, M. Goldschmidt-Clermont, and S. Merchant, eds (Dordrecht: Springer Netherlands), pp. 549-567.
- Ball, S.G., Colleoni, C., Kadouche, D., Ducatez, M., Arias, M.-C., and Tirtiaux, C.** (2015). Toward an understanding of the function of Chlamydiales in plastid endosymbiosis. *Biochimica et Biophysica Acta (BBA)-Bioenergetics* **1847**, 495-504.
- Baunsgaard, L., Lütken, H., Mikkelsen, R., Glaring, M.A., Pham, T.T., and Blennow, A.** (2005). A novel isoform of glucan, water dikinase phosphorylates pre-phosphorylated  $\alpha$ -glucans and is involved in starch degradation in *Arabidopsis*. *The Plant Journal* **41**, 595-605.
- Bender, H., and Wallenfels, K.** (1966). [95] Pullulanase (an amylopectin and glycogen debranching enzyme) from *Aerobacter aerogenes*. In *Methods in Enzymology* (Academic Press), pp. 555-559.
- Biliaderis, C., Grant, D., and Vose, J.** (1979). Molecular weight distributions of legume starches by gel chromatography. *Cereal Chem* **56**, 475-480.
- Black, R., Loerch, J., McArdle, F., and Creech, R.** (1966). Genetic interactions affecting maize phytoglycogen and the phytoglycogen-forming branching enzyme. *Genetics* **53**, 661.

- Blauth, S.L., Yao, Y., Klucinec, J.D., Shannon, J.C., Thompson, D.B., and Guiltinan, M.J.** (2001). Identification of Mutator insertional mutants of starch-branching enzyme 2a in corn. *Plant Physiology* **125**, 1396-1405.
- Blauth, S.L., Kim, K.-N., Klucinec, J., Shannon, J.C., Thompson, D., and Guiltinan, M.** (2002). Identification of Mutator insertional mutants of starch-branching enzyme 1 (sbe1) in *Zea mays* L. *Plant molecular biology* **48**, 287-297.
- Blennow, A., and Engelsen, S.B.** (2010). Helix-breaking news: fighting crystalline starch energy deposits in the cell. *Trends in Plant Science* **15**, 236-240.
- Borén, M., Larsson, H., Falk, A., and Jansson, C.** (2004). The barley starch granule proteome—internalized granule polypeptides of the mature endosperm. *Plant Science* **166**, 617-626.
- Boyer, C.D., and Preiss, J.** (1978). Multiple forms of (1→4)- $\alpha$ -D-glucan,(1→4)- $\alpha$ -D-glucan-6-glycosyl transferase from developing *Zea mays* L. kernels. *Carbohydrate Research* **61**, 321-334.
- Boyer, C.D., and Preiss, J.** (1979). Properties of citrate-stimulated starch synthesis catalyzed by starch synthase I of developing maize kernels. *Plant physiology* **64**, 1039-1042.
- Brummell, D.A., Watson, L.M., Zhou, J., McKenzie, M.J., Hallett, I.C., Simmons, L., Carpenter, M., and Timmerman-Vaughan, G.M.** (2015). Overexpression of STARCH BRANCHING ENZYME II increases short-chain branching of amylopectin and alters the physicochemical properties of starch from potato tuber. *BMC biotechnology* **15**, 28.
- Buchanan, B.B.** (1984). The ferredoxin/thioredoxin system: a key element in the regulatory function of light in photosynthesis. *Bioscience* **34**, 378-383.
- Buckley, M., Collins, M., Thomas-Oates, J., and Wilson, J.C.** (2009). Species identification by analysis of bone collagen using matrix-assisted laser desorption/ionisation time-of-flight mass spectrometry. *Rapid Communications in Mass Spectrometry: An International Journal Devoted to the Rapid Dissemination of Up-to-the-Minute Research in Mass Spectrometry* **23**, 3843-3854.
- Buléon, A., Colonna, P., Planchot, V., and Ball, S.** (1998). Starch granules: structure and biosynthesis. *International journal of biological macromolecules* **23**, 85-112.
- Bull, S.E., Seung, D., Chanez, C., Mehta, D., Kuon, J.E., Truernit, E., Hochmuth, A., Zurkirchen, I., Zeeman, S.C., Gruissem, W., and Vanderschuren, H.** (2018). Accelerated ex situ breeding of GBSS- and PTST1-edited cassava for modified starch. *Sci Adv* **4**, eaat6086.
- Burton, R.A., Bewley, J.D., Smith, A.M., Bhattacharyya, M.K., Tatge, H., Ring, S., Bull, V., Hamilton, W.D., and Martin, C.** (1995). Starch branching enzymes belonging to distinct enzyme families are differentially expressed during pea embryo development. *The Plant Journal* **7**, 3-15.
- Burton, R.A., Jenner, H., Carrangis, L., Fahy, B., Fincher, G.B., Hylton, C., Laurie, D.A., Parker, M., Waite, D., and Van Wegen, S.** (2002). Starch granule initiation

- and growth are altered in barley mutants that lack isoamylase activity. *The Plant Journal* **31**, 97-112.
- Bustos, R., Fahy, B., Hylton, C.M., Seale, R., Nebane, N.M., Edwards, A., Martin, C., and Smith, A.M.** (2004). Starch granule initiation is controlled by a heteromultimeric isoamylase in potato tubers. *Proceedings of the National Academy of Sciences* **101**, 2215-2220.
- Cao, H., Imparl-Radosevich, J., Guan, H., Keeling, P.L., James, M.G., and Myers, A.M.** (1999). Identification of the soluble starch synthase activities of maize endosperm. *Plant physiology* **120**, 205-216.
- Caspar, T., Lin, T.-P., Kakefuda, G., Benbow, L., Preiss, J., and Somerville, C.** (1991). Mutants of *Arabidopsis* with Altered Regulation of Starch Degradation. *Plant Physiology* **95**, 1181-1188.
- Chia, T., Chirico, M., King, R., Ramirez-Gonzalez, R., Saccomanno, B., Seung, D., Simmonds, J., Trick, M., Uauy, C., and Verhoeven, T.** (2019a). A carbohydrate-binding protein, FLOURY ENDOSPERM 6 influences the initiation of A- and B-type starch granules in wheat. *bioRxiv*, 643759.
- Chia, T., Chirico, M., King, R., Ramirez-Gonzalez, R., Saccomanno, B., Seung, D., Simmonds, J., Trick, M., Uauy, C., Verhoeven, T., and Trafford, K.** (2019b). A carbohydrate-binding protein, B-GRANULE CONTENT 1, influences starch granule size distribution in a dose-dependent manner in polyploid wheat. *Journal of Experimental Botany* **71**, 105-115.
- Commuri, P.D., and Keeling, P.L.** (2001). Chain-length specificities of maize starch synthase I enzyme: studies of glucan affinity and catalytic properties. *The Plant Journal* **25**, 475-486.
- Comparot-Moss, S., Kötting, O., Stettler, M., Edner, C., Graf, A., Weise, S.E., Streb, S., Lue, W.-L., MacLean, D., and Mahlow, S.** (2010). A putative phosphatase, LSF1, is required for normal starch turnover in *Arabidopsis* leaves. *Plant physiology* **152**, 685-697.
- Coutinho, P.M., Deleury, E., and Henrissat, B.** (2003). The families of carbohydrate-active enzymes in the genomic era. *Journal of applied glycoscience* **50**, 241-244.
- Craig, J., Lloyd, J.R., Tomlinson, K., Barber, L., Edwards, A., Wang, T.L., Martin, C., Hedley, C.L., and Smith, A.M.** (1998). Mutations in the gene encoding starch synthase II profoundly alter amylopectin structure in pea embryos. *The Plant Cell* **10**, 413-426.
- Critchley, J.H., Zeeman, S.C., Takaha, T., Smith, A.M., and Smith, S.M.** (2001). A critical role for disproportionating enzyme in starch breakdown is revealed by a knock-out mutation in *Arabidopsis*. *The Plant Journal* **26**, 89-100.
- Crumpton-Taylor, M., Grandison, S., Png, K.M., Bushby, A.J., and Smith, A.M.** (2012). Control of starch granule numbers in *Arabidopsis* chloroplasts. *Plant Physiology* **158**, 905-916.
- Crumpton-Taylor, M., Pike, M., Lu, K.J., Hylton, C.M., Feil, R., Eicke, S., Lunn, J.E., Zeeman, S.C., and Smith, A.M.** (2013). Starch synthase 4 is essential for

- coordination of starch granule formation with chloroplast division during *Arabidopsis* leaf expansion. *New Phytologist* **200**, 1064-1075.
- Dang, P.L., and Boyer, C.D.** (1988). Maize leaf and kernel starch synthases and starch branching enzymes. *Phytochemistry* **27**, 1255-1259.
- Dang, P.L., and Boyer, C.D.** (1989). Comparison of soluble starch synthases and branching enzymes from leaves and kernels of normal and amylose-extender maize. *Biochemical genetics* **27**, 521-532.
- Dauvillée, D., Colleoni, C., Mouille, G., Morell, M.K., d'Hulst, C., Wattedled, F., Liénard, L., Delvallé, D., Ral, J.-P., and Myers, A.M.** (2001a). Biochemical characterization of wild-type and mutant isoamylases of *Chlamydomonas reinhardtii* supports a function of the multimeric enzyme organization in amylopectin maturation. *Plant Physiology* **125**, 1723-1731.
- Dauvillée, D., Colleoni, C., Mouille, G., Buléon, A., Gallant, D.J., Bouchet, B., Morell, M.K., d'Hulst, C., Myers, A.M., and Ball, S.G.** (2001b). Two loci control phytoglycogen production in the monocellular green alga *Chlamydomonas reinhardtii*. *Plant Physiology* **125**, 1710-1722.
- Dauvillée, D., Chochois, V., Steup, M., Haebel, S., Eckermann, N., Ritte, G., Ral, J.P., Colleoni, C., Hicks, G., and Wattedled, F.** (2006). Plastidial phosphorylase is required for normal starch synthesis in *Chlamydomonas reinhardtii*. *The Plant Journal* **48**, 274-285.
- Delatte, T., Trevisan, M., Parker, M.L., and Zeeman, S.C.** (2005). *Arabidopsis* mutants *Atisa1* and *Atisa2* have identical phenotypes and lack the same multimeric isoamylase, which influences the branch point distribution of amylopectin during starch synthesis. *The Plant Journal* **41**, 815-830.
- Delatte, T., Umhang, M., Trevisan, M., Eicke, S., Thorneycroft, D., Smith, S.M., and Zeeman, S.C.** (2006). Evidence for Distinct Mechanisms of Starch Granule Breakdown in Plants. *Journal of Biological Chemistry* **281**, 12050-12059.
- Delrue, B., Fontaine, T., Routier, F., Decq, A., Wieruszkeski, J.M., Van Den Koornhuysse, N., Maddelein, M.L., Fournet, B., and Ball, S.** (1992). Waxy *Chlamydomonas reinhardtii*: monocellular algal mutants defective in amylose biosynthesis and granule-bound starch synthase activity accumulate a structurally modified amylopectin. *Journal of Bacteriology* **174**, 3612-3620.
- Delvallé, D., Dumez, S., Wattedled, F., Roldán, I., Planchot, V., Berbezy, P., Colonna, P., Vyas, D., Chatterjee, M., and Ball, S.** (2005). Soluble starch synthase I: a major determinant for the synthesis of amylopectin in *Arabidopsis thaliana* leaves. *The Plant Journal* **43**, 398-412.
- Denyer, K., and Smith, A.M.** (1992). The purification and characterisation of the two forms of soluble starch synthase from developing pea embryos. *Planta* **186**, 609-617.
- Denyer, K., Sidebottom, C., Hylton, C.M., and Smith, A.M.** (1993). Soluble isoforms of starch synthase and starch-branching enzyme also occur within starch granules in developing pea embryos. *The Plant Journal* **4**, 191-198.

- Denyer, K., Hylton, C., Jenner, C.F., and Smith, A.** (1995a). Identification of multiple isoforms of soluble and granule-bound starch synthase in developing wheat endosperm. *Planta* **196**, 256-265.
- Denyer, K., Waite, D., Motawia, S., Møller, B.L., and Smith, A.M.** (1999). Granule-bound starch synthase I in isolated starch granules elongates malto-oligosaccharides processively. *Biochemical Journal* **340**, 183-191.
- Denyer, K., Barber, L., Burton, R., Hedley, C., Hylton, C., Johnson, S., Jones, D.A., Marshall, J., Smith, A., Tatge, H., Tomlinson, K., and Wang, T.** (1995b). The isolation and characterization of novel low-amylose mutants of *Pisum sativum* L. *Plant, Cell & Environment* **18**, 1019-1026.
- Deschamps, P., Moreau, H., Worden, A.Z., Dauvillée, D., and Ball, S.G.** (2008). Early gene duplication within chloroplasts and its correspondence with relocation of starch metabolism to chloroplasts. *Genetics* **178**, 2373-2387.
- Dian, W., Jiang, H., and Wu, P.** (2005). Evolution and expression analysis of starch synthase III and IV in rice. *Journal of Experimental Botany* **56**, 623-632.
- Dinges, J.R., Colleoni, C., Myers, A.M., and James, M.G.** (2001). Molecular structure of three mutations at the *maizesugary1* locus and their allele-specific phenotypic effects. *Plant Physiology* **125**, 1406-1418.
- Dinges, J.R., Colleoni, C., James, M.G., and Myers, A.M.** (2003). Mutational analysis of the pullulanase-type debranching enzyme of maize indicates multiple functions in starch metabolism. *The Plant Cell* **15**, 666-680.
- Doehlert, D.C., and Knutson, C.A.** (1991). Two classes of starch debranching enzymes from developing maize kernels. *Journal of Plant Physiology* **138**, 566-572.
- Dong, X., Zhang, D., Liu, J., Liu, Q.Q., Liu, H., Tian, L., Jiang, L., and Qu, L.Q.** (2015). Plastidial Disproportionating Enzyme Participates in Starch Synthesis in Rice Endosperm by Transferring Maltooligosyl Groups from Amylose and Amylopectin to Amylopectin. *Plant Physiology* **169**, 2496-2512.
- Dry, I., Smith, A., Edwards, A., Bhattacharyya, M., Dunn, P., and Martin, C.** (1992). Characterization of cDNAs encoding two isoforms of granule-bound starch synthase which show differential expression in developing storage organs of pea and potato. *The Plant Journal* **2**, 193-202.
- Dumez, S., Wattebled, F., Dauvillee, D., Delvalle, D., Planchot, V., Ball, S.G., and d'Hulst, C.** (2006). Mutants of *Arabidopsis* lacking starch branching enzyme II substitute plastidial starch synthesis by cytoplasmic maltose accumulation. *The Plant Cell* **18**, 2694-2709.
- Edwards, A., Marshall, J., Sidebottom, C., Visser, R.G., Smith, A.M., and Martin, C.** (1995). Biochemical and molecular characterization of a novel starch synthase from potato tubers. *The Plant Journal* **8**, 283-294.
- Emanuelsson, O., Nielsen, H., and von Heijne, G.** (1999). ChloroP, a neural network-based method for predicting chloroplast transit peptides and their cleavage sites. *Protein Sci* **8**, 978-984.



- Erde, J., Loo, R.R.O., and Loo, J.A.** (2014). Enhanced FASP (eFASP) to increase proteome coverage and sample recovery for quantitative proteomic experiments. *Journal of proteome research* **13**, 1885-1895.
- Facon, M., Lin, Q., Azzaz, A.M., Hennen-Bierwagen, T.A., Myers, A.M., Putaux, J.-L., Roussel, X., D'Hulst, C., and Wattedled, F.** (2013). Distinct functional properties of isoamylase-type starch debranching enzymes in monocot and dicot leaves. *Plant physiology* **163**, 1363-1375.
- Fan, K.-T., Rendahl, A.K., Chen, W.-P., Freund, D.M., Gray, W.M., Cohen, J.D., and Hegeman, A.D.** (2016). Proteome Scale-Protein Turnover Analysis Using High Resolution Mass Spectrometric Data from Stable-Isotope Labeled Plants. *Journal of Proteome Research* **15**, 851-867.
- Fannon, J.E., Hauber, R.J., and BeMILLER, J.N.** (1992). Surface pores of starch granules. *Cereal Chem* **69**, 284-288.
- Fannon, J.E., Shull, J.M., and BeMILLER, J.N.** (1993). Interior channels of starch granules. *Cereal chemistry* **70**, 611-611.
- Feike, D., Seung, D., Graf, A., Bischof, S., Ellick, T., Coiro, M., Soyk, S., Eicke, S., Mettler-Altmann, T., Lu, K.J., Trick, M., Zeeman, S.C., and Smith, A.M.** (2016). The Starch Granule-Associated Protein EARLY STARVATION1 Is Required for the Control of Starch Degradation in *Arabidopsis thaliana* Leaves. *The Plant Cell* **28**, 1472-1489.
- Ferreira, S.J., Senning, M., Fischer-Stettler, M., Streb, S., Ast, M., Neuhaus, H.E., Zeeman, S.C., Sonnewald, S., and Sonnewald, U.** (2017). Simultaneous silencing of isoamylases ISA1, ISA2 and ISA3 by multi-target RNAi in potato tubers leads to decreased starch content and an early sprouting phenotype. *PloS one* **12**.
- Fettke, J., Eckermann, N., Poeste, S., Pauly, M., and Steup, M.** (2004). The glycan substrate of the cytosolic (Pho 2) phosphorylase isozyme from *Pisum sativum* L.: identification, linkage analysis and subcellular localization. *Plant J* **39**, 933-946.
- Fettke, J., Poeste, S., Eckermann, N., Tiessen, A., Pauly, M., Geigenberger, P., and Steup, M.** (2005). Analysis of cytosolic heteroglycans from leaves of transgenic potato (*Solanum tuberosum* L.) plants that under- or overexpress the Pho 2 phosphorylase isozyme. *Plant Cell Physiol* **46**, 1987-2004.
- Fischer, M., Schreiber, L., Colby, T., Kuckenberg, M., Tacke, E., Hofferbert, H.-R., Schmidt, J., and Gebhardt, C.** (2013). Novel candidate genes influencing natural variation in potato tuber cold sweetening identified by comparative proteomics and association mapping. *BMC plant biology* **13**, 113.
- Fisher, D.K., Gao, M., Kim, K.-N., Boyer, C.D., and Guiltinan, M.J.** (1996). Two closely related cDNAs encoding starch branching enzyme from *Arabidopsis thaliana*. *Plant molecular biology* **30**, 97-108.
- Fontaine, T., d'Hulst, C., Maddelein, M.-L., Routier, F., Pepin, T.M., Decq, A., Wieruszeski, J.-M., Delrue, B., Van den Koornhuyse, N., and Bossu, J.-P.** (1993). Toward an understanding of the biogenesis of the starch granule.

- Evidence that *Chlamydomonas* soluble starch synthase II controls the synthesis of intermediate size glucans of amylopectin. *Journal of Biological Chemistry* **268**, 16223-16230.
- Francisco, P., Li, J., and Smith, S.M.** (2010). The gene encoding the catalytically inactive  $\beta$ -amylase BAM4 involved in starch breakdown in *Arabidopsis* leaves is expressed preferentially in vascular tissues in source and sink organs. *Journal of plant physiology* **167**, 890-895.
- French, D.** (1972). Fine structure of starch and its relationship to the organization of starch granules. *Journal of the Japanese Society of Starch Science* **19**, 8-25.
- Fujita, N., Kubo, A., Francisco, P.B., Nakakita, M., Harada, K., Minaka, N., and Nakamura, Y.** (1999). Purification, characterization, and cDNA structure of isoamylase from developing endosperm of rice. *Planta* **208**, 283-293.
- Fujita, N., Yoshida, M., Asakura, N., Ohdan, T., Miyao, A., Hirochika, H., and Nakamura, Y.** (2006). Function and characterization of starch synthase I using mutants in rice. *Plant physiology* **140**, 1070-1084.
- Fulton, D.C., Stettler, M., Mettler, T., Vaughan, C.K., Li, J., Francisco, P., Gil, M., Reinhold, H., Eicke, S., and Messerli, G.** (2008).  $\beta$ -AMYLASE4, a noncatalytic protein required for starch breakdown, acts upstream of three active  $\beta$ -amylases in *Arabidopsis* chloroplasts. *The Plant Cell* **20**, 1040-1058.
- Gallant, D.J., Bouchet, B., and Baldwin, P.M.** (1997). Microscopy of starch: evidence of a new level of granule organization. *Carbohydrate polymers* **32**, 177-191.
- Gómez-Arjona, F.M., Raynaud, S., Ragel, P., and Mérida, A.** (2014). Starch synthase 4 is located in the thylakoid membrane and interacts with plastoglobule-associated proteins in *Arabidopsis*. *Plant J* **80**, 305-316.
- Gao, M., Fisher, D.K., Kim, K.-N., Shannon, J.C., and Guiltinan, M.J.** (1996). Evolutionary conservation and expression patterns of maize starch branching enzyme I and IIIb genes suggests isoform specialization. *Plant molecular biology* **30**, 1223-1232.
- Gao, M., Wanat, J., Stinard, P.S., James, M.G., and Myers, A.M.** (1998). Characterization of *dull1*, a maize gene coding for a novel starch synthase. *The Plant Cell* **10**, 399-412.
- Geiger, D.R., and Servaites, J.C.** (1994). Diurnal regulation of photosynthetic carbon metabolism in C3 plants. *Annual review of plant biology* **45**, 235-256.
- Gentry, M.S., Downen, R.H., III, Worby, C.A., Mattoo, S., Ecker, J.R., and Dixon, J.E.** (2007). The phosphatase laforin crosses evolutionary boundaries and links carbohydrate metabolism to neuronal disease. *Journal of Cell Biology* **178**, 477-488.
- Gibson, T., Solah, V., and McCleary, B.** (1997). A procedure to measure amylose in cereal starches and flours with concanavalin A. *Journal of cereal science* **25**, 111-119.
- Glaring, M.A., Zygadlo, A., Thorneycroft, D., Schulz, A., Smith, S.M., Blennow, A., and Baunsgaard, L.** (2007). An extra-plastidial  $\alpha$ -glucan, water dikinase from

- Arabidopsis phosphorylates amylopectin in vitro and is not necessary for transient starch degradation. *Journal of experimental botany* **58**, 3949-3960.
- Godet, M.C., Bizot, H., and Buléon, A.** (1995). Crystallization of amylose—fatty acid complexes prepared with different amylose chain lengths. *Carbohydrate Polymers* **27**, 47-52.
- Gordon, R.W., Manners, D.J., and Stark, J.R.** (1975). The limit dextrinase of the broad bean (*Vicia faba* L.). *Carbohydrate Research* **42**, 125-134.
- Goren, A., Ashlock, D., and Tetlow, I.J.** (2018). Starch formation inside plastids of higher plants. *Protoplasma* **255**, 1855-1876.
- Grimaud, F., Rogniaux, H., James, M.G., Myers, A.M., and Planchot, V.** (2008). Proteome and phosphoproteome analysis of starch granule-associated proteins from normal maize and mutants affected in starch biosynthesis. *Journal of Experimental Botany* **59**, 3395-3406.
- Grimaud, F., Lancelon-Pin, C., Rolland-Sabaté, A., Roussel, X., Laguerre, S., Viksø-Nielsen, A., Putaux, J.-L., Guilois, S., Buléon, A., D'Hulst, C., and Potocki-Véronèse, G.** (2013). In Vitro Synthesis of Hyperbranched  $\alpha$ -Glucans Using a Biomimetic Enzymatic Toolbox. *Biomacromolecules* **14**, 438-447.
- Grzes, M., Szczerbal, I., Fijak-Nowak, H., Szydlowski, M., and Switonski, M.** (2011). Two candidate genes (FTO and INSIG2) for fat accumulation in four canids: chromosome mapping, gene polymorphisms and association studies of body and skin weight of red foxes. *Cytogenet Genome Res* **135**, 25-32.
- Guan, H., Li, P., Imparl-Radosevich, J., Preiss, J., and Keeling, P.** (1997). Comparing the properties of *Escherichia coli* branching enzyme and maize branching enzyme. *Archives of Biochemistry and Biophysics* **342**, 92-98.
- Han, X.Z., Benmoussa, M., Gray, J.A., BeMiller, J.N., and Hamaker, B.R.** (2005). Detection of proteins in starch granule channels. *Cereal Chemistry* **82**, 351-355.
- Han, Y., Sun, F.J., Rosales-Mendoza, S., and Korban, S.S.** (2007). Three orthologs in rice, *Arabidopsis*, and *Populus* encoding starch branching enzymes (SBEs) are different from other SBE gene families in plants. *Gene* **401**, 123-130.
- Harada, T., Misaki, A., Akai, H., Yokobayashi, K., and Sugimoto, K.** (1972). Characterization of *Pseudomonas* isoamylase by its actions on amylopectin and glycogen: Comparison with *Aerobacter* pullulanase. *Biochimica et Biophysica Acta (BBA) - Enzymology* **268**, 497-505.
- Harn, C., Knight, M., Ramakrishnan, A., Guan, H., Keeling, P.L., and Wasserman, B.P.** (1998). Isolation and characterization of the zSSIIa and zSSIIb starch synthase cDNA clones from maize endosperm. *Plant molecular biology* **37**, 639-649.
- Hayashi, M., Crofts, N., Oitome, N.F., and Fujita, N.** (2018). Analyses of starch biosynthetic protein complexes and starch properties from developing mutant rice seeds with minimal starch synthase activities. *BMC Plant Biol* **18**, 59.
- He, S., Hao, X., Wang, S., Zhou, W., Ma, Q., Lu, X., Chen, L., and Zhang, P.** (2020). A newly-identified inactive starch synthase simultaneously regulates starch

- synthesis and carbon allocation in storage roots of cassava (*Manihot esculenta* Crantz). *bioRxiv*, 2020.2003.2025.006957.
- Heibges, A., Salamini, F., and Gebhardt, C.** (2003). Functional comparison of homologous members of three groups of Kunitz-type enzyme inhibitors from potato tubers (*Solanum tuberosum* L.). *Molecular genetics and genomics* **269**, 535-541.
- Hejazi, M., Fettke, J., Kötting, O., Zeeman, S.C., and Steup, M.** (2010). The Laforin-like dual-specificity phosphatase SEX4 from *Arabidopsis* hydrolyzes both C6- and C3-phosphate esters introduced by starch-related dikinases and thereby affects phase transition of  $\alpha$ -glucans. *Plant physiology* **152**, 711-722.
- Helle, S., Bray, F., Putaux, J.-L., Verbeke, J., Flament, S., Rolando, C., D'Hulst, C., and Szydlowski, N.** (2019). Intra-Sample Heterogeneity of Potato Starch Reveals Fluctuation of Starch-Binding Proteins According to Granule Morphology. *Plants* **8**, 324.
- Helle, S., Bray, F., Verbeke, J., Devassine, S., Courseaux, A., Facon, M., Tokarski, C., Rolando, C., and Szydlowski, N.** (2018). Proteome Analysis of Potato Starch Reveals the Presence of New Starch Metabolic Proteins as Well as Multiple Protease Inhibitors. *Frontiers in Plant Science* **9**.
- Helm, S., Dobritzsch, D., Rödiger, A., Agne, B., and Baginsky, S.** (2014). Protein identification and quantification by data-independent acquisition and multi-parallel collision-induced dissociation mass spectrometry (MS(E)) in the chloroplast stroma proteome. *J Proteomics* **98**, 79-89.
- Hennen-Bierwagen, T.A., Liu, F., Marsh, R.S., Kim, S., Gan, Q., Tetlow, I.J., Emes, M.J., James, M.G., and Myers, A.M.** (2008). Starch biosynthetic enzymes from developing maize endosperm associate in multisubunit complexes. *Plant Physiology* **146**, 1892-1908.
- Hirose, T., and Terao, T.** (2004). A comprehensive expression analysis of the starch synthase gene family in rice (*Oryza sativa* L.). *Planta* **220**, 9-16.
- Hizukuri, S.** (1986). Polymodal distribution of the chain lengths of amylopectins, and its significance. *Carbohydr. Res.* **147**, 342-347.
- Hizukuri, S., Takeda, Y., Abe, J., Hanashiro, I., Matsunobu, G., and Kiyota, H.** (1997). Analytical developments: molecular and microstructural characterization. *Starch: structure and functionality*.
- Hong, J., An, D., Li, L., Liu, C., Li, M., Buckow, R., Zheng, X., and Bian, K.** (2020). Structural, rheological and gelatinization properties of wheat starch granules separated from different noodle-making process. *Journal of Cereal Science* **91**, 102897.
- Hou, J., Zhang, H., Liu, J., Reid, S., Liu, T., Xu, S., Tian, Z., Sonnewald, U., Song, B., and Xie, C.** (2017). Amylases StAmy23, StBAM1 and StBAM9 regulate cold-induced sweetening of potato tubers in distinct ways. *Journal of experimental botany* **68**, 2317-2331.
- Hsieh, C.-F., Liu, W., Whaley, J.K., and Shi, Y.-C.** (2019). Structure and functional properties of waxy starches. *Food Hydrocolloids* **94**, 238-254.

- Hussain, H., Mant, A., Seale, R., Zeeman, S., Hinchliffe, E., Edwards, A., Hylton, C., Bornemann, S., Smith, A.M., and Martin, C.** (2003). Three isoforms of isoamylase contribute different catalytic properties for the debranching of potato glucans. *The Plant Cell* **15**, 133-149.
- Imberty, A., Chanzy, H., Pérez, S., Buléon, A., and Tran, V.** (1988). The double-helical nature of the crystalline part of A-starch. *Journal of molecular biology* **201**, 365-378.
- Ishikawa, A., Ohta, S., Matsuoka, K., Hattori, T., and Nakamura, K.** (1994). A family of potato genes that encode Kunitz-type proteinase inhibitors: structural comparisons and differential expression. *Plant Cell Physiol* **35**, 303-312.
- Ishizaki, Y., Taniguchi, H., Maruyama, Y., and Nakamura, M.** (1983). Debranching Enzymes of Potato Tubers (*Solanum tuberosum* L.). I. Purification and Some Properties of Potato Isoamylase. *Agricultural and biological chemistry* **47**, 771-779.
- Ito, H., Hamada, S., Isono, N., Yoshizaki, T., Ueno, H., Yoshimoto, Y., Takeda, Y., and Matsui, H.** (2004). Functional characteristics of C-terminal regions of starch-branching enzymes from developing seeds of kidney bean (*Phaseolus vulgaris* L.). *Plant science* **166**, 1149-1158.
- James, M.G., Robertson, D.S., and Myers, A.M.** (1995). Characterization of the maize gene *sugary1*, a determinant of starch composition in kernels. *The Plant Cell* **7**, 417-429.
- Jane, J.-l., and Shen, J.J.** (1993). Internal structure of the potato starch granule revealed by chemical gelatinization. *Carbohydrate Research* **247**, 279-290.
- Jane, J.L., Kasemsuwan, T., Leas, S., Zobel, H., and Robyt, J.F.** (1994). Anthology of starch granule morphology by scanning electron microscopy. *Starch-Stärke* **46**, 121-129.
- Jenkins, P., and Donald, A.** (1995). The influence of amylose on starch granule structure. *International Journal of Biological Macromolecules* **17**, 315-321.
- Jeong, S.Y., Rose, A., and Meier, I.** (2003). MFP1 is a thylakoid-associated, nucleoid-binding protein with a coiled-coil structure. *Nucleic Acids Research* **31**, 5175-5185.
- Jiang, H., Dian, W., Liu, F., and Wu, P.** (2004). Molecular cloning and expression analysis of three genes encoding starch synthase II in rice. *Planta* **218**, 1062-1070.
- Kaneko, Y., Matsuda, S.-i., and Kadokawa, J.-i.** (2007). Chemoenzymatic Syntheses of Amylose-Grafted Chitin and Chitosan. *Biomacromolecules* **8**, 3959-3964.
- Kasemsuwan, T., and Jane, J.** (1994). Location of amylose in normal starch granules. II. Locations of phosphodiester cross-linking revealed by phosphorus-31 nuclear magnetic resonance. *Cereal chemistry* **71**, 282-286.
- Khalid, S., Yu, L., Meng, L., Liu, H., Ali, A., and Chen, L.** (2017). Poly(lactic acid)/starch composites: Effect of microstructure and morphology of starch granules on performance. *Journal of Applied Polymer Science* **134**, 45504.

- Kim, K.-N., and Gultinan, M.J.** (1999). Identification of cis-acting elements important for expression of the starch-branching enzyme I gene in maize endosperm. *Plant physiology* **121**, 225-236.
- Kossmann, J., Abel, G.J., Springer, F., Lloyd, J.R., and Willmitzer, L.** (1999). Cloning and functional analysis of a cDNA encoding a starch synthase from potato (*Solanum tuberosum* L.) that is predominantly expressed in leaf tissue. *Planta* **208**, 503-511.
- Kötting, O., Pusch, K., Tiessen, A., Geigenberger, P., Steup, M., and Ritte, G.** (2005). Identification of a novel enzyme required for starch metabolism in Arabidopsis leaves. The phosphoglucan, water dikinase. *Plant physiology* **137**, 242-252.
- Kötting, O., Santelia, D., Edner, C., Eicke, S., Marthaler, T., Gentry, M.S., Comparot-Moss, S., Chen, J., Smith, A.M., and Steup, M.** (2009). STARCH-EXCESS4 is a laforin-like phosphoglucan phosphatase required for starch degradation in Arabidopsis thaliana. *The Plant Cell* **21**, 334-346.
- Kubo, A., Fujita, N., Harada, K., Matsuda, T., Satoh, H., and Nakamura, Y.** (1999). The starch-debranching enzymes isoamylase and pullulanase are both involved in amylopectin biosynthesis in rice endosperm. *Plant physiology* **121**, 399-410.
- Kubo, A., Colleoni, C., Dinges, J.R., Lin, Q., Lappe, R.R., Rivenbark, J.G., Meyer, A.J., Ball, S.G., James, M.G., and Hennen-Bierwagen, T.A.** (2010). Functions of heteromeric and homomeric isoamylase-type starch-debranching enzymes in developing maize endosperm. *Plant physiology* **153**, 956-969.
- Kulak, N.A., Pichler, G., Paron, I., Nagaraj, N., and Mann, M.** (2014). Minimal, encapsulated proteomic-sample processing applied to copy-number estimation in eukaryotic cells. *Nature Methods* **11**, 319.
- Larsson, C.-T., Hofvander, P., Khoshnoodi, J., Ek, B., Rask, L., and Larsson, H.** (1996). Three isoforms of starch synthase and two isoforms of branching enzyme are present in potato tuber starch. *Plant Science* **117**, 9-16.
- Leterrier, M., Holappa, L.D., Broglie, K.E., and Beckles, D.M.** (2008). Cloning, characterisation and comparative analysis of a starch synthase IV gene in wheat: functional and evolutionary implications. *BMC Plant Biology* **8**, 98.
- Letunic, I., and Bork, P.** (2017). 20 years of the SMART protein domain annotation resource. *Nucleic Acids Research* **46**, D493-D496.
- Li, Z., Rahman, S., Kosar-Hashemi, B., Mouille, G., Appels, R., and Morell, M.** (1999a). Cloning and characterization of a gene encoding wheat starch synthase I. *Theoretical and Applied Genetics* **98**, 1208-1216.
- Li, Z., Chu, X., Mouille, G., Yan, L., Kosar-Hashemi, B., Hey, S., Napier, J., Shewry, P., Clarke, B., and Appels, R.** (1999b). The localization and expression of the class II starch synthases of wheat. *Plant physiology* **120**, 1147-1156.
- Lin, D.-G., and Jeang, C.-L.** (2005). Cloning, expression, and characterization of soluble starch synthase I cDNA from taro (*Colocasia esculenta* var. *esculenta*). *Journal of agricultural and food chemistry* **53**, 7985-7990.

- Lin, L., Guo, K., Zhang, L., Zhang, C., Liu, Q., and Wei, C.** (2019). Effects of molecular compositions on crystalline structure and functional properties of rice starches with different amylopectin extra-long chains. *Food Hydrocolloids* **88**, 137-145.
- Liu, F., Makhmoudova, A., Lee, E.A., Wait, R., Emes, M.J., and Tetlow, I.J.** (2009). The amylose extender mutant of maize conditions novel protein–protein interactions between starch biosynthetic enzymes in amyloplasts. *Journal of Experimental Botany* **60**, 4423-4440.
- Liu, H., Yu, G., Wei, B., Wang, Y., Zhang, J., Hu, Y., Liu, Y., Yu, G., Zhang, H., and Huang, Y.** (2015). Identification and phylogenetic analysis of a novel starch synthase in maize. *Frontiers in plant science* **6**, 1013.
- Lloyd, J.R., Blennow, A., Burhenne, K., and Kossmann, J.** (2004). Repression of a novel isoform of disproportionating enzyme (stDPE2) in potato leads to inhibition of starch degradation in leaves but not tubers stored at low temperature. *Plant Physiol* **134**, 1347-1354.
- Lohmeier-Vogel, E.M., Kerk, D., Nimick, M., Wrobel, S., Vickerman, L., Muench, D.G., and Moorhead, G.B.** (2008). Arabidopsis At5g39790 encodes a chloroplast-localized, carbohydrate-binding, coiled-coil domain-containing putative scaffold protein. *BMC plant biology* **8**, 120.
- Lombard, V., Golaconda Ramulu, H., Drula, E., Coutinho, P.M., and Henrissat, B.** (2014). The carbohydrate-active enzymes database (CAZy) in 2013. *Nucleic acids research* **42**, D490-D495.
- López, J.L., Marina, A., Álvarez, G., and Vázquez, J.** (2002). Application of proteomics for fast identification of species-specific peptides from marine species. *Proteomics* **2**, 1658-1665.
- Lorberth, R., Ritte, G., Willmitzer, L., and Kossmann, J.** (1998). Inhibition of a starch-granule-bound protein leads to modified starch and repression of cold sweetening. *Nature Biotechnology* **16**, 473-477.
- Lu, K.-J., Pfister, B., Jenny, C., Eicke, S., and Zeeman, S.C.** (2018). Distinct Functions of STARCH SYNTHASE 4 Domains in Starch Granule Formation. *Plant Physiology* **176**, 566-581.
- Luo, W., Zhao, Z., Pan, H., Zhao, L., Xu, C., and Yu, X.** (2018). Feasibility of butanol production from wheat starch wastewater by *Clostridium acetobutylicum*. *Energy* **154**, 240-248.
- Ma, U.V.L., Floros, J.D., and Ziegler, G.R.** (2011). Effect of starch fractions on spherulite formation and microstructure. *Carbohydrate Polymers* **83**, 1757-1765.
- Mahlow, S., Orzechowski, S., and Fettke, J.** (2016). Starch phosphorylation: insights and perspectives. *Cellular and molecular life sciences* **73**, 2753-2764.
- Malinova, I., Mahto, H., Brandt, F., AL-Rawi, S., Qasim, H., Brust, H., Hejazi, M., and Fettke, J.** (2018). EARLY STARVATION1 specifically affects the phosphorylation action of starch-related dikinases. *The Plant Journal* **95**, 126-137.

- Marmuse, L., Nepogodiev, S.A., and Field, R.A.** (2005). "Click chemistry" en route to pseudo-starch. *Organic & Biomolecular Chemistry* **3**, 2225-2227.
- Marshall, J., Sidebottom, C., Debet, M., Martin, C., Smith, A.M., and Edwards, A.** (1996). Identification of the major starch synthase in the soluble fraction of potato tubers. *The Plant Cell* **8**, 1121-1135.
- McDonnell, A.V., Jiang, T., Keating, A.E., and Berger, B.** (2006). Paircoil2: improved prediction of coiled coils from sequence. *Bioinformatics* **22**, 356-358.
- Meekins, D.A., Guo, H.-F., Husodo, S., Paasch, B.C., Bridges, T.M., Santelia, D., Kötting, O., Vander Kooi, C.W., and Gentry, M.S.** (2013). Structure of the *Arabidopsis* glucan phosphatase like sex four2 reveals a unique mechanism for starch dephosphorylation. *The Plant Cell* **25**, 2302-2314.
- Mérida, A., Rodríguez-Galán, J.M., Vincent, C., and Romero, J.M.** (1999). Expression of the granule-bound starch synthase I (Waxy) gene from snapdragon is developmentally and circadian clock regulated. *Plant physiology* **120**, 401-410.
- Mikkelsen, R., and Blennow, A.** (2005). Functional domain organization of the potato  $\alpha$ -glucan, water dikinase (GWD): evidence for separate site catalysis as revealed by limited proteolysis and deletion mutants. *Biochemical Journal* **385**, 355-361.
- Mikkelsen, R., Baunsgaard, L., and Blennow, A.** (2004). Functional characterization of alpha-glucan, water dikinase, the starch phosphorylating enzyme. *Biochemical Journal* **377**, 525-532.
- Mizuno, K., Kawasaki, T., Shimada, H., Satoh, H., Kobayashi, E., Okumura, S., Arai, Y., and Baba, T.** (1993). Alteration of the structural properties of starch components by the lack of an isoform of starch branching enzyme in rice seeds. *Journal of Biological Chemistry* **268**, 19084-19091.
- Møller, M.S., and Svensson, B.** (2016). Structural biology of starch-degrading enzymes and their regulation. *Current opinion in structural biology* **40**, 33-42.
- Monroe, J.D.** (2020). Involvement of five catalytically active *Arabidopsis*  $\beta$ -amylases in leaf starch metabolism and plant growth. *Plant Direct* **4**, e00199.
- Monroe, J.D., and Storm, A.R.** (2018). The *Arabidopsis*  $\beta$ -amylase (BAM) gene family: diversity of form and function. *Plant Science* **276**, 163-170.
- Monroe, J.D., Breault, J.S., Pope, L.E., Torres, C.E., Gebrejesus, T.B., Berndsen, C.E., and Storm, A.R.** (2017). *Arabidopsis*  $\beta$ -Amylase2 Is a  $K^+$ -Requiring, Catalytic Tetramer with Sigmoidal Kinetics. *Plant Physiology* **175**, 1525-1535.
- Morell, M.K., Blennow, A., Kosar-Hashemi, B., and Samuel, M.S.** (1997). Differential expression and properties of starch branching enzyme isoforms in developing wheat endosperm. *Plant Physiology* **113**, 201-208.
- Morell, M.K., Kosar-Hashemi, B., Cmiel, M., Samuel, M.S., Chandler, P., Rahman, S., Buleon, A., Batey, I.L., and Li, Z.** (2003). Barley sex6 mutants lack starch synthase IIa activity and contain a starch with novel properties. *The Plant Journal* **34**, 173-185.
- Mori, H., Tanizawa, K., and Fukui, T.** (1991). Potato tuber type H phosphorylase isozyme. Molecular cloning, nucleotide sequence, and expression of a full-



- length cDNA in *Escherichia coli*. *Journal of Biological Chemistry* **266**, 18446-18453.
- Morrison, W.R., Milligan, T.P., and Azudin, M.N.** (1984). A relationship between the amylose and lipid contents of starches from diploid cereals. *Journal of Cereal Science* **2**, 257-271.
- Mouille, G., Maddelein, M.-L., Libessart, N., Talaga, P., Decq, A., Delrue, B., and Ball, S.** (1996). Preamylopectin processing: a mandatory step for starch biosynthesis in plants. *The Plant Cell* **8**, 1353-1366.
- Mu-Forster, C., Huang, R., Powers, J.R., Harriman, R.W., Knight, M., Singletary, G.W., Keeling, P.L., and Wasserman, B.P.** (1996). Physical association of starch biosynthetic enzymes with starch granules of maize endosperm (granule-associated forms of starch synthase I and starch branching enzyme II). *Plant Physiology* **111**, 821-829.
- Mu, C., Harn, C., Ko, Y.T., Singletary, G.W., Keeling, P.L., and Wasserman, B.P.** (1994). Association of a 76 kDa polypeptide with soluble starch synthase I activity in maize (cv B73) endosperm. *The Plant Journal* **6**, 151-159.
- Mugford, S.T., Fernandez, O., Brinton, J., Flis, A., Krohn, N., Encke, B., Feil, R., Sulpice, R., Lunn, J.E., Stitt, M., and Smith, A.M.** (2014). Regulatory Properties of ADP Glucose Pyrophosphorylase Are Required for Adjustment of Leaf Starch Synthesis in Different Photoperiods. *Plant Physiology* **166**, 1733-1747.
- Muscat, D., Adhikari, B., Adhikari, R., and Chaudhary, D.S.** (2012). Comparative study of film forming behaviour of low and high amylose starches using glycerol and xylitol as plasticizers. *Journal of Food Engineering* **109**, 189-201.
- Mutisya, J., Sathish, P., Sun, C., Andersson, L., Ahlandsberg, S., Baguma, Y., Palmqvist, S., Odhiambo, B., Åman, P., and Jansson, C.** (2003). Starch branching enzymes in sorghum (*Sorghum bicolor*) and barley (*Hordeum vulgare*): comparative analyses of enzyme structure and gene expression. *Journal of plant physiology* **160**, 921-930.
- Nakamura, Y.** (2002). Towards a better understanding of the metabolic system for amylopectin biosynthesis in plants: rice endosperm as a model tissue. *Plant and Cell Physiology* **43**, 718-725.
- Nakamura, Y., Utsumi, Y., Sawada, T., Aihara, S., Utsumi, C., Yoshida, M., and Kitamura, S.** (2010). Characterization of the reactions of starch branching enzymes from rice endosperm. *Plant and cell physiology* **51**, 776-794.
- Nepogodiev, S.A., Dedola, S., Marmuse, L., de Oliveira, M.T., and Field, R.A.** (2007). Synthesis of triazole-linked pseudo-starch fragments. *Carbohydrate Research* **342**, 529-540.
- Niittylä, T., Messerli, G., Trevisan, M., Chen, J., Smith, A.M., and Zeeman, S.C.** (2004). A previously unknown maltose transporter essential for starch degradation in leaves. *Science* **303**, 87-89.
- Niittylä, T., Comparot-Moss, S., Lue, W.-L., Messerli, G., Trevisan, M., Seymour, M.D., Gatehouse, J.A., Villadsen, D., Smith, S.M., and Chen, J.** (2006). Similar

- protein phosphatases control starch metabolism in plants and glycogen metabolism in mammals. *Journal of Biological Chemistry* **281**, 11815-11818.
- Nishi, A., Nakamura, Y., Tanaka, N., and Satoh, H.** (2001). Biochemical and genetic analysis of the effects of amylose-extender mutation in rice endosperm. *Plant physiology* **127**, 459-472.
- Noctor, G., Mhamdi, A., Chaouch, S., Han, Y., Neukermans, J., Marquez-Garcia, B., Queval, G., and Foyer, C.H.** (2012). Glutathione in plants: an integrated overview. *Plant Cell Environ* **35**, 454-484.
- O'Farrell, P.H.** (1975). High resolution two-dimensional electrophoresis of proteins. *J Biol Chem* **250**, 4007-4021.
- O'Neill, E.C., and Field, R.A.** (2015). Underpinning starch biology with in vitro studies on carbohydrate-active enzymes and biosynthetic glycomaterials. *Frontiers in Bioengineering and Biotechnology* **3**.
- O'Neill, E.C., and Field, R.A.** (2015). Underpinning Starch Biology with in vitro Studies on Carbohydrate-Active Enzymes and Biosynthetic Glycomaterials. *Frontiers in Bioengineering and Biotechnology* **3**.
- Okagaki, R.J., and Wessler, S.R.** (1988). Comparison of non-mutant and mutant waxy genes in rice and maize. *Genetics* **120**, 1137-1143.
- Ong, S.E., Blagoev, B., Kratchmarova, I., Kristensen, D.B., Steen, H., Pandey, A., and Mann, M.** (2002). Stable isotope labeling by amino acids in cell culture, SILAC, as a simple and accurate approach to expression proteomics. *Mol Cell Proteomics* **1**, 376-386.
- Oostergetel, G.T., and van Bruggen, E.F.** (1993). The crystalline domains in potato starch granules are arranged in a helical fashion. *Carbohydrate Polymers* **21**, 7-12.
- Peat, S., Whelan, W.J., and Rees, W.** (1956). 10. The enzymic synthesis and degradation of starch. Part XX. The disproportionating enzyme (D-enzyme) of the potato. *Journal of the Chemical Society (Resumed)*, 44-53.
- Peng, C., Wang, Y., Liu, F., Ren, Y., Zhou, K., Lv, J., Zheng, M., Zhao, S., Zhang, L., and Wang, C.** (2014). FLOURY ENDOSPERM 6 encodes a CBM 48 domain-containing protein involved in compound granule formation and starch synthesis in rice endosperm. *The Plant Journal* **77**, 917-930.
- Peng, M., Gao, M., Baga, M., Hucl, P., and Chibbar, R.N.** (2000). Starch-Branching Enzymes Preferentially Associated with A-Type Starch Granules in Wheat Endosperm. *Plant Physiology* **124**, 265-272.
- Pérez, S., and Bertoft, E.** (2010). The molecular structures of starch components and their contribution to the architecture of starch granules: A comprehensive review. *Starch-Stärke* **62**, 389-420.
- Pfister, B., and Zeeman, S.C.** (2016). Formation of starch in plant cells. *Cellular and Molecular Life Sciences* **73**, 2781-2807.
- Pfister, B., Sánchez-Ferrer, A., Diaz, A., Lu, K., Otto, C., Holler, M., Shaik, F.R., Meier, F., Mezzenga, R., and Zeeman, S.C.** (2016). Recreating the synthesis of starch granules in yeast. *eLife* **5**, e15552.

- Phan, T.D., Debeaufort, F., Luu, D., and Voilley, A.** (2005). Functional Properties of Edible Agar-Based and Starch-Based Films for Food Quality Preservation. *Journal of Agricultural and Food Chemistry* **53**, 973-981.
- Piperno, D.R.** (2011). The Origins of Plant Cultivation and Domestication in the New World Tropics: Patterns, Process, and New Developments. *Current Anthropology* **52**, S453-S470.
- Piperno, D.R., and Dillehay, T.D.** (2008). Starch grains on human teeth reveal early broad crop diet in northern Peru. *Proceedings of the National Academy of Sciences* **105**, 19622-19627.
- Pirone, C., Gurrieri, L., Gaiba, I., Adamiano, A., Valle, F., Trost, P., and Sparla, F.** (2017). The analysis of the different functions of starch-phosphorylating enzymes during the development of *Arabidopsis thaliana* plants discloses an unexpected role for the cytosolic isoform GWD2. *Physiologia Plantarum* **160**, 447-457.
- Poltronieri, P., Liu, S., Cimaglia, F., Santino, A., and Wang, Y.** (2012). Characterization of Kunitz-type inhibitor B1 performance using protein chips and AFM. *Sensors and Actuators B: Chemical* **168**, 231-237.
- Poulsen, P., and Kreiberg, J.D.** (1993). Starch branching enzyme cDNA from *Solanum tuberosum*. *Plant physiology* **102**, 1053.
- Putaux, J.-L., Potocki-Véronèse, G., Remaud-Simeon, M., and Buleon, A.** (2006).  $\alpha$ -D-Glucan-Based Dendritic Nanoparticles Prepared by in Vitro Enzymatic Chain Extension of Glycogen. *Biomacromolecules* **7**, 1720-1728.
- Ragel, P., Streb, S., Feil, R., Sahrawy, M., Annunziata, M.G., Lunn, J.E., Zeeman, S., and Mérida, Á.** (2013). Loss of starch granule initiation has a deleterious effect on the growth of *Arabidopsis* plants due to an accumulation of ADP-glucose. *Plant Physiology* **163**, 75-85.
- Rahman, S., Regina, A., Li, Z., Mukai, Y., Yamamoto, M., Kosar-Hashemi, B., Abrahams, S., and Morell, M.K.** (2001). Comparison of starch-branching enzyme genes reveals evolutionary relationships among isoforms. characterization of a gene for starch-branching enzyme IIa from the wheat d genome Donor *Aegilops tauschii*. *Plant Physiology* **125**, 1314-1324.
- Rahman, S., Kosar-Hashemi, B., Samuel, M., Hill, A., Abbott, D., Skerritt, J., Preiss, J., Appels, R., and Morell, M.** (1995). The major proteins of wheat endosperm starch granules. *Functional Plant Biology* **22**, 793-803.
- Rapaille, A., and Vanhemelrijck, J.** (1997). Modified starches. In *Thickening and Gelling Agents for Food*, A.P. Imeson, ed (Boston, MA: Springer US), pp. 199-229.
- Regina, A., Kosar-Hashemi, B., Li, Z., Pedler, A., Mukai, Y., Yamamoto, M., Gale, K., Sharp, P.J., Morell, M.K., and Rahman, S.** (2005). Starch branching enzyme IIb in wheat is expressed at low levels in the endosperm compared to other cereals and encoded at a non-syntenic locus. *Planta* **222**, 899.
- Regina, A., Bird, A., Topping, D., Bowden, S., Freeman, J., Barsby, T., Kosar-Hashemi, B., Li, Z., Rahman, S., and Morell, M.** (2006). High-amylose wheat

- generated by RNA interference improves indices of large-bowel health in rats. *Proceedings of the National Academy of Sciences* **103**, 3546-3551.
- Reinhold, H., Soyk, S., Šimková, K., Hostettler, C., Marafino, J., Mainiero, S., Vaughan, C.K., Monroe, J.D., and Zeeman, S.C.** (2011).  $\beta$ -Amylase-Like Proteins Function as Transcription Factors in *Arabidopsis*, Controlling Shoot Growth and Development. *The Plant Cell* **23**, 1391-1403.
- Ridout, M.J., Parker, M.L., Hedley, C.L., Bogracheva, T.Y., and Morris, V.J.** (2003). Atomic force microscopy of pea starch granules: granule architecture of wild-type parent, r and rb single mutants, and the rrb double mutant. *Carbohydrate research* **338**, 2135-2147.
- Ritte, G., Lloyd, J.R., Eckermann, N., Rottmann, A., Kossmann, J., and Steup, M.** (2002). The starch-related R1 protein is an  $\alpha$ -glucan, water dikinase. *Proceedings of the National Academy of Sciences* **99**, 7166-7171.
- Ritte, G., Heydenreich, M., Mahlow, S., Haebel, S., Kötting, O., and Steup, M.** (2006). Phosphorylation of C6- and C3-positions of glucosyl residues in starch is catalysed by distinct dikinases. *FEBS Letters* **580**, 4872-4876.
- Robin, J.** (1974). Lint-nerized starches. Gel filtration and enzymatic studies of insoluble residues from prolonged acid treatment of potato starch. *Cereal Chem.* **51**, 389-406.
- Roldán, I., Wattebled, F., Mercedes Lucas, M., Delvallé, D., Planchot, V., Jiménez, S., Pérez, R., Ball, S., d'Hulst, C., and Mérida, Á.** (2007). The phenotype of soluble starch synthase IV defective mutants of *Arabidopsis thaliana* suggests a novel function of elongation enzymes in the control of starch granule formation. *The Plant Journal* **49**, 492-504.
- Rydberg, U., Andersson, L., Andersson, R., Åman, P., and Larsson, H.** (2001). Comparison of starch branching enzyme I and II from potato. *European Journal of Biochemistry* **268**, 6140-6145.
- Santelia, D., and Lunn, J.E.** (2017). Transitory Starch Metabolism in Guard Cells: Unique Features for a Unique Function. *Plant Physiol* **174**, 539-549.
- Santelia, D., Kötting, O., Seung, D., Schubert, M., Thalmann, M., Bischof, S., Meekins, D.A., Lutz, A., Patron, N., and Gentry, M.S.** (2011). The phosphoglucan phosphatase like sex Four2 dephosphorylates starch at the C3-position in *Arabidopsis*. *The Plant Cell* **23**, 4096-4111.
- Satoh, H., Nishi, A., Yamashita, K., Takemoto, Y., Tanaka, Y., Hosaka, Y., Sakurai, A., Fujita, N., and Nakamura, Y.** (2003). Starch-branching enzyme I-deficient mutation specifically affects the structure and properties of starch in rice endosperm. *Plant physiology* **133**, 1111-1121.
- Sauer, M.-L.A., Xu, B., and Sutton, F.** (2014). Metabolic labeling with stable isotope nitrogen ( $^{15}\text{N}$ ) to follow amino acid and protein turnover of three plastid proteins in *Chlamydomonas reinhardtii*. *Proteome Science* **12**, 14.
- Sawada, T., Itoh, M., and Nakamura, Y.** (2018). Contributions of three starch branching enzyme isozymes to the fine structure of amylopectin in rice endosperm. *Frontiers in plant science* **9**, 1536.

- Sawada, T., Nakamura, Y., Ohdan, T., Saitoh, A., Francisco Jr, P.B., Suzuki, E., Fujita, N., Shimonaga, T., Fujiwara, S., and Tsuzuki, M.** (2014). Diversity of reaction characteristics of glucan branching enzymes and the fine structure of  $\alpha$ -glucan from various sources. *Archives of biochemistry and biophysics* **562**, 9-21.
- Scheidig, A., Fröhlich, A., Schulze, S., Lloyd, J.R., and Kossmann, J.** (2002). Downregulation of a chloroplast-targeted  $\beta$ -amylase leads to a starch-excess phenotype in leaves. *The Plant Journal* **30**, 581-591.
- Schliebner, I., Pribil, M., Zuhlke, J., Dietzmann, A., and Leister, D.** (2008). A survey of chloroplast protein kinases and phosphatases in *Arabidopsis thaliana*. *Current genomics* **9**, 184-190.
- Schreier, T.B., Umhang, M., Lee, S.-K., Lue, W.-L., Shen, Z., Silver, D., Graf, A., Müller, A., Eicke, S., Stadler-Waibel, M., Seung, D., Bischof, S., Briggs, S.P., Kötting, O., Moorhead, G.B.G., Chen, J., and Zeeman, S.C.** (2019). LIKE SEX4 1 Acts as a  $\beta$ -Amylase-Binding Scaffold on Starch Granules during Starch Degradation. *The Plant Cell* **31**, 2169-2186.
- Senoura, T., Asao, A., Takashima, Y., Isono, N., Hamada, S., Ito, H., and Matsui, H.** (2007). Enzymatic characterization of starch synthase III from kidney bean (*Phaseolus vulgaris* L.). *The FEBS journal* **274**, 4550-4560.
- Seo, B.-s., Kim, S., Scott, M.P., Singletary, G.W., Wong, K.-s., James, M.G., and Myers, A.M.** (2002). Functional interactions between heterologously expressed starch-branching enzymes of maize and the glycogen synthases of Brewer's yeast. *Plant physiology* **128**, 1189-1199.
- Seung, D., Schreier, T.B., Bürgy, L., Eicke, S., and Zeeman, S.C.** (2018). Two Plastidial Coiled-Coil Proteins Are Essential for Normal Starch Granule Initiation in *Arabidopsis*. *The Plant Cell* **30**, 1523-1542.
- Seung, D., Soyk, S., Coiro, M., Maier, B.A., Eicke, S., and Zeeman, S.C.** (2015). PROTEIN TARGETING TO STARCH Is Required for Localising GRANULE-BOUND STARCH SYNTHASE to Starch Granules and for Normal Amylose Synthesis in *Arabidopsis*. *PLOS Biology* **13**, e1002080.
- Seung, D., Boudet, J., Monroe, J., Schreier, T.B., David, L.C., Abt, M., Lu, K.-J., Zanella, M., and Zeeman, S.C.** (2017). Homologs of PROTEIN TARGETING TO STARCH Control Starch Granule Initiation in *Arabidopsis* Leaves. *The Plant Cell* **29**, 1657-1677.
- Sevestre, F., Facon, M., Wattedled, F., and Szydlowski, N.** (2020). Facilitating gene editing in potato: a Single-Nucleotide Polymorphism (SNP) map of the *Solanum tuberosum* L. cv. Desiree genome. *Scientific Reports* **10**, 1-8.
- Skryhan, K., Cuesta-Seijo, J.A., Nielsen, M.M., Marri, L., Mellor, S.B., Glaring, M.A., Jensen, P.E., Palcic, M.M., and Blennow, A.** (2015). The role of cysteine residues in redox regulation and protein stability of *Arabidopsis thaliana* starch synthase 1. *PloS one* **10**.
- Sonnewald, U., Basner, A., Greve, B., and Steup, M.** (1995). A second L-type isozyme of potato glucan phosphorylase: cloning, antisense inhibition and expression analysis. *Plant Molecular Biology* **27**, 567-576.

- Steidle, E.A.** (2010). Investigation of the role of BAM9 in starch metabolism in *Arabidopsis thaliana*.
- Stensballe, A., Hald, S., Bauw, G., Blennow, A., and Welinder, K.G.** (2008). The amyloplast proteome of potato tuber. *The FEBS journal* **275**, 1723-1741.
- Stettler, M., Eicke, S., Mettler, T., Messerli, G., Hörtensteiner, S., and Zeeman, S.C.** (2009). Blocking the metabolism of starch breakdown products in *Arabidopsis* leaves triggers chloroplast degradation. *Molecular Plant* **2**, 1233-1246.
- Steup, M.** (1981). Purification of chloroplast  $\alpha$ -1, 4-glucan phosphorylase from spinach leaves by chromatography on sepharose-bound starch. *Biochimica et Biophysica Acta (BBA)-Enzymology* **659**, 123-131.
- Steup, M., and Latzko, E.** (1979). Intracellular localization of phosphorylases in spinach and pea leaves. *Planta* **145**, 69-75.
- Steup, M., Peavey, D.G., and Gibbs, M.** (1976). The regulation of starch metabolism by inorganic phosphate. *Biochemical and biophysical research communications* **72**, 1554-1561.
- Streb, S., Delatte, T., Umhang, M., Eicke, S., Schorderet, M., Reinhardt, D., and Zeeman, S.C.** (2008). Starch granule biosynthesis in *Arabidopsis* is abolished by removal of all debranching enzymes but restored by the subsequent removal of an endoamylase. *The Plant Cell* **20**, 3448-3466.
- Szabo de Edelenyi, F., Julia, C., Courtois, F., Méjean, C., Péneau, S., Galan, P., Hercberg, S., and Kesse-Guyot, E.** (2014). Starchy Food Consumption in French Adults: A Cross-Sectional Analysis of the Profile of Consumers and Contribution to Nutritional Intake in a Web-Based Prospective Cohort. *Annals of Nutrition and Metabolism* **64**, 28-37.
- Szydlowski, N., Ragel, P., Hennen-Bierwagen, T.A., Planchot, V., Myers, A.M., Merida, A., d'Hulst, C., and Wattedled, F.** (2011). Integrated functions among multiple starch synthases determine both amylopectin chain length and branch linkage location in *Arabidopsis* leaf starch. *J Exp Bot* **62**, 4547-4559.
- Szydlowski, N., Ragel, P., Raynaud, S., Lucas, M.M., Roldán, I., Montero, M., Muñoz, F.J., Ovecka, M., Bahaji, A., Planchot, V., Pozueta-Romero, J., D'Hulst, C., and Mérida, Á.** (2009). Starch Granule Initiation in *Arabidopsis* Requires the Presence of Either Class IV or Class III Starch Synthases. *The Plant Cell* **21**, 2443-2457.
- Tai, H.H., Lagüe, M., Thomson, S., Aourousseau, F., Neilson, J., Murphy, A., Bizimungu, B., Davidson, C., Deveaux, V., Bègue, Y., Wang, H.Y., Xiong, X., and Jacobs, J.M.E.** (2020). Tuber transcriptome profiling of eight potato cultivars with different cold-induced sweetening responses to cold storage. *Plant Physiology and Biochemistry* **146**, 163-176.
- Takaha, T., Yanase, M., Okada, S., and Smith, S.M.** (1993). Disproportionating enzyme (4- $\alpha$ -glucanotransferase; EC 2.4.1.25) of potato. Purification, molecular cloning, and potential role in starch metabolism. *J Biol Chem* **268**, 1391-1396.

- Takeda, Y.** (1982). Location of phosphate groups in potato amylopectin. *Carbohydr. Res.* **102**, 321-327.
- Takeda, Y., Guan, H.-P., and Preiss, J.** (1993). Branching of amylose by the branching isoenzymes of maize endosperm. *Carbohydrate research* **240**, 253-263.
- Takeda, Y., Hizukuri, S., Takeda, C., and Suzuki, A.** (1987). Structures of branched molecules of amyloses of various origins, and molar fractions of branched and unbranched molecules. *Carbohydrate Research* **165**, 139-145.
- Tanaka, E., Ral, J.-P.F., Li, S., Gaire, R., Cavanagh, C.R., Cullis, B.R., and Whan, A.** (2017). Increased accuracy of starch granule type quantification using mixture distributions. *Plant methods* **13**, 107.
- Tester, R., Karkalas, J., and Qi, X.** (2004). Starch structure and digestibility enzyme-substrate relationship. *World's Poultry Science Journal* **60**, 186-195.
- Tetlow, I.J., Wait, R., Lu, Z., Akkasaeng, R., Bowsher, C.G., Esposito, S., Kosar-Hashemi, B., Morell, M.K., and Emes, M.J.** (2004). Protein phosphorylation in amyloplasts regulates starch branching enzyme activity and protein-protein interactions. *The Plant Cell* **16**, 694-708.
- Tetlow, I.J., Beisel, K.G., Cameron, S., Makhmoudova, A., Liu, F., Bresolin, N.S., Wait, R., Morell, M.K., and Emes, M.J.** (2008). Analysis of protein complexes in wheat amyloplasts reveals functional interactions among starch biosynthetic enzymes. *Plant physiology* **146**, 1878-1891.
- Thalmann, M., Coiro, M., Meier, T., Wicker, T., Zeeman, S.C., and Santelia, D.** (2019). The evolution of functional complexity within the  $\beta$ -amylase gene family in land plants. *BMC evolutionary biology* **19**, 66.
- Thalmann, M., Pazmino, D., Seung, D., Horrer, D., Nigro, A., Meier, T., Kölling, K., Pfeifhofer, H.W., Zeeman, S.C., and Santelia, D.** (2016). Regulation of Leaf Starch Degradation by Abscisic Acid Is Important for Osmotic Stress Tolerance in Plants. *Plant Cell* **28**, 1860-1878.
- Toyosawa, Y., Kawagoe, Y., Matsushima, R., Crofts, N., Ogawa, M., Fukuda, M., Kumamaru, T., Okazaki, Y., Kusano, M., Saito, K., Toyooka, K., Sato, M., Ai, Y., Jane, J.-L., Nakamura, Y., and Fujita, N.** (2016). Deficiency of Starch Synthase IIIa and IVb Alters Starch Granule Morphology from Polyhedral to Spherical in Rice Endosperm. *Plant Physiology* **170**, 1255-1270.
- Tsai, C.-Y.** (1974). The function of the Waxy locus in starch synthesis in maize endosperm. *Biochemical Genetics* **11**, 83-96.
- Turra, D., and Lorito, M.** (2011). Potato type I and II proteinase inhibitors: modulating plant physiology and host resistance. *Current Protein and Peptide Science* **12**, 374-385.
- Tyson, R.H., and ap Rees, T.** (1988). Starch synthesis by isolated amyloplasts from wheat endosperm. *Planta* **175**, 33-38.
- Tyynelä, J., Stitt, M., Lönneborg, A., Smeekens, S., and Schulman, A.H.** (1995). Metabolism of starch synthesis in developing grains of the shx shrunken mutant of barley (*Hordeum vulgare*). *Physiologia Plantarum* **93**, 77-84.

- Umemoto, T., and Aoki, N.** (2005). Single-nucleotide polymorphisms in rice starch synthase IIa that alter starch gelatinisation and starch association of the enzyme. *Functional Plant Biology* **32**, 763-768.
- Umemoto, T., Yano, M., Satoh, H., Shomura, A., and Nakamura, Y.** (2002). Mapping of a gene responsible for the difference in amylopectin structure between japonica-type and indica-type rice varieties. *Theoretical and Applied Genetics* **104**, 1-8.
- Utsumi, Y., and Nakamura, Y.** (2006). Structural and enzymatic characterization of the isoamylase1 homo-oligomer and the isoamylase1–isoamylase2 hetero-oligomer from rice endosperm. *Planta* **225**, 75-87.
- Utsumi, Y., Utsumi, C., Sawada, T., Fujita, N., and Nakamura, Y.** (2011). Functional diversity of isoamylase oligomers: the ISA1 homo-oligomer is essential for amylopectin biosynthesis in rice endosperm. *Plant Physiology* **156**, 61-77.
- Valdez, H.A., Busi, M.V., Wayllace, N.Z., Parisi, G., Ugalde, R.A., and Gomez-Casati, D.F.** (2008). Role of the N-terminal starch-binding domains in the kinetic properties of starch synthase III from *Arabidopsis thaliana*. *Biochemistry* **47**, 3026-3032.
- Valerio, C., Costa, A., Marri, L., Issakidis-Bourguet, E., Pupillo, P., Trost, P., and Sparla, F.** (2011). Thioredoxin-regulated  $\beta$ -amylase (BAM1) triggers diurnal starch degradation in guard cells, and in mesophyll cells under osmotic stress. *Journal of experimental botany* **62**, 545-555.
- van de Wal, M., D'Hulst, C., Vincken, J.-P., Buléon, A., Visser, R., and Ball, S.** (1998). Amylose Is Synthesized in Vitro by Extension of and Cleavage from Amylopectin. *Journal of Biological Chemistry* **273**, 22232-22240.
- van der Leij, F., Visser, R., Oosterhaven, K., Van der Kop, D., Jacobsen, E., and Feenstra, W.** (1991a). Complementation of the amylose-free starch mutant of potato (*Solanum tuberosum*.) by the gene encoding granule-bound starch synthase. *Theoretical and applied genetics* **82**, 289-295.
- Van Der Leij, F.R., Visser, R.G., Ponstein, A.S., Jacobsen, E., and Feenstra, W.J.** (1991b). Sequence of the structural gene for granule-bound starch synthase of potato (*Solanum tuberosum* L.) and evidence for a single point deletion in the amf allele. *Molecular and General Genetics MGG* **228**, 240-248.
- Van Harselaar, J.K., Lorenz, J., Senning, M., Sonnewald, U., and Sonnewald, S.** (2017). Genome-wide analysis of starch metabolism genes in potato (*Solanum tuberosum* L.). *BMC genomics* **18**, 37.
- Vandenbrink, J.P., and Kiss, J.Z.** (2019). Plant responses to gravity. *Seminars in Cell & Developmental Biology* **92**, 122-125.
- Vandromme, C., Spriet, C., Dauvillée, D., Courseaux, A., Putaux, J.L., Wychowski, A., Krzewinski, F., Facon, M., d'Hulst, C., and Wattedled, F.** (2019). PII1: a protein involved in starch initiation that determines granule number and size in *Arabidopsis* chloroplast. *New Phytologist* **221**, 356-370.
- Veillet, F., Chauvin, L., Kermarrec, M.-P., Sevestre, F., Merrer, M., Terret, Z., Szydłowski, N., Devaux, P., Gallois, J.-L., and Chauvin, J.-E.** (2019). The



- Solanum tuberosum* GBSSI gene: a target for assessing gene and base editing in tetraploid potato. *Plant cell reports* **38**, 1065-1080.
- Verbeke, J., Penverne, C., D'Hulst, C., Rolando, C., and Szydlowski, N.** (2016). Rapid and sensitive quantification of C3- and C6-phosphoesters in starch by fluorescence-assisted capillary electrophoresis. *Carbohydr Polym* **152**, 784-791.
- Wang, C., Tang, C.-H., Fu, X., Huang, Q., and Zhang, B.** (2016). Granular size of potato starch affects structural properties, octenylsuccinic anhydride modification and flowability. *Food chemistry* **212**, 453-459.
- Wang, Q., Monroe, J., and Sjolund, R.D.** (1995). Identification and Characterization of a Phloem-Specific [ $\beta$ ]-Amylase. *Plant Physiology* **109**, 743-750.
- Wang, X., Xue, L., Sun, J., and Zuo, J.** (2010). The Arabidopsis BE1 gene, encoding a putative glycoside hydrolase localized in plastids, plays crucial roles during embryogenesis and carbohydrate metabolism. *Journal of integrative plant biology* **52**, 273-288.
- Wattebled, F., Planchot, V., Dong, Y., Szydlowski, N., Pontoire, B., Devin, A., Ball, S., and d'Hulst, C.** (2008). Further evidence for the mandatory nature of polysaccharide debranching for the aggregation of semicrystalline starch and for overlapping functions of debranching enzymes in Arabidopsis leaves. *Plant physiology* **148**, 1309-1323.
- Wattebled, F., Ral, J.P., Dauvillée, D., Myers, A.M., James, M.G., Schlichting, R., Giersch, C., Ball, S.G., and D'Hulst, C.** (2003). STA11, a *Chlamydomonas reinhardtii* locus required for normal starch granule biogenesis, encodes disproportionating enzyme. Further evidence for a function of  $\alpha$ -1,4 glucanotransferases during starch granule biosynthesis in green algae. *Plant Physiol* **132**, 137-145.
- Wattebled, F., Dong, Y., Dumez, S., Delvallé, D., Planchot, V., Berbezy, P., Vyas, D., Colonna, P., Chatterjee, M., Ball, S., and D'Hulst, C.** (2005). Mutants of Arabidopsis lacking a chloroplastic isoamylase accumulate phytoglycogen and an abnormal form of amylopectin. *Plant Physiol* **138**, 184-195.
- Wheeler, R.M.** (2017). Agriculture for space: People and places paving the way. *Open Agriculture* **2**, 14-32.
- Wilkins, M.R., Pasquali, C., Appel, R.D., Ou, K., Golaz, O., Sanchez, J.C., Yan, J.X., Gooley, A.A., Hughes, G., Humphery-Smith, I., Williams, K.L., and Hochstrasser, D.F.** (1996). From proteins to proteomes: large scale protein identification by two-dimensional electrophoresis and amino acid analysis. *Biotechnology (N Y)* **14**, 61-65.
- Wong, K.-S., Kubo, A., Jane, J.-I., Harada, K., Satoh, H., and Nakamura, Y.** (2003). Structures and properties of amylopectin and phytoglycogen in the endosperm of sugary-1 mutants of rice. *Journal of Cereal Science* **37**, 139-149.
- Wychofski, A., Bompard, C., Grimaud, F., Potocki-Véronèse, G., d'Hulst, C., Wattebled, F., and Roussel, X.** (2017). Biochemical characterization of Arabidopsis thaliana starch branching enzyme 2.2 reveals an enzymatic positive cooperativity. *Biochimie* **140**, 146-158.

- Xing, S., Meng, X., Zhou, L., Mujahid, H., Zhao, C., Zhang, Y., Wang, C., and Peng, Z. (2016). Proteome profile of starch granules purified from rice (*Oryza sativa*) endosperm. *PloS one* 11.
- Xu, X., Pan, S., Cheng, S., Zhang, B., Mu, D., Ni, P., Zhang, G., Yang, S., Li, R., Wang, J., Orjeda, G., Guzman, F., Torres, M., Lozano, R., Ponce, O., Martinez, D., De la Cruz, G., Chakrabarti, S.K., Patil, V.U., Skryabin, K.G., Kuznetsov, B.B., Ravin, N.V., Kolganova, T.V., Beletsky, A.V., Mardanov, A.V., Di Genova, A., Bolser, D.M., Martin, D.M.A., Li, G., Yang, Y., Kuang, H., Hu, Q., Xiong, X., Bishop, G.J., Sagredo, B., Mejía, N., Zagorski, W., Gromadka, R., Gawor, J., Szczesny, P., Huang, S., Zhang, Z., Liang, C., He, J., Li, Y., He, Y., Xu, J., Zhang, Y., Xie, B., Du, Y., Qu, D., Bonierbale, M., Ghislain, M., del Rosario Herrera, M., Giuliano, G., Pietrella, M., Perrotta, G., Facella, P., O'Brien, K., Feingold, S.E., Barreiro, L.E., Massa, G.A., Diambra, L., Whitty, B.R., Vaillancourt, B., Lin, H., Massa, A.N., Geoffroy, M., Lundback, S., DellaPenna, D., Robin Buell, C., Sharma, S.K., Marshall, D.F., Waugh, R., Bryan, G.J., Destefanis, M., Nagy, I., Milbourne, D., Thomson, S.J., Fiers, M., Jacobs, J.M.E., Nielsen, K.L., Sønderkær, M., Iovene, M., Torres, G.A., Jiang, J., Veilleux, R.E., Bachem, C.W.B., de Boer, J., Borm, T., Kloosterman, B., van Eck, H., Datema, E., te Lintel Hekkert, B., Goverse, A., van Ham, R.C.H.J., Visser, R.G.F., The Potato Genome Sequencing, C., The Potato Genome, C., Shenzhen, B.G.I., Cayetano Heredia, U., Central Potato Research, I., Centre Bioengineering, R.A.S., Cgr-Cmm, U.d.C., College of Life Sciences, U.o.D., High Technology Research Center, S.A.o.A.S., Huazhong Agriculture, U., Hunan Agricultural, U., Imperial College, L., Instituto de Investigaciones, A., Institute of, B., Biophysics, Institute of, V., Flowers, C.A.o.A.S., International Potato, C., Italian National Agency for New Technologies, E., Sustainable, D., Institute, J.C.V., Laboratorio de Agrobiotecnología, I.N.d.T.A., Laboratorio de Biología de Sistemas, U.N.d.L., Michigan State, U., Scottish Crop Research, I., Teagasc Crops Research, C., The New Zealand Institute for, P., Food Research, L., University of, A., University of, W., Virginia Polytechnic, I., State, U., Wageningen, U., and Research, C. (2011). Genome sequence and analysis of the tuber crop potato. *Nature* 475, 189-195.
- Xuan, W., Beeckman, T., and Xu, G. (2017). Plant nitrogen nutrition: sensing and signaling. *Current Opinion in Plant Biology* 39, 57-65.
- Yamaguchi, M., Kainuma, K., and French, D. (1979). Electron microscopic observations of waxy maize starch. *Journal of ultrastructure research* 69, 249-261.
- Yamashita, Y., and Monobe, K. (1971). Single crystals of amylose V complexes. III. Crystals with 81 helical configuration. *Journal of Polymer Science Part A-2: Polymer Physics* 9, 1471-1481.
- Yan, H.-B., Pan, X.-X., Jiang, H.-W., and Wu, G.-J. (2009). Comparison of the starch synthesis genes between maize and rice: copies, chromosome location and expression divergence. *Theoretical and applied genetics* 119, 815-825.

- Yanase, M., Takaha, T., and Kuriki, T.** (2007). Developing and Engineering Enzymes for Manufacturing Amylose. *Journal of Applied Glycoscience* **54**, 125-131.
- Yandeau-Nelson, M.D., Laurens, L., Shi, Z., Xia, H., Smith, A.M., and Guiltinan, M.J.** (2011). Starch-branching enzyme IIa is required for proper diurnal cycling of starch in leaves of maize. *Plant physiology* **156**, 479-490.
- Yoon, J.-W., and Lim, S.-T.** (2003). Molecular fractionation of starch by density-gradient ultracentrifugation. *Carbohydrate Research* **338**, 611-617.
- Yu, J., Chang, P.R., and Ma, X.** (2010). The preparation and properties of dialdehyde starch and thermoplastic dialdehyde starch. *Carbohydrate Polymers* **79**, 296-300.
- Yu, J.W., Choi, J.S., Upadhyaya, C.P., Kwon, S.O., Gururani, M.A., Nookaraju, A., Nam, J.H., Choi, C.W., Kim, S.I., Ajappala, H., Kim, H.S., Jeon, J.H., and Park, S.W.** (2012). Dynamic proteomic profile of potato tuber during its in vitro development. *Plant Sci* **195**, 1-9.
- Yu, T.-S., Zeeman, S.C., Thorneycroft, D., Fulton, D.C., Dunstan, H., Lue, W.-L., Hegemann, B., Tung, S.-Y., Umemoto, T., and Chapple, A.** (2005).  $\alpha$ -Amylase is not required for breakdown of transitory starch in Arabidopsis leaves. *Journal of Biological Chemistry* **280**, 9773-9779.
- Yu, T.S., Kofler, H., Häusler, R.E., Hille, D., Flügge, U.I., Zeeman, S.C., Smith, A.M., Kossmann, J., Lloyd, J., Ritte, G., Steup, M., Lue, W.L., Chen, J., and Weber, A.** (2001a). The Arabidopsis *sex1* mutant is defective in the R1 protein, a general regulator of starch degradation in plants, and not in the chloroplast hexose transporter. *Plant Cell* **13**, 1907-1918.
- Yu, Y., Mu, H.H., Wasserman, B.P., and Carman, G.M.** (2001b). Identification of the maize amyloplast stromal 112-kD protein as a plastidic starch phosphorylase. *Plant Physiology* **125**, 351-359.
- Yun, M.-S., Umemoto, T., and Kawagoe, Y.** (2011). Rice Debranching Enzyme Isoamylase3 Facilitates Starch Metabolism and Affects Plastid Morphogenesis. *Plant and Cell Physiology* **52**, 1068-1082.
- Yusuph, M., Tester, R.F., Ansell, R., and Snape, C.E.** (2003). Composition and properties of starches extracted from tubers of different potato varieties grown under the same environmental conditions. *Food Chemistry* **82**, 283-289.
- Zeeman, S.C., Northrop, F., Smith, A.M., and Rees, T.a.** (1998a). A starch-accumulating mutant of Arabidopsis thaliana deficient in a chloroplastic starch-hydrolysing enzyme. *The Plant Journal* **15**, 357-365.
- Zeeman, S.C., Umemoto, T., Lue, W.-L., Au-Yeung, P., Martin, C., Smith, A.M., and Chen, J.** (1998b). A mutant of Arabidopsis lacking a chloroplastic isoamylase accumulates both starch and phytyglycogen. *The Plant Cell* **10**, 1699-1711.
- Zeeman, S.C., Thorneycroft, D., Schupp, N., Chapple, A., Weck, M., Dunstan, H., Haldimann, P., Bechtold, N., Smith, A.M., and Smith, S.M.** (2004). Plastidial  $\alpha$ -glucan phosphorylase is not required for starch degradation in Arabidopsis leaves but has a role in the tolerance of abiotic stress. *Plant Physiology* **135**, 849-858.

- Zhang, J., Xin, L., Shan, B., Chen, W., Xie, M., Yuen, D., Zhang, W., Zhang, Z., Lajoie, G.A., and Ma, B.** (2012). PEAKS DB: de novo sequencing assisted database search for sensitive and accurate peptide identification. *Molecular & cellular proteomics* **11**.
- Zhang, X., Myers, A.M., and James, M.G.** (2005). Mutations affecting starch synthase III in *Arabidopsis* alter leaf starch structure and increase the rate of starch synthesis. *Plant physiology* **138**, 663-674.
- Zhang, X., Colleoni, C., Ratushna, V., Sirghie-Colleoni, M., James, M., and Myers, A.** (2004). Molecular characterization demonstrates that the *Zea mays* gene *sugary2* codes for the starch synthase isoform SSIIa. *Plant molecular biology* **54**, 865-879.
- Zhang, X., Szydlowski, N., Delvallé, D., d'Hulst, C., James, M.G., and Myers, A.M.** (2008). Overlapping functions of the starch synthases SSII and SSIII in amylopectin biosynthesis in *Arabidopsis*. *BMC plant biology* **8**, 96.
- Zhang, Y., Rempel, C., and McLaren, D.** (2014). Chapter 16 - Thermoplastic Starch. In *Innovations in Food Packaging (Second Edition)*, J.H. Han, ed (San Diego: Academic Press), pp. 391-412.
- Zhang, Y., Fonslow, B.R., Shan, B., Baek, M.-C., and Yates III, J.R.** (2013). Protein analysis by shotgun/bottom-up proteomics. *Chemical reviews* **113**, 2343-2394.
- Zhong, Y., Blennow, A., Kofoed-Enevoldsen, O., Jiang, D., and Hebelstrup, K.H.** (2019). Protein Targeting to Starch 1 is essential for starchy endosperm development in barley. *Journal of experimental botany* **70**, 485-496.
- Ziegler, G.R., Creek, J.A., and Runt, J.** (2005). Spherulitic crystallization in starch as a model for starch granule initiation. *Biomacromolecules* **6**, 1547-1554.
- Zybilov, B., Rutschow, H., Friso, G., Rudella, A., Emanuelsson, O., Sun, Q., and van Wijk, K.J.** (2008). Sorting signals, N-terminal modifications and abundance of the chloroplast proteome. *PLoS One* **3**, e1994.

# Annexes



# Proteome Analysis of Potato Starch Reveals the Presence of New Starch Metabolic Proteins as Well as Multiple Protease Inhibitors

Stanislas Helle<sup>1,2†</sup>, Fabrice Bray<sup>2†</sup>, Jérémy Verbeke<sup>2</sup>, Stéphanie Devassine<sup>2</sup>, Adeline Courseaux<sup>1</sup>, Maud Facon<sup>1,2</sup>, Caroline Tokarski<sup>2</sup>, Christian Rolando<sup>2</sup> and Nicolas Szydlowski<sup>1,2\*</sup>

<sup>1</sup> Univ. Lille, CNRS, UMR8576 – UGSF – Unité de Glycobiologie Structurale et Fonctionnelle, Lille, France, <sup>2</sup> Univ. Lille, CNRS, USR 3290 – MSAP – Miniaturisation pour la Synthèse, l'Analyse et la Protéomique, Lille, France

## OPEN ACCESS

### Edited by:

Michael James Emes,  
University of Guelph, Canada

### Reviewed by:

Naoko Crofts,  
Akita Prefectural University, Japan  
Robert D. Hancock,  
The James Hutton Institute,  
United Kingdom

### \*Correspondence:

Nicolas Szydlowski  
nicolas.szydlowski@univ-lille1.fr

† These authors have contributed  
equally to this work.

### Specialty section:

This article was submitted to  
Plant Metabolism  
and Chemodiversity,  
a section of the journal  
Frontiers in Plant Science

**Received:** 06 April 2018

**Accepted:** 15 May 2018

**Published:** 15 June 2018

### Citation:

Helle S, Bray F, Verbeke J,  
Devassine S, Courseaux A, Facon M,  
Tokarski C, Rolando C and  
Szydlowski N (2018) Proteome  
Analysis of Potato Starch Reveals  
the Presence of New Starch  
Metabolic Proteins as Well as Multiple  
Protease Inhibitors.  
*Front. Plant Sci.* 9:746.  
doi: 10.3389/fpls.2018.00746

Starch bound proteins mainly include enzymes from the starch biosynthesis pathway. Recently, new functions in starch molecular assembly or active protein targeting were also proposed for starch associated proteins. The potato genome sequence reveals 77 loci encoding starch metabolizing enzymes with the identification of previously unknown putative isoforms. Here we show by bottom-up proteomics that most of the starch biosynthetic enzymes in potato remain associated with starch even after washing with SDS or protease treatment of the granule surface. Moreover, our study confirmed the presence of PTST1 (Protein Targeting to Starch), ESV1 (Early StarVation1) and LESV (Like ESV), that have recently been identified in Arabidopsis. In addition, we report on the presence of a new isoform of starch synthase, SS6, containing both K-X-G-G-L catalytic motifs. Furthermore, multiple protease inhibitors were also identified that are cleared away from starch by SDS and thermolysin treatments. Our results indicate that SS6 may play a yet uncharacterized function in starch biosynthesis and open new perspectives both in understanding storage starch metabolism as well as breeding improved potato lines.

**Keywords:** starch, potato, proteomics, starch-binding protein, starch synthase, *Solanum tuberosum*

## INTRODUCTION

Starch is a storage form of carbohydrates in plants and the main source of calories in human and animal diets. Moreover, this polymer is used in many industrial applications for food and non-food purpose. Amylose and amylopectin, the two polymers composing starch, are made of  $\alpha$ -1,4- and  $\alpha$ -1,6-bound glucose residues (Buléon et al., 1998). The former is a rather linear molecule containing about 1% of  $\alpha$ -1,6 linkages while the latter is moderately branched with 5–6% of  $\alpha$ -1,6 linkages. Potato starch contains on average 0.06% of proteins on a weight basis (Jobling, 2004). To date, most of the identified starch-bound proteins are starch metabolic enzymes (Grimaud et al., 2008).

Granule Bound Starch Synthase 1 (GBSS1) elongates the  $\alpha$ -1,4 glucans composing amylose (Ball et al., 1998). GBSS1 represents up to 95% of the granule bound proteins and is only fully active under this granule-associated form (Rongine De Fekete et al., 1960). On the other hand, amylopectin is synthesized by the concerted activities of soluble starch synthases (SSs),

starch branching enzymes (BEs) and starch debranching enzymes (DBEs). To our knowledge, among the SS isoforms, only SS1, SS2 and SS3 were identified by proteomic analysis of maize starch as well as in starches from potato, rice, pea, barley and wheat (Denyer et al., 1993; Peng et al., 2000; Borén et al., 2004; Grimaud et al., 2008; Stensballe et al., 2008; Xing et al., 2016). A series of evidence show that SS4 is involved in starch initiation although the precise mechanism is still under investigation (Roldan et al., 2007; Szydłowski et al., 2009; Raynaud et al., 2016; Seung et al., 2016). However, the enzyme was observed at the edge of starch granules by confocal microscopy analysis of Arabidopsis leaf chloroplasts and potato amyloplasts (Szydłowski et al., 2009; Gámez-Arjona et al., 2011). A fifth SS isoform, SS5, is phylogenetically related to SS4 but lacks the C-terminal X-X-G-G-L motif conserved within other SSSs (Deschamps et al., 2008; Liu et al., 2015). The corresponding gene is conserved in plants (Deschamps et al., 2008) and is expressed during grain filling in maize (Liu et al., 2015). Thus, a function in starch synthesis was proposed but it remains to be investigated (Liu et al., 2015). An additional putative isoform, SS6 was found in the potato genome (Yu et al., 2012; Van Harsselaar et al., 2017). The corresponding gene displays a specific dynamic expression profile during tuberization (Yu et al., 2012). The putative function of this isoform in starch biosynthesis as well as its repartition within the plant kingdom is unknown.

In potato, two isoforms of branching enzymes, BE1 and BE2, which introduce  $\alpha$ -1,6 linkages within amylopectin molecules, have been characterized (Schwall et al., 2000; Rydberg et al., 2001). Similar to SSSs, starch branching enzymes were found in association with the starch granule in numerous plant species (Mu-Forster et al., 1996; Borén et al., 2004; Regina et al., 2005; Umemoto and Aoki, 2005; Grimaud et al., 2008). Interestingly, wheat BE2b interacts with SS1 and SS2a to form a trimeric complex associated with the granule (Tetlow et al., 2004, 2008b). Another complex including BE2b, BE1 and PHS1 (Starch Phosphorylase 1) was also identified although it was associated with starch only in a *be2b* mutant background (Tetlow et al., 2004; Liu et al., 2009; Subasinghe et al., 2014). Debranching enzymes comprise isoamylases and pullulanase. Both hydrolyze the  $\alpha$ -1,6 bonds of amylopectin (Wattebled et al., 2005, 2008). Plants contain one pullulanase (PU) and three isoamylases (ISA1, ISA2 and ISA3). ISA1 and ISA2 participate in starch synthesis and interact to form hetero and homo complexes where the catalytic activity is carried by ISA1 (Delatte et al., 2005). DBEs seem to be predominantly soluble but ISA2 and PU were recently identified in association with starch in rice (Xing et al., 2016; Yu and Wang, 2016).

A set of enzymes (i.e., GWD, PWD, LSF1, LSF2 and SEX4) participate in starch breakdown via glucan phosphorylation/dephosphorylation (Ritte et al., 2000; Kötting et al., 2005; Comparot-Moss et al., 2010; Hejazi et al., 2010; Santelia et al., 2011). GWD (Glucan Water Dikinase) and PWD (Phosphoglucan Water Dikinase) phosphorylate starch at C6- and C3-position of the glucose residues, respectively (Ritte et al., 2006). GWD was observed

in internal association with purified potato starch granules while PWD was shown to bind the surface of starch granules in Arabidopsis (Ritte et al., 2000; Kötting et al., 2005). LSF2 (Like SEX Four 2) and SEX4 (Starch Excess 4) also bind to starch granules as demonstrated both *in vitro* and with native starch granules isolated from Arabidopsis (Santelia et al., 2011). On the other hand, LSF1 (Like SEX Four 1) is likely associated with the granule surface according to the suborganellar distribution of the corresponding GFP-tagged proteins in Arabidopsis protoplasts (Comparot-Moss et al., 2010). Noteworthy, apart from GWD that appears to be entrapped in the starch matrix, these enzymes are located at the surface of the granules, consistent with the current model for phosphorylation/dephosphorylation driven starch breakdown (Silver et al., 2014).

In addition to starch metabolic enzymes, a series of proteins devoid of known catalytic domains were recently identified (Peng et al., 2014; Seung et al., 2015; Feike et al., 2016). Floury Endosperm 6 (FLO6) and Protein Targeting to Starch (PTST1) both contain a CBM48 (Carbohydrate Binding Module 48) that drives protein binding to starch (Peng et al., 2014; Seung et al., 2015). PTST1 also binds to GBSS1 and was proposed to target the amylose-synthesizing enzyme to starch polysaccharides (Seung et al., 2015). On the other hand, FLO6 interacts with ISA1 and is likely regulating its binding to starch although the exact mechanism remains to be uncovered (Peng et al., 2014). In both cases, inactivation of the corresponding gene leads to a phenotype similar to those of *gbss1* and *isa1* mutants, respectively (Peng et al., 2014; Seung et al., 2015). Furthermore, Early Starvation 1 (ESV1) and its homolog Like ESV1 (LESV) do not display any characterized domain (Feike et al., 2016). Both proteins are involved in the regulation of starch breakdown and likely play antagonistic roles (Feike et al., 2016). The molecular mechanisms underlying these functions are still under investigation. Nevertheless, it was proposed that both proteins modulate the organization of starch glucans and consequently affect their accessibility to catabolic enzymes (Feike et al., 2016).

These recent investigations highlight that non-catalytic starch binding proteins can also be involved in starch metabolism as well as its regulation and that some minor proteins remain to be characterized. In consequence, exhaustive proteomic analysis of starch is likely to lead to the identification of yet unknown functions in starch metabolism. The nuclear genome of a homozygous doubled-monoploid potato clone has been sequenced (Xu et al., 2011). *In silico* analysis revealed 77 genomic loci encoding enzymes related to starch metabolism with numerous novel putative isoforms (Van Harsselaar et al., 2017). Based on the potato genome annotation, the starch bound proteins were analyzed by mass spectrometry in this study. In addition to the already known starch metabolizing enzymes, we report on so far non-described starch bound proteins that are likely involved in the metabolism. Most strikingly, these include a novel isoform of starch synthase, SS6 that contains both K-X-G-G-L and X-X-G-G-L motifs, and a series of protease inhibitors. The latter are associated with the granule surface and are removed by SDS or thermolysin treatments.

## MATERIALS AND METHODS

### Plant Material and Starch Isolation

Starch granules were isolated from tubers of *Solanum tuberosum* (Mona Lisa) cultivated in the field in Villeneuve d'Ascq (50.607735, 3.143431), France, between March and July 2014. Potato tubers were washed with tap water and peeled prior extraction. Tubers were ground with a blender in 200 mL of ultrapure water. Tuber extracts were then filtered through a nylon net (100  $\mu$ m mesh) and left for sedimentation of starch granules for 3 h. The supernatant was then removed and sedimented starch was resuspended in 500 mL of ultrapure water. Starch suspensions were subsequently washed three times with 1 L of ultrapure water and stored in 20% ethanol at 4 °C.

### Surface Treatment of Starch and Protein Extraction

300 mg of starch granules were either washed five times with 2 mL of 2% SDS or submitted to thermolysin treatment. 60  $\mu$ g of thermolysin (Promega, France) were added to 300 mg of starch in 50 mM Tris, 0.5 mM CaCl<sub>2</sub>, pH 8.0 prior to incubation overnight at 37°C. Starch granules were then washed five times with 2 mL of ultrapure water and protein extraction was carried out with the addition of 5 mL of extraction buffer (0.2 M Tris, 2% SDS, 20% glycerol, 50 mM DTT, pH 6.8) and subsequent incubation at 100°C for 20 min with regular vortexing. After centrifugation (20000 g, 10 min), the supernatant was concentrated on an ultrafiltration column (Amicon Ultra-15, Merck, Germany) by centrifugation for 30 min at 7500 g. Bottom-up identification was then performed either after protein separation by SDS-PAGE or by shotgun proteomics.

### Sodium Dodecyl Sulfate Polyacrylamide Gel Electrophoresis (SDS-PAGE)

All chemicals were purchased from Sigma-Aldrich (St. Louis, MO, United States). A stock solution of acrylamide (50 mL) was prepared using 14.6 g acrylamide and 390 mg *N,N,N',N'*-methylene-bis-acrylamide as a cross-linking agent. 330 mg of Dextran (MW 500 000) was also incorporated to the previous solution in order to improve gel separation according to the protocol developed in (Ainseba-Chirani et al., 2011). The polyacrylamide concentrations were 4% and 7 or 10% in stacking (0.5 M Tris-Base, 14 mM SDS, pH 6.8) and running (1.1 M Tris-Base, 0.4 M Tris-HCl and 20 mM SDS, pH 8.8) gels, respectively. Protein samples were dissolved in a denaturing buffer containing 0.375 M Tris, 2 mM DTT, 1.4 M SDS, 2% glycerol (v/v), and 0.02% bromophenol blue (w/v). Low- and high-molecular weight (LMW and HMW, respectively) protein markers were used for gel calibration. Electrophoresis was carried out using a SE 600 Ruby system (GE Lifescience, Velizy-Villacoublay, France) at 40 mA per gel in a migration buffer composed of 0.025 M Tris, 0.192 M glycine, 0.1% SDS (w/v). Gels were fixed in a 20% ethanol, 7% acetic acid (v/v) solution for 20 min prior to overnight incubation in a Ru(BPS)<sub>3</sub>-containing staining solution under continuous gentle agitation. Gels were finally washed in 25% ethanol,

14% (v/v) acetic acid and digitized using a Typhoon® 9000 scanner (GE Lifescience, Velizy-Villacoublay, France). Image acquisition was performed at a resolution of 100 nm, photomultipliers 580 V with the blue laser and filters corresponding to  $\lambda_{Exc} = 488$  nm for excitation and  $\lambda_{Em} = 610$  nm for emission.

### In-Gel Tryptic Digestion

The protein bands were excised from the gel and destained in 50% acetonitrile, 25 mM NH<sub>4</sub>HCO<sub>3</sub> prior to dehydration in 100% acetonitrile. Proteins were reduced by incubating the gel bands in 10 mM DTT, 25 mM NH<sub>4</sub>HCO<sub>3</sub> at 56°C for 1 h, and alkylation was performed with 55 mM iodoacetamide, 25 mM NH<sub>4</sub>HCO<sub>3</sub> at room temperature for 45 min in the dark. Gel pieces were dehydrated in 50% acetonitrile, 25 mM NH<sub>4</sub>HCO<sub>3</sub> twice for 30 min, and then in 100% acetonitrile during 5 min. Gel samples were dried for 30 min at room temperature prior to rehydration with a solution of sequencing-grade trypsin (10 ng/ $\mu$ L in 25 mM NH<sub>4</sub>HCO<sub>3</sub>) on ice for 30 min and subsequently submitted to overnight incubation at 37°C. The resulting peptides were extracted with 100% acetonitrile containing 0.1% trifluoroacetic acid (TFA) for 30 min. The extracts were finally dried in a vacuum concentrator and dissolved in a solution of 0.1% formic acid for mass spectrometric analysis.

### Enhanced Filter Aided Sample Preparation (eFASP) Tryptic Digestion

Enhanced Filter Aided Sample Preparation protocol was performed following the protocol of (Erde et al., 2014), with some modifications. Protein extracts were incubated in 50  $\mu$ L of reducing buffer (4% SDS, 0.2% deoxycholic acid, 50 mM DTT, 200 mM NH<sub>4</sub>HCO<sub>3</sub>) overnight at 4°C prior to centrifugation at 13,000 g for 15 min. The supernatants were then mixed with 200  $\mu$ L of exchange buffer (8 M urea, 0.2% deoxycholic acid, 100 mM NH<sub>4</sub>HCO<sub>3</sub>, pH 8) and transferred on ultrafiltration units (Amicon®, 10 kDa cutoff; Millipore, Billerica, MA, United States). The latter were centrifuged at 13000 g during 30 min and an additional buffer exchange step was performed. Reduced proteins were alkylated within the filtration units with the addition of 100  $\mu$ L of 8 M urea, 50 mM iodoacetamide, 100 mM NH<sub>4</sub>HCO<sub>3</sub>, pH 8 followed by incubation at 37°C at room temperature in darkness. Proteins were then washed once with 200  $\mu$ L of exchange buffer and twice with 200  $\mu$ L of eFASP digestion buffer (50 mM NH<sub>4</sub>HCO<sub>3</sub>, 0.2% deoxycholic acid pH 8). Tryptic digestion was carried out by incubating protein samples with 1  $\mu$ g of trypsin in 120  $\mu$ L of eFASP digestion buffer at 37°C for 16 h and under constant agitation. Peptides were recovered in a new collection tube by centrifugation at 13000 g for 20 min. For complete peptide recovery, ultrafiltration units were subsequently washed twice with 50  $\mu$ L of 50 mM NH<sub>4</sub>HCO<sub>3</sub>.

### Phase Transfer

Peptides were precipitated with the addition of 200  $\mu$ L of ethyl acetate and 2.5  $\mu$ L of TFA. Peptide precipitates were then washed three times with 800  $\mu$ L of ethyl acetate with centrifugation at 13000 g for 10 min after each wash. Ethyl acetate was evaporated



by placing the peptide pellets at 60°C, in a fume hood, for 5 min and residual organic and volatile salts were removed by vacuum drying. This vacuum drying step was then repeated twice after the addition of 50% methanol and dried peptides were dissolved in 10 µL of milliQ water. One microliter of peptide solution, corresponding to 5 µg of protein digest, was diluted in 9 µL of nano-HPLC buffer (5% acetonitrile and 0.1% formic acid).

## Protein Identification Using LC-MS/MS

A nanoflow HPLC instrument (U3000 RSLC Thermo Fisher Scientific™) was coupled on-line to a Q Exactive plus (Thermo Scientific™) with a nano-electrospray ion source. 1 µL of gel-band peptide extracts or 1 µg of eFASP digests were loaded onto the preconcentration trap (Thermo Scientific™, Acclaim PepMap100 C18, 5 µm, 300 µm i.d. × 5 mm) using partial loop injection, for 5 min at a flow rate of 10 µL·min<sup>-1</sup> with buffer A (5% acetonitrile and 0.1% formic acid) and separated on a reversed phase column (Acclaim PepMap100 C18, 3 µm, 75 mm i.d. × 500 mm) with a linear gradient of 5–50% buffer B (75% acetonitrile and 0.1% formic acid) at a flow rate of 250 nL·min<sup>-1</sup> and temperature of 45°C. Gradient length was 100 min and 240 min for gel bands and eFASP samples, respectively. The column was washed with 99% of buffer B for 10 min and reconditioned with buffer A. The total time for an LC MS/MS run was about 120 min long for gel band analysis and 270 min for eFASP digestion.

## Mass Spectrometry

The MS data was acquired on a Q-Exactive™ plus instrument (Thermo Scientific™) using a data-dependent top 20 method, dynamically choosing the most abundant precursor ions from the survey scan (350–1600 m/z) for Higher energy Collisional Dissociation (HCD) fragmentation. Dynamic exclusion duration was 60 s. Isolation of precursors was performed with a 1.6 m/z window and MS/MS scans were acquired with a starting mass of 80 m/z. Survey scans were acquired at a resolution of 70,000 at m/z 400 (AGC set to 10<sup>6</sup> ions with a maximum fill time of 100 ms). Resolution for HCD spectra was set to 70,000 at m/z 200 (AGC set to 10<sup>5</sup> ions with a maximum fill time of 200 ms). Normalized collision energy was 28 eV. The underfill ratio, which specifies the minimum percentage of the target value likely to be reached at maximum fill time, was defined as 0.4%. The instruments was run with peptide recognition mode (i.e., from 2 to 8 charge), exclusion of singly charged and of unassigned precursor ions enabled.

## Protein Identification With PEAKS 7.0

The acquired raw files were analyzed with PEAKS Studio 7.0 (Bioinformatics Solutions Inc.) (Zhang et al., 2012) using a custom made database including PGSC\_DM\_v3.4\_pep\_non-redundant (April 2017, 52,570 entries<sup>1</sup>), and the addition of unannotated and miss annotated known proteins according to (Van Harsselaar et al., 2017). The peptide mass tolerance was set to 10 ppm and 0.01 Da for MS/MS. Variable modifications included were as follows: Oxidation of M, Y, H, deamidation of N,

Q, carbamidomethylation of C, phosphorylation of Y, S, T, pyro-G. For high-confidence, peptide threshold was FDR 1%, protein –10lgP > 20 and protein were identified with two peptides.

## Label-Free Quantification With MaxQuant and Perseus

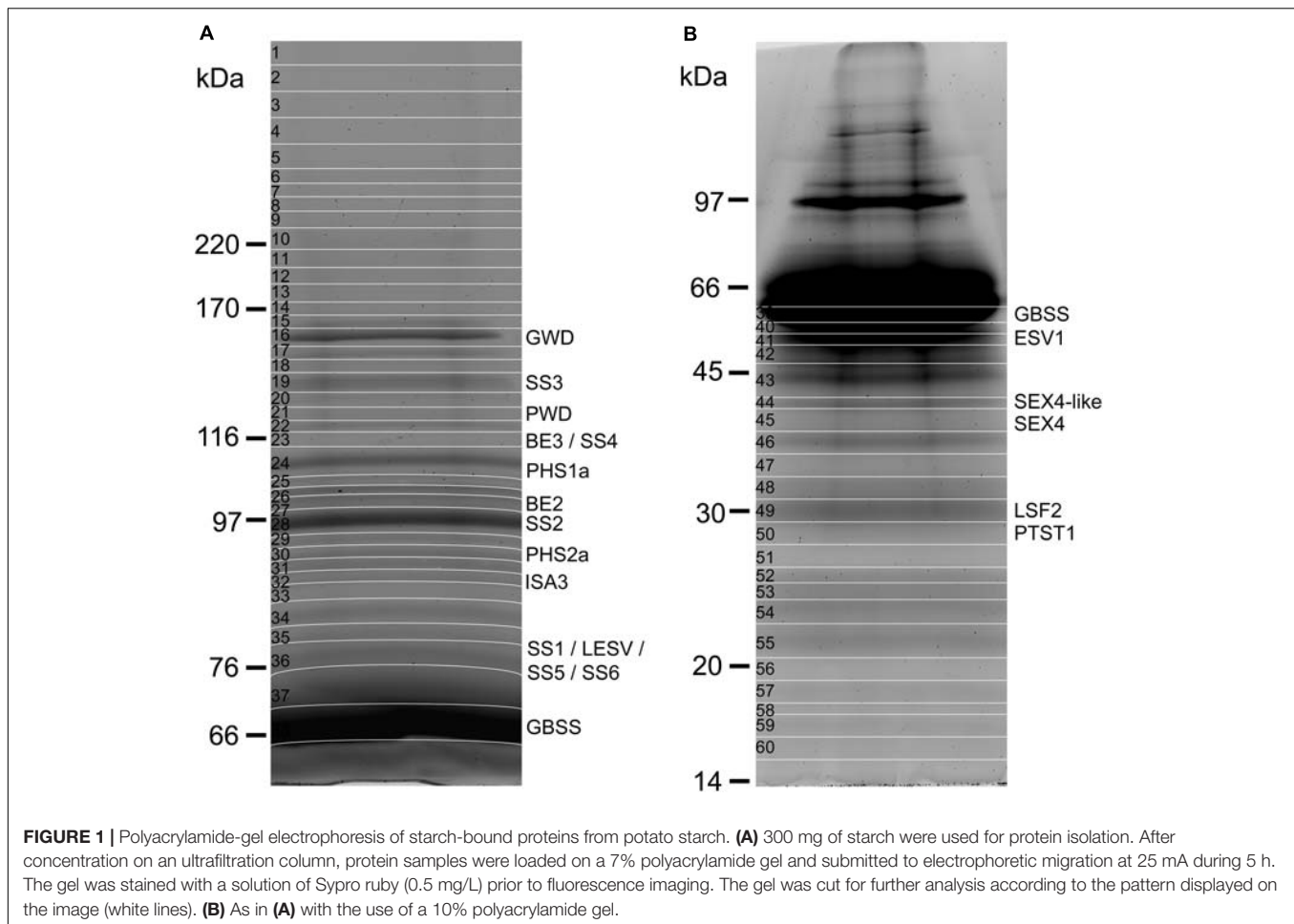
The acquired raw files were analyzed with MaxQuant 1.5.3.30 using the Andromeda search engine (Cox and Mann, 2008; Tyanova et al., 2016a). Proteins were identified by searching MS and MS/MS data of peptides against the UniProt-Solanum tuberosum database (April 2017, 49,664 entries) and PGSC\_DM\_v3.4\_pep\_non-redundant (April 2017, 52,570 entries). The precursor mass and fragment mass were identified with an initial mass tolerance of 10 ppm and 20 ppm, respectively. The search included variable modifications of methionine and proline oxidation, asparagine and glutamine deamidation, tyrosine, serine and threonine phosphorylation, lysine acetylation, glutamine to pyroglutamate conversion and fixed modifications of carbamidomethyl cysteine. Minimal peptide length was set to six amino acids and a maximum of three mis-cleavages was allowed. The false discovery rate (FDR) was set to 0.01 for peptide and protein identifications. MS runs were analyzed with the “match between runs” option 2 min and a 30-min retention time window. In the case of identified peptides that are all shared between two proteins, these were combined and reported as one protein group. Proteins matching to the reverse database were filtered out. LFQ intensities for respective protein groups were uploaded in Perseus and analyzed (Tyanova et al., 2016b). We filter out the reverse (decoy) database hits (proteins match again a reversed or scrambled database), the contaminants (proteins match again contaminants database) and the proteins with a localization probability < 0.75. Raw LFQ intensities were logarithmized by Log2. At least three LFQ values per protein group needed to be present for the analysis. To replace non-quantified values with low intensities, data imputation was performed based on normal distribution of LFQ intensities. Proteins were identified with two peptides minimum. Significant interactors were determined using a two-sample analysis *t*-test and multiple sample test with Benjamini-Hochberg FDR at 0.05. Normalization with *Z*-score was determined and hierarchical clustering was constructed with Euclidean distance.

## RESULTS

### Granule-Bound Proteins Related to Starch Metabolism

Starch-bound proteins were analyzed by SDS-PAGE and fluorescence imaging after isolation from 300 mg of potato starch and concentration by ultrafiltration columns (Figure 1). Appropriate separation of bands with an apparent MW above 60 kDa and below 45 kDa was achieved with the use of 7 and 10% polyacrylamide gels, respectively (Figures 1A,B). Apparent MW of each band was determined from their relative migration distance (Rf) using the linear range of

<sup>1</sup>[http://solanaceae.plantbiology.msu.edu/pgsc\\_download.shtml](http://solanaceae.plantbiology.msu.edu/pgsc_download.shtml)



a standard curve constructed with the Rfs of the protein ladder (Table 1 and Supplementary Figure S1). Sixty gel bands were cut for further analysis according to the patterns displayed in Figures 1A,B. Each band was submitted to in-gel trypsin hydrolysis prior to nanoLC-nanoESI-MS/MS analysis. Protein identification was then carried out with the use of a homemade peptide database including non-redundant CDS annotated by the potato genome sequencing consortium (PGSC<sup>2</sup>) (Hirsch et al., 2014) (Supplementary Data Sheets 1, 2). Several starch metabolism isoforms including BE2 were identified in previous studies but have not been annotated in this database (Van Harsselaar et al., 2017). On the other hand, known starch metabolism genes were miss annotated or annotated under a truncated form (e.g., SS3, *PHS1a*, *ISA3*). Therefore, the iTAG ID, Sotub09g011090.1.1 (BE2), Sotub02g012780.1.1 (AMY3-like), Sotub07g025820.1.1 (BE1.2) and UniProt references, P30924 (BE1), Q43846 (SS3), P32811 (*PHS2a*), Q84YG5 (*ISA3*), and P04045 (*PHS1a*) were manually added to the database. The protein confidence threshold ( $-10\text{LgP}$ ) was empirically set to 100 and proteins were identified with at least two peptides for a confident identification.

<sup>2</sup>[http://solanaceae.plantbiology.msu.edu/pgsc\\_download.shtml](http://solanaceae.plantbiology.msu.edu/pgsc_download.shtml)

A total of 36 proteins were identified in this study (Tables 1, 2 and Supplementary Data Sheets 1, 2), 20 of which were related to starch metabolism (Table 1). GBSS is the major starch-binding protein (Stensballe et al., 2008). According to previous descriptions, a band of high intensity was observed at 62.8 kDa, similar to the predicted MW of potato GBSS (Figures 1A,B and Table 1). In addition, while only SS2 and GWD were previously detected in association with potato starch, SS1, SS2, SS3, BE1, BE2, GWD, and *PHS1* have been described in maize, wheat, rice, pea, and barley (Grimaud et al., 2008; Stensballe et al., 2008). These isoforms were also found in our study (Table 1). Noteworthy, we observed two additional isoforms of SS, SS5 and SS6, with 10 and 9 unique peptides, respectively, as well as a protein highly homologous to SEX4 annotated as SEX4-like (Table 1) (Van Harsselaar et al., 2017). Both SS5 and SS6 genes were annotated in potato but their functions have not yet been investigated (Deschamps et al., 2008; Yu et al., 2012; Liu et al., 2015; Van Harsselaar et al., 2017). BLAST search<sup>3</sup> with the use of the potato SS6 sequence, XP\_006353746.1, reveals two other genes similarly annotated in *Solanum lycopersicum* (NP\_001234387.1) and *Vitis vinifera* (NP\_001268021.1). Furthermore, this analysis showed

<sup>3</sup><https://blast.ncbi.nlm.nih.gov/Blast.cgi>

**TABLE 1** | Proteins related to starch metabolism.

Isoform	PGSC/UniProt/iTAG IDs	Predicted MW (kDa)	Observed MW (kDa)	Band number	Number of matched unique peptides	Coverage (%)
GWD	PGSC0003DMP400013565	163.2	160.7	17	78	39
SS3	Q43846	139	142.7	19	57	37
PWD	PGSC0003DMP400029040	132.1	129.3	21	27	25
SS4	PGSC0003DMP400014590	113.8	122.9	23	4	4
BE3	PGSC0003DMP400017627	104.1	122.9	23	9	9
PHS1a	P04045	109.4	118.2	24	23	28
BE2	Sotub09g011090.1.1	100	105.7	27	54	41
SS2	PGSC0003DMP400002384	85.2	103.0	28	203	87
PHS2a	P32811	95.1	96	30	15	21
ISA3	Q84YG5	87.4	88.9	32	31	45
SS1	PGSC0003DMP400032329	70.6	74.1	36	46	46
LESV	PGSC0003DMP400028510	63.4	74.1	36	72	77
SS6	PGSC0003DMP400023948	69.5	74.1	36	9	29
SS5	PGSC0003DMP400053324	76.5	74.1	36	10	27
GBSS1	PGSC0003DMP400021398	66.5	62.8	39	870	94
ESV1	PGSC0003DMP400017605	48.9	51.8	41	6	17
SEX4-like	PGSC0003DMP400047519	41.5	42.8	44	12	38
SEX4	PGSC0003DMP400026742	41.5	38.3	45	11	33
LSF2	PGSC0003DMP400050640	24.9	30.7	49	13	50
PTST1	PGSC0003DMP400053308	28.5	28.6	50	21	58

Observed MW was estimated from the relative migration distance (*Rf*) of each band with the use of a calibration curve constructed with the *Rf*s of the protein ladder.

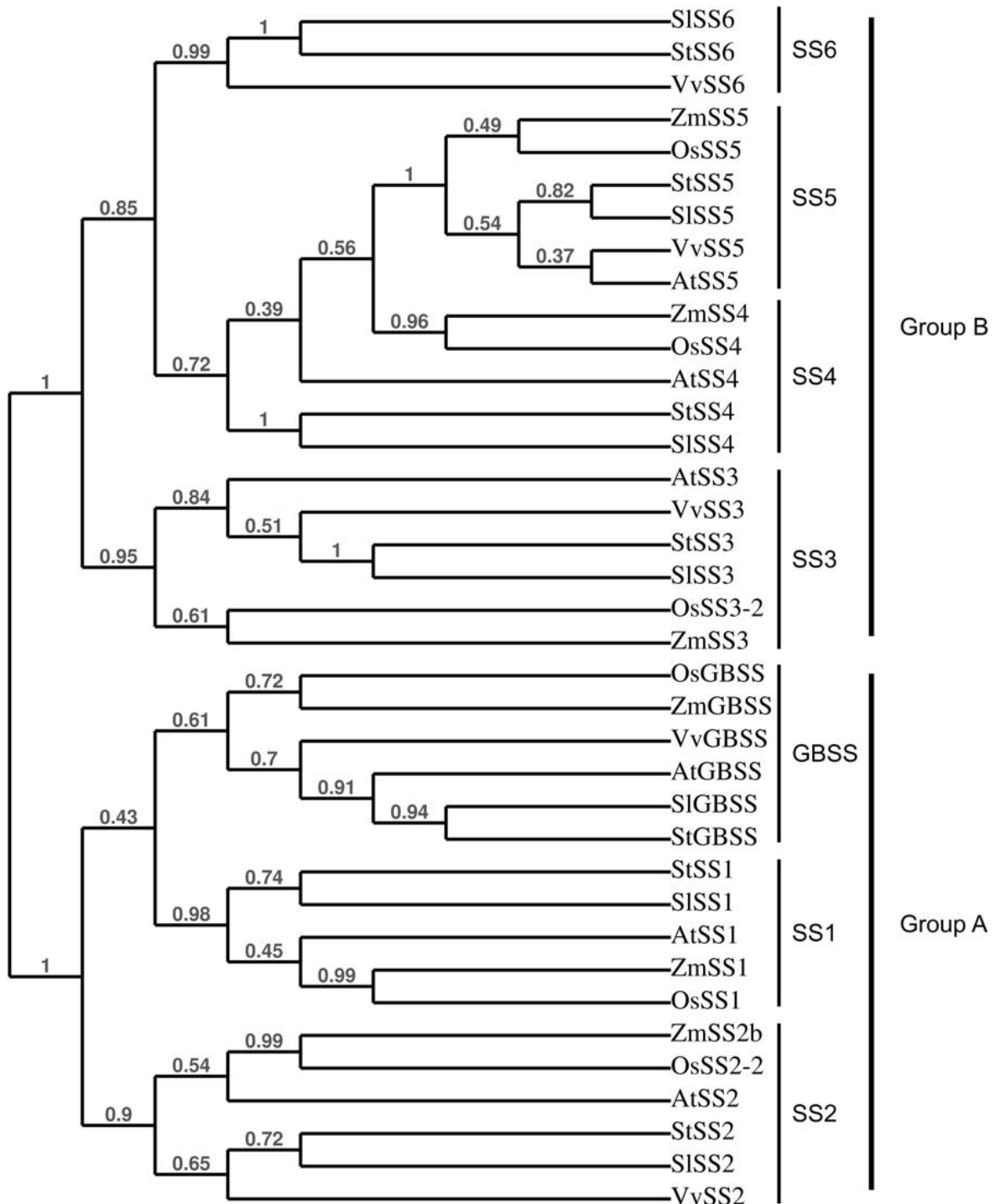
90 sequences of hypothetical proteins with a score ranging from 572 to 1301 and an *E*-value of 0 in the genomes of numerous plant species, indicating that the gene is conserved. Phylogenetic analysis including all classes of Ss confirmed that SS6 proteins define a new class of starch synthases (Figure 2). SS6 class belongs to group B-starch synthases similar to SS3-5, and most closely resembles to SS4 while the protein lacks the long N-terminal extension usually observed in SS4 proteins (Leterrier et al., 2008). Moreover, the SS6 sequence contains both K-X-G-G-L and X-X-G-G-L highly conserved motifs, whereas the latter are missing in SS5 (Figure 3) (Cao et al., 1999).

Two isoforms of PHS2, PHS2a and PHS2b, were identified in the potato genome (Van Harsselaar et al., 2017). This enzyme is also called PHS2 (High glycogen activity) and was reported as a cytosolic form of starch phosphorylase participating to the catabolism of starch degradation products in the cytosol (Fettke et al., 2004, 2005). On the other hand, the protein was also identified in proteomic studies of the plastid stroma, suggesting multiple sub-cellular localization for this enzyme (Zybailov et al., 2008; Helm et al., 2014). Interestingly, PHS2a was found attached to the starch granule in this study (Table 1). Moreover, PTST1, a protein responsible for targeting GBSS to the starch granule, as well as ESV1 and LESV, two regulatory proteins involved in starch degradation through unknown mechanisms, were recently identified in Arabidopsis (Seung et al., 2015; Feike et al., 2016). The potato genome contains homologs of PTST1, LESV, and ESV1 that were observed in our analysis at 28.6, 74.1, and 51.8 kDa, respectively, which is in good accordance with their predicted MW (Figure 3 and Table 1). Among starch degradation

enzymes, PWD, SEX4, ISA3, and LSF2 were detected with a number of unique peptides ranging from 11 to 31 (Table 1), thus corroborating that these enzymes interact with starch granules.

## Other Starch-Bound Proteins

In addition to proteins related to starch metabolism, we identified 10 sequences of predicted protease inhibitors (Table 2). Protein sequence alignment and phylogenetic analysis established a classification comprising four groups within which the level of protein identity ranges from 77.55 to 99.09% (Supplementary Figure S2A). Nine of these proteins belonged to the potato Kunitz-type proteinase inhibitor (PKPI) family and grouped specifically with one of the three PKPI classes, A, B or C (Ishikawa et al., 1994) (Supplementary Figure S2B). The two other proteinase inhibitors contained 2 or 3 proteinase inhibitor II (PINII) domains and belonged to the potato type II protease inhibitor family (Supplementary Figure S2). Intriguingly, this analysis also revealed the presence of a thioredoxin (TRX) protein (PGSC0003DMP400021964) homologous to Arabidopsis thioredoxins M as well as a glutathione peroxidase (GPX, PGSC0003DMP400045625) similar to Arabidopsis GPX1 (Table 2 and Supplementary Data Sheet S2). Both enzymes are involved in redox regulation in plants with TRXM participating to the regulation of metabolic processes in function of light and GPX acting during plant-cell stress response (Buchanan, 1984; Noctor et al., 2012). Additionally, an acid phosphatase (PGSC0003DMP400028845, stem 28 kDa glycoprotein) was identified with five unique peptides and a Peptidyl-prolyl *cis-trans* isomerase (PGSC0003DMP400009292) belonging to the cyclophilin family with eight unique peptides (Table 2). Finally,



**FIGURE 2 |** Phylogenetical tree showing the relationships between plant SSs. The tree was constructed with the use of PhyML (Maximum-likelihood based) and confidence limits were assigned by bootstrapping the alignment with 100 trials. At, *Arabidopsis thaliana*; St, *Solanum tuberosum*; Zm, *Zea mays*; Os, *Oryza sativa*; Vv, *Vitis vinifera*; Sl, *Solanum lycopersicum*. SISS6: NP\_001234387.1, StSS6: XP\_006353746.1, VvSS6: NP\_001268021.1, ZmSS5: NP\_001123603.1, OsSS5: XP\_015626202.1, StSS5: PGSC0003DMP400053324, SISS5: XP\_019067616.1, VvSS5: NP\_001290013.1, AtSS5: NP\_569018.1, ZmSS4: NP\_001123590.1, OsSS4: XP\_015639005.1, AtSS4: NP\_193558.3, StSS4: PGSC0003DMP400014590, SISS4: NP\_001234617.2, AtSS3: NP\_001184965.1, VvSS3: XP\_002269011.2, StSS3: Q43846, SISS3: NP\_001234623.1, OsSS3-2: XP\_015650668.1, ZmSS3: NP\_001104881.2, OsGBSS: XP\_015644490.1, ZmGBSS: NP\_001105001.3, VvGBSS: XP\_010660257.1, AtGBSS: NP\_174566.1, SlGBSS: NP\_001311458.1, StGBSS: PGSC0003DMP400021398, StSS1: PGSC0003DMP400032328, SISS1: XP\_010318024.1, AtSS1: NP\_197818.1, ZmSS1: NP\_001104892.1, OsSS1: XP\_015644241.1, ZmSS2b: NP\_001106014.1, OsSS2-2: XP\_015627452.1, AtSS2: NP\_186767.1, StSS2: PGSC0003DMP400002383, SISS2: XP\_004232219.1, VvSS2: XP\_010661072.1.

STGBSS	95	KTGGLGDVVLGGLPPALAAARGHRVMTISPRYDQ-----YKDAWDTSV---AVEVKVGD	144
STSS1	145	KTGGLGDVCGSLPMALAAARGHRVMVSPRYLNGGSPDEKYANAVDLDV---RATVHCFGD	201
STSS2	290	KTGGLGDVAGALPKALARRGHRVMVAPRYDN-----YPEPQDSGV---RKIYKVDGQ	339
STSS3	794	KVGGLGDVVTSLSRAVQDLNHNVDIILPKYDCLK-----MNNVKDFRFHK---NYFWGG	844
STSS5	211	SVGSLAPYVTGLSRALQRKGNLVEVILPKYASLN-----LNEVHGLREVEAEFHSYFNGQ	265
STSS4	518	KVGGLGDVVAGLQKALQEKGLVEIVLPKYDCMQ-----YESIKDMKVLVVVSVESYFDGR	572
STSS6	139	KVGGLGDVVVTGLGRACLNRGHKVDVMIFFYECIP-----KHCINELALMK-TYNSYHDGN	192
STGBSS	145	IEIVRFFHCYKRGVDRVFDHMPFLEKVMGKTGSKIYGPKAGLDYLDNELRFSLLCQAA	204
STSS1	202	AQEVAFYHEYRAGVDWVFDHSSYCRP--G----TPYGDIYGAFG-DNQFRFTLLSHAAC	254
STSS2	340	DVEVYFQAFIDGVDFVFDISHMFRHI--G----NN---IYGGNRVDILKRMVLFCKAAI	390
STSS3	845	-TEIKVWFGKVEGLSVYFLEPQN--GL-----FSKGCVYG-CS-NDGERFGFFCHAAL	892
STSS5	266	LHGNRIWTGVVCGIGVTFIEPLYSAF-----FGCENIYG-YS-NDFERFTYFSRSL	316
STSS4	573	LYKNNIWTGIVEGLPVYFIEPQHPGKF-----FCRGQLYG-EH-DDFKRFSFFSRVAL	623
STSS6	193	WVACKAYRGEVSGVPVILIEPSN--HF-----FKGKNYGGSY-NELDAYLFFSRACL	242
STGBSS	265	AYQGRFSFSDPFLNLPDEFRGSFDFIDGY----EKPVKGRKINWMKAGILESHRVTVS	320
STSS1	310	AHQGVEPAVTYNNLGLPPQWYGAWEWIFPTWARAHALDTGETVNVLKGAIIVADRILTVS	369
STSS2	447	AHQGRGPLEDFSVDLPPHYMDPFKLYDP-----VGGEHFNIFAAGLKTADRVTVS	498
STSS3	941	EFGADLIG-----RAMTNADKATTVS	961
STSS5	365	ESQCVQPEKLLALCGLDPYGLHCSDRLQDN-----NKSHLVNVLKAGVYVSNNVIIMS	417
STSS4	672	EYQGTAAASELTSCGLDAYHLNKPDRMQDN-----SANDRINPVKGAIVFSNIVTVS	724
STSS6	291	EHYGECRQEQLSKFGLDGSAYATEDKAVDD---RTVGHNPERLSLLKGGIVYSNAIVTVS	347
STGBSS	321	PYYAQELVSAVDKGVELDSV--LRKTCITGIVNGMDTQEWNPATDKYTDVK---YDITT	374
STSS1	370	QGYSWEITPE-GGYGLHELLSSRQSVLNGITNGIDVNDWNPSTDEHIASHY---SIND	424
STSS2	499	HGYSWELKTSQ-GGWGLHQIINENDWKLQGI VNGIDTKAWNPELDVHLQSDGMNYSLDT	557
STSS3	962	PTYSQEVSG-----NPVIAPHLHKFHGIVNGIDPDIWDPLNDKFIPIPI---YTSN	1009
STSS5	418	SMQTKGQIIHA-TSHGLEPTLTIHKDKLVVAPPGFSSAWDPSVDMFLPQN---YSA--	470
STSS4	725	PTYAQEVRSQAQ-GGKGLHATINSHSKKFAGILNGIDTDAWNPNASDNFLKVQ---YSAS-	778
STSS6	345	PTYLKETLCSG----WLSGALMRNRDKYSGI LINGIDTEMWNPATDIYLPK---FDAS-	398
STGBSS	375	VMDAKPLLKEALQAAVGLPVDK-----KIPLIGFIGRLEEQKGS DILVAAIHKFIG	425
STSS1	425	-LSGKVQCKTDLQKELGLPIRP-----DCPLIGFIGRLDYQKGVDIILSAIPELMQ	474
STSS2	556	LQTGKPCCKAALQKELGLPVRD-----DVPLIGFIGRLDPQKGVDLIAEAVPMMG	608
STSS3	1010	VVEGKTAAKEALQRKLGKQAD-----LPLVGIITRTHQKGIHLIKHAIWRTLE	1059
STSS5	471	DLKGSVCKVSLQQLGLQEKAK-----SIVLVGCI FSDISDIELENLKTLIWMASR	521
STSS4	779	DIDGKIQNKALRRLGLSSSDP-----RQPLVSCITRLVPQKGVHLIRHAIYRTLE	830
STSS6	399	KTEGKRICKQFVQRGLGLPFQGIKHGICVADQIPLVVCITRLVAQKGLHLITHAIKHVEE	458
STGBSS	481	RFEPGCLIQLHAMRYGTVPICASTGGILVDTVKEGY--TG---FHMGA FNVECDVVDPA	535
STSS1	530	RFEPGCLNQLYAMRYGTIPIVHSTGGILRDTVKDFN--PY---AQEGNGEGTGWTFSP	584
STSS2	664	RFEPGCLNQLYAMRYGTIPVVAVGGILRDTVQPFD--PF---NE---SGLGWTFSRAEA	714
STSS3	1120	IFEPGCLTQLTAMRYGSI PVVVKTTGGLYDTVFD--VDHKERAQCGLEPNFSFDGADA	1177
STSS5	579	FDGTVLQIPLKAMRYGAMPILLDFTD---SKYGHSVDRDLEGT-----EFS-----RY	626
STSS4	889	IFEPGCLTQMIAMRYGSIPIARKTTGGLNDSVFD--VDDDSIPL----QFRNGFTFVT	942
STSS6	517	MYEPGCLAQMIAMRYGAVPIVVKTTGGIADTVFD--MDDQSH-T----EIANGVFVFE	569

**FIGURE 3** | Sequence alignment of the catalytic domains of starch synthase proteins from *Solanum tuberosum*. The protein alignment was constructed with the use of clustal omega (<https://www.ebi.ac.uk/Tools/msa/clustalo/>). The outlined boxes indicate the highly conserved K-X-G-G-L and X-X-G-G-L motifs. The unique peptides identified by mass spectrometry are highlighted in gray within SS4, SS5, and SS6 sequences. St, *Solanum tuberosum*, StGBSS: PGSC0003DMP400021398, StSS1: PGSC0003DMP400032328, StSS3: Q43846, StSS4: PGSC0003DMP400014590, StSS5: XP\_006357749.1, StSS6: XP\_006357746.1.

four unique peptides of a Ci21A protein as well as two peptide from b-tubulin were found in this study (Table 2).

## Surface Treatment of the Starch Granules

To investigate the nature of protein interactions with the starch granule, the samples were submitted to SDS-washes or thermolysin treatment prior to protein isolation and MS/MS shotgun analysis after eFASP tryptic digestion (Figure 4). Proteins were quantified by the MaxLFQ procedure with the

use of the MaxQuant software (Cox et al., 2014) (Figure 4 and Supplementary Data Sheet S3). As expected, label-free quantification of starch-bound proteins confirmed that GBSS is the predominant isoform, representing 87–69% of total proteins in non-treated and SDS-treated samples, respectively (Figure 4A). This disparity was due to a significant decrease ranging from 75 to 92% in the content of the protease inhibitors following surface treatment whereas all starch metabolism proteins including GBSS remained constant (Figures 4A,B). These data allow building two groups according to protein behaviors following surface treatment of starch granules.

**TABLE 2** | Other proteins identified in this study.

PGSC ID	Protein annotation	Predicted MW (kDa)	Number of matched unique peptides	Coverage (%)
PGSC0003DMP400016823	Kunitz-type proteinase inhibitor	20.1	5	35
PGSC0003DMP400017933	Serine protease inhibitor 7	24	7	35
PGSC0003DMP400017950	Cysteine protease inhibitor 1	24.8	5	37
PGSC0003DMP400017953	Kunitz-type tuber invertase inhibitor	24.5	16	67
PGSC0003DMP400017939	Cysteine protease inhibitor 1	21.8	10	26
PGSC0003DMP400016824	Aspartic protease inhibitor 5	23.9	2	22
PGSC0003DMP400008029	Proteinase inhibitor type-2 P303.51	16.6	11	63
PGSC0003DMP400017942	Cysteine protease inhibitor 1	25.1	2	15
PGSC0003DMP400017952	Cysteine protease inhibitor 9	13.8	6	66
PGSC0003DMP400016822	Aspartic protease inhibitor 8	24.1	3	52
PGSC0003DMP400021964	Thioredoxin	19.5	4	38
PGSC0003DMP400028845	Stem 28 kDa glycoprotein	29.5	5	26
PGSC0003DMP400045625	Glutathione peroxidase	26.1	7	33
PGSC0003DMP400009292	Peptidyl-prolyl <i>cis-trans</i> isomerase	26.5	8	29
PGSC0003DMP400011774	Ci21A protein	12.4	4	30
PGSC0003DMP400051081	Beta-tubulin	50.5	2	16

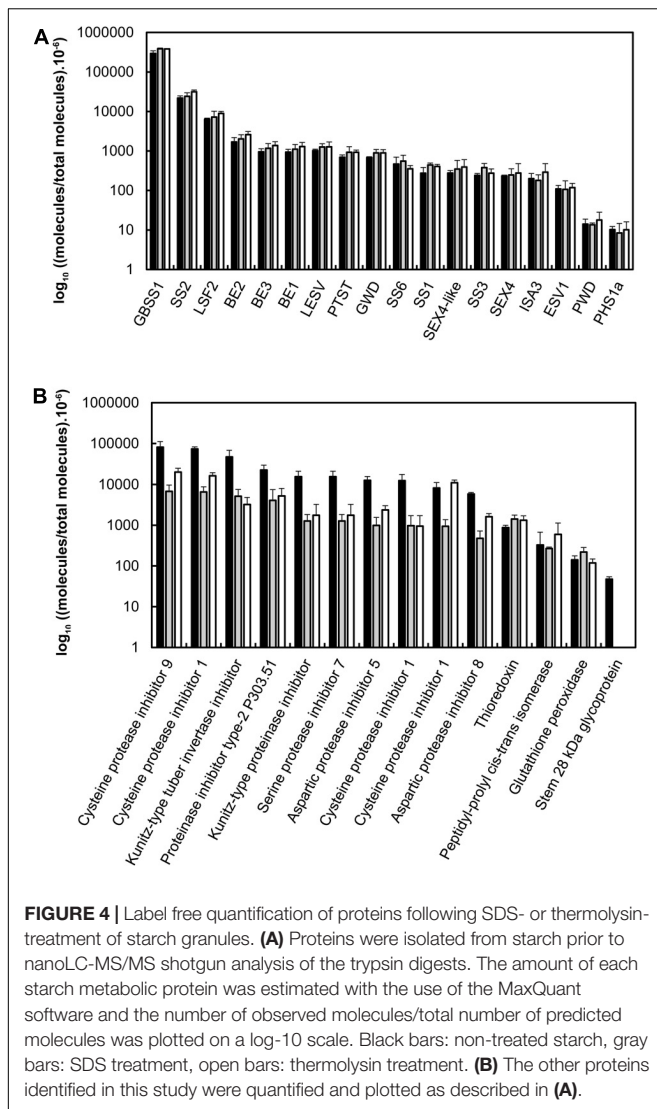
A drastic decrease was observed in the first group including protease inhibitors and the acid phosphatase indicating weak interaction of these proteins with starch polymers (**Figure 4B**). On the other hand, the second group comprised the enzymes of the starch metabolic pathway that remained attached to starch (**Figure 4A**). Interestingly, the thioredoxin, cyclophilin and the glutathione peroxidase displayed a profile similar to that of starch biosynthetic enzymes, suggesting that their interaction with starch is of the same nature.

## DISCUSSION

The main objective of this study was to investigate exhaustively the proteome associated with potato starch granules. This work relied on the publication of the potato genome (Xu et al., 2011) as well as on a previous genomic analysis of potato genes related to starch metabolism (Van Harsseelaar et al., 2017). In addition to the 77 loci identified in the latter study, bottom-up proteomics including *in silico* analysis of all predicted genes allowed us to seek for new, previously uncharacterized starch-bound proteins. Our data confirm previous findings of GBSS, SS2, and GWD as the major proteins bound to potato starch (Stensballe et al., 2008). In addition, we report that SS1, SS3, BE1, BE2, and PHS1a, already observed in other plant species, are also found in association with starch in potato (Grimaud et al., 2008). This result is particularly striking since SSs, BEs as well as PHS1 are known to form protein complexes that are likely modulating individual protein activities (Tetlow et al., 2004, 2008a). Our study lends support both on the conservation of these protein-protein interactions in potato and the interaction of these complexes with the granule. Technical limitations did not allow investigating protein complexes from starch-isolated proteins. Studies on protein-protein complexes usually examine the soluble fraction (i.e., stromal

proteins) since physicochemical treatments applied during the process of starch-bound protein extraction alters protein-protein interactions. Apart from GBSS, it is not yet clear, *in planta*, whether starch synthesizing enzymes are active under their soluble forms, when physically bound to the granule, or both. Developing milder extraction methods for the isolation of the starch-bound proteome would contribute answering this question by providing information on the maintenance of protein-protein complexes within the starch matrix.

Recent investigation showed that GBSS is actively targeted to starch granules by PTST1 through reciprocal coiled-coil interaction and glucan binding via a carbohydrate binding module (CBM) present in the N-terminal region of PTST1 (Seung et al., 2015). PTST1 was identified in our analysis thus confirming physical interaction with starch in potato. The question whether PTST1 interacts with GBSS in potato tubers as well as with other starch metabolic proteins arises and requires further investigation. Several other starch-binding proteins contain coiled-coil domains that could drive physical association with PTST1. Noteworthy, this is the case of SS4 that was recently shown to interact with PTST2, a protein homologous to PTST1 and an ortholog of the rice FLO6, in Arabidopsis leaves (Seung et al., 2017). The proposed function of PTST2 is redundant with that of PTST3 for binding and delivering maltooligosaccharides to SS4 during starch initiation (Seung et al., 2017). These findings point out that protein-protein and protein-glucan interactions are actively controlled both at the steps of starch initiation and granule growth. Interestingly, whereas traces of SS4 were detected in the present work, neither PTST2/FLO6 nor PTST3 orthologs were observed. However, plastid physiology, and likely starch metabolism, are substantially different between Arabidopsis leaf chloroplasts and potato tuber amyloplasts. Moreover, while the function of SS4 in the initiation and the control of starch granule number



have been studied extensively in Arabidopsis, its function in storage starch synthesis of the potato tuber remains to be investigated.

Proteins of the starch degradation pathway were also identified in this work. These comprise PWD, GWD, SEX4, ISA3, LSF2 as well as the recently discovered degradation regulators ESV1 and LESV. Interaction of PWD and GWD with the starch granule is well documented (Ritte et al., 2000; Kötting et al., 2005; Stensballe et al., 2008). Similarly, the present work confirms that both SEX4 and LSF2, already observed in association with Arabidopsis starch, bind to the starch granules in potato tubers (Santelia et al., 2011). On the other hand, to our knowledge, this is the first report on the physical association of ISA3 with starch. Like ISA1, ISA3 hydrolyzes  $\alpha$ -1,6 linkages of amylopectin. However, whereas ISA1 participate to starch synthesis in complex with ISA2, ISA3 is involved in transitory starch breakdown in leaves of Arabidopsis and rice (Wattebled et al., 2005; Yun et al., 2011). The function of ISA3 in amyloplasts is not yet fully

elucidated as illustrated by the phenotype of an *isa3* mutant of rice (Yun et al., 2011). In the latter, plastid morphogenesis is altered as well as the morphology of starch granules. The authors proposed that ISA3 impacts plastid division in rice leading to pleomorphic amyloplasts and starch granules when the gene is knocked out (Yun et al., 2011). Simultaneous antisense inhibition of ISA1, ISA2, and ISA3 in potato do not alter leaf starch metabolism (Ferreira et al., 2017). On the other hand, starch accumulation in tubers of the latter lines is significantly reduced (Ferreira et al., 2017). Furthermore, granule morphology is altered with an increase in the proportion of the smallest starch granules, which was already observed in antisense lines where the expression of only ISA1 and ISA2 is reduced (Bustos et al., 2004; Ferreira et al., 2017). Thus, while the function of the ISA1/ISA2 complex in starch synthesis in plants including potato is supported by several lines of evidence, elucidating that of ISA3 would require further studies such as the characterization of knock out mutants (Hussain et al., 2003). One possible explanation for the presence of ISA3 within the starch granule is that the protein interacts with a CBM-containing protein similar to ISA1 interacting with FLO6 in the rice endosperm (Peng et al., 2014). The authors proposed that while not directly binding to starch, rice ISA1 binds to FLO6 which, in turn, bridges between starch and the enzyme (Peng et al., 2014).

The finding of an isoform of PHS2, usually referred to as the cytosolic starch phosphorylase, in the present analysis was intriguing. This starch phosphorylase is strongly interacting with a complex cytosolic heteroglycan in several plant species comprising potato (Yang and Steup, 1990; Fettke et al., 2008). This heteroglycan is a glucosyl acceptor for PHS2 activity in starch catabolism consistent with a cytosolic localization of the enzyme, which was confirmed by immunofluorescence studies of pea leaves as well as potato leaves overexpressing PHS2 from *Vicia faba* (Conrads et al., 1986; Fettke et al., 2004, 2005). On the other hand, the protein was identified in the chloroplast stroma of Arabidopsis by proteomics, suggesting a dual-targeting of the enzyme (Zybailov et al., 2008; Helm et al., 2014). Interestingly, contrary to Arabidopsis, the potato genome contains two isoforms of PHS2, namely, PHS2a and PHS2b (Van Harsselaar et al., 2017). Our observation of starch-bound PHS2a corroborates that this isoform is not restricted to the cytosol. However, it is also possible that the enzyme, known to strongly interact with polysaccharides was copurified with starch during the experimental procedure. Unfortunately, our MaxLFQ analysis following surface treatment of starch did not allow quantifying this particular isoform and consequently to investigate the type of interaction between the enzyme and the starch granule.

Identification of the two isoforms of starch synthase, SS5 and SS6, in this study was particularly fascinating. Indeed, while both isoforms were already observed in the genomes of numerous plant species, evidence for their participation to starch metabolism are sporadic (Deschamps et al., 2008; Yu et al., 2012; Liu et al., 2015; Van Harsselaar et al., 2017). These only relied on the gene expression profiles during grain filling in maize and potato tuberization for SS5 and SS6, respectively (Yu et al., 2012;

Liu et al., 2015). Our study is the first report on the presence of these isoforms in association with the starch granule, arguing for a function in starch metabolism. Phylogenetic analysis showed that both proteins have a monophyletic origin and belong to Group-B starch synthase similar to SS3 and SS4 (Letierrier et al., 2008). Although a detailed phylogenomic analysis of SS6 remains to be performed, the gene appears to be conserved among dicots with the noticeable exception of Arabidopsis. The Arabidopsis genome contains a truncated version of SS5 and no SS6 gene was observed in this species. The potato SS5 sequence lacks the two X-X-G-G-L motifs suggesting that this protein is not catalytically active. However, one cannot exclude that this isoform diverged functionally during plant evolution and carries a specific enzymatic activity that remains to be characterized. On the other hand, it is also possible that, while devoid of any enzymatic activity, the protein is responsible for the maintenance of protein–protein, protein–glucan, or glucan–glucan interactions, as suggested for ESV1 and LESV (Feike et al., 2016). Contrarily to SS5, SS6 contains the catalytic amino acid residues of starch synthases suggesting a function in the elongation of  $\alpha$ -1,4 starch polysaccharides (Cao et al., 1999). This isoform is most closely related to SS4, which is involved in starch initiation and the control of granule number and shape in Arabidopsis (Roldan et al., 2007). Interestingly, SS6 differs from SS4 by the absence of the long N-terminal extension but both proteins share the catalytic glucosyltransferase domain. The former is known to modulate the subcellular localization of SS4 as well as conditioning granule shape while the latter is responsible for starch initiation, thus controlling the number of starch granules in Arabidopsis (Raynaud et al., 2016; Lu et al., 2017). Potato amyloplasts contain a single, relatively huge, spherical starch granule, similar to *ss4-* leaf chloroplasts when compared to wild type plants (usually accumulating 5–7 granules per chloroplast) (Roldan et al., 2007). It is thus tempting to speculate that the function of SS6 supplants that of SS4 in starch initiation in potato tubers leading to the initiation of a single, rounded-shape starch granule, analogous to those of Arabidopsis *ss4-* mutants. Nevertheless, deciphering the functions of SS5 and SS6 will require further investigations including the phenotypic analysis of mutant plants.

In addition to starch metabolic enzymes, the present study revealed the presence of several protease inhibitors belonging to the PKPI or PINII family. The PKPI family comprises three groups, namely A, B, and C. PKPI-A group contains inhibitors of aspartic proteases and chymotrypsin while PKPI-B group includes dual inhibitors of serine proteases, trypsin and chymotrypsin (Heibges et al., 2003). On the other hand, PKPI-C group comprises inhibitors of cysteine proteases and inhibitors of the invertase (Heibges et al., 2003; Poltronieri et al., 2012). However, a high functional polymorphism between PKPIs regardless of the group they belong to was previously suggested (Heibges et al., 2003). On the other hand, PINII proteins are serine protease inhibitors containing eight cysteine residues forming disulfide bounds and are involved in plant defense against pathogens (Turra and Lorito, 2011). Noteworthy, the contents of these

proteins were drastically reduced after treating starch granules with SDS or thermolysin indicating weak interaction with starch polysaccharides. These protease inhibitors may either specifically interact with starch or being co-purified during granule isolation. In the former case, one could hypothesize that they protect starch-bound hydrolyzing enzymes from proteolysis for efficient starch degradation concomitant with tuber sprouting. However, the nature of the interaction between starch and the protease inhibitors remains to be characterized, which could help elucidating their potential function in starch metabolism.

Contrary to the protease inhibitors, all starch metabolism proteins that could be quantified in this study remained attached to starch treated with SDS or thermolysin. This result indicate that these proteins are entrapped within the starch matrix. This was also the case of GPX, a glutathione peroxidase and TRX, a thioredoxin protein homologous to Arabidopsis TRX M. The latter protein is involved in the ferredoxin/thioredoxin pathway that regulates metabolic processes in function of light (Buchanan, 1984). Thioredoxin M4 was shown to participate to the redox regulation of starch metabolism in Arabidopsis by activating SS1 and BAM1 (Valerio et al., 2010; Skryhan et al., 2015). On the other hand, glutathione peroxidases are thought to contribute to cellular redox homeostasis during plant-cell stress responses (Noctor et al., 2012). Overall, both proteins are involved in redox regulation of enzyme activities. Their presence in the starch granule suggests that they play a role in regulating starch metabolism and open new perspectives to studying these aspects in storage organs of one of the most cultivated crops.

## AUTHOR CONTRIBUTIONS

AC and MF cultivated and harvested plants. JV and SD prepared the samples. SH and FB performed the experiments and analyzed the data. NS, CT, and CR conceived and designed the study. NS wrote the manuscript.

## FUNDING

This research was supported by ANR grants through the PDOC-2013 and the “Investment for the Future” programs (ANR-13-PDOC-0030-01 and ANR-10-IEED-0004-01, respectively) as well as the Hauts-de-France region through the CPER “Ali Biotech”.

## ACKNOWLEDGMENTS

The authors thank Christophe D’Hulst and Fabrice Wattedbled for fruitful discussions and valuable suggestions.

## SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: <https://www.frontiersin.org/articles/10.3389/fpls.2018.00746/full#supplementary-material>



## REFERENCES

- Ainseba-Chirani, N., Dembahri, Z., Tokarski, C., Rolando, C., and Benmouna, M. (2011). Newly designed polyacrylamide/dextran gels for electrophoresis protein separation: synthesis and characterization. *Polym. Int.* 60, 1024–1029. doi: 10.1002/pi.3035
- Ball, S. G., Van De Wal, M. H. B. J., and Visser, R. G. F. (1998). Progress in understanding the biosynthesis of amylose. *Trends Plant Sci.* 3, 462–467. doi: 10.1016/S1360-1385(98)01342-9
- Borén, M., Larsson, H., Falk, A., and Jansson, C. (2004). The barley starch granule proteome—internalized granule polypeptides of the mature endosperm. *Plant Sci.* 166, 617–626. doi: 10.1016/j.plantsci.2003.10.028
- Buchanan, B. B. (1984). The ferredoxin/thioredoxin system: a key element in the regulatory function of light in photosynthesis. *Bioscience* 34, 378–383. doi: 10.2307/1309730
- Buléon, A., Colonna, P., Planchot, V., and Ball, S. (1998). Starch granules: structure and biosynthesis. *Int. J. Biol. Macromol.* 23, 85–112. doi: 10.1016/S0141-8130(98)00040-3
- Bustos, R., Fahy, B., Hylton, C. M., Seale, R., Nebane, N. M., Edwards, A., et al. (2004). Starch granule initiation is controlled by a heteromultimeric isoamylase in potato tubers. *Proc. Natl. Acad. Sci. U.S.A.* 101, 2215–2220. doi: 10.1073/pnas.0305920101
- Cao, H., Imparl-Radosevich, J., Guan, H., Keeling, P. L., James, M. G., and Myers, A. M. (1999). Identification of the soluble starch synthase activities of maize endosperm. *Plant Physiol.* 120, 205–216. doi: 10.1104/pp.120.1.205
- Comparot-Moss, S., Kötting, O., Stettler, M., Edner, C., Graf, A., Weise, S. E., et al. (2010). A putative phosphatase, LSF1, is required for normal starch turnover in Arabidopsis leaves. *Plant Physiol.* 152, 685–697. doi: 10.1104/pp.109.148981
- Conrads, J., Van Berkel, J., Schächtele, C., and Steup, M. (1986). Non-chloroplast  $\alpha$ -1,4-glucan phosphorylase from pea leaves: characterization and in situ localization by indirect immunofluorescence. *Biochim. Biophys. Acta* 882, 452–463. doi: 10.1016/0304-4165(86)90270-9
- Cox, J., Hein, M. Y., Luber, C. A., Paron, I., Nagaraj, N., and Mann, M. (2014). Accurate proteome-wide label-free quantification by delayed normalization and maximal peptide ratio extraction, termed MaxLFQ. *Mol. Cell. Proteomics* 13, 2513–2526. doi: 10.1074/mcp.M113.031591
- Cox, J., and Mann, M. (2008). MaxQuant enables high peptide identification rates, individualized p.p.b.-range mass accuracies and proteome-wide protein quantification. *Nat. Biotechnol.* 26, 1367–1372. doi: 10.1038/nbt.1511
- Delatte, T., Trevisan, M., Parker, M. L., and Zeeman, S. C. (2005). Arabidopsis mutants *Atisa1* and *Atisa2* have identical phenotypes and lack the same multimeric isoamylase, which influences the branch point distribution of amylopectin during starch synthesis. *Plant J.* 41, 815–830. doi: 10.1111/j.1365-313X.2005.02348.x
- Denyer, K., Sidebottom, C., Hylton, C. M., and Smith, A. M. (1993). Soluble isoforms of starch synthase and starch-branching enzyme also occur within starch granules in developing pea embryos. *Plant J.* 4, 191–198. doi: 10.1046/j.1365-313X.1993.04010191.x
- Deschamps, P., Moreau, H., Worden, A. Z., Dauvillée, D., and Ball, S. G. (2008). Early gene duplication within chloroplastida and its correspondence with relocation of starch metabolism to chloroplasts. *Genetics* 178, 2373–2387. doi: 10.1534/genetics.108.087205
- Erde, J., Loo, R. R. O., and Loo, J. A. (2014). Enhanced FASP (eFASP) to increase proteome coverage and sample recovery for quantitative proteomic experiments. *J. Proteome Res.* 13, 1885–1895. doi: 10.1021/pr4010019
- Feike, D., Seung, D., Graf, A., Bischof, S., Ellick, T., Coiro, M., et al. (2016). The starch Granule-associated protein EARLY STARVATION1 is required for the control of starch degradation in *Arabidopsis thaliana* leaves. *Plant Cell* 28, 1472–1489. doi: 10.1105/tpc.16.00011
- Ferreira, S. J., Senning, M., Fischer-Stettler, M., Streb, S., Ast, M., Neuhaus, H. E., et al. (2017). Simultaneous silencing of isoamylases ISA1, ISA2 and ISA3 by multi-target RNAi in potato tubers leads to decreased starch content and an early sprouting phenotype. *PLoS One* 12:e0181444. doi: 10.1371/journal.pone.0181444
- Fettke, J., Eckermann, N., Poeste, S., Pauly, M., and Steup, M. (2004). The glycan substrate of the cytosolic (Pho 2) phosphorylase isozyme from *Pisum sativum* L.: identification, linkage analysis and subcellular localization. *Plant J.* 39, 933–946. doi: 10.1111/j.1365-313X.2004.02181.x
- Fettke, J., Nunes-Nesi, A., Alpers, J., Szkop, M., Fernie, A. R., and Steup, M. (2008). Alterations in cytosolic glucose-phosphate metabolism affect structural features and biochemical properties of starch-related heteroglycans. *Plant Physiol.* 148, 1614–1629. doi: 10.1104/pp.108.127969
- Fettke, J., Poeste, S., Eckermann, N., Tiessen, A., Pauly, M., Geigenberger, P., et al. (2005). Analysis of cytosolic heteroglycans from leaves of transgenic potato (*Solanum tuberosum* L.) plants that under- or overexpress the Pho 2 phosphorylase isozyme. *Plant Cell Physiol.* 46, 1987–2004. doi: 10.1093/pcp/pci214
- Gámez-Arjona, F. M., Li, J., Raynaud, S., Baroja-Fernández, E., Muñoz, F. J., Ovecka, M., et al. (2011). Enhancing the expression of starch synthase class IV results in increased levels of both transitory and long-term storage starch. *Plant Biotechnol. J.* 9, 1049–1060. doi: 10.1111/j.1467-7652.2011.00626.x
- Grimaud, F., Rogniaux, H., James, M. G., Myers, A. M., and Planchot, V. (2008). Proteome and phosphoproteome analysis of starch granule-associated proteins from normal maize and mutants affected in starch biosynthesis. *J. Exp. Bot.* 59, 3395–3406. doi: 10.1093/jxb/ern198
- Heibges, A., Salamini, F., and Gebhardt, C. (2003). Functional comparison of homologous members of three groups of Kunitz-type enzyme inhibitors from potato tubers (*Solanum tuberosum* L.). *Mol. Genet. Genomics* 269, 535–541. doi: 10.1007/s00438-003-0861-z
- Hejazi, M., Fettke, J., Kötting, O., Zeeman, S. C., and Steup, M. (2010). The Laforin-like dual-specificity phosphatase SEX4 from Arabidopsis hydrolyzes both C6- and C3-phosphate esters introduced by starch-related dikinases and thereby affects phase transition of alpha-glucans. *Plant Physiol.* 152, 711–722. doi: 10.1104/pp.109.149914
- Helm, S., Dobritzsch, D., Rödiger, A., Agne, B., and Baginsky, S. (2014). Protein identification and quantification by data-independent acquisition and multi-parallel collision-induced dissociation mass spectrometry (MSE) in the chloroplast stroma proteome. *J. Proteomics* 98, 79–89. doi: 10.1016/j.jprot.2013.12.007
- Hirsch, C. D., Hamilton, J. P., Childs, K. L., Cepela, J., Crisovan, E., Vaillancourt, B., et al. (2014). Spud DB: a resource for mining sequences, genotypes, and phenotypes to accelerate potato breeding. *Plant Genome* 7:1. doi: 10.3835/plantgenome2013.12.0042
- Hussain, H., Mant, A., Seale, R., Zeeman, S., Hinchliffe, E., Edwards, A., et al. (2003). Three isoforms of isoamylase contribute different catalytic properties for the debranching of potato glucans. *Plant Cell* 15, 133–149. doi: 10.1105/tpc.006635
- Ishikawa, A., Ohta, S., Matsuoka, K., Hattori, T., and Nakamura, K. (1994). A family of potato genes that encode Kunitz-type proteinase inhibitors: structural comparisons and differential expression. *Plant Cell Physiol.* 35, 303–312.
- Jobling, S. (2004). Improving starch for food and industrial applications. *Curr. Opin. Plant Biol.* 7, 210–218. doi: 10.1016/j.pbi.2003.12.001
- Kötting, O., Pusch, K., Tiessen, A., Geigenberger, P., Steup, M., and Ritte, G. (2005). Identification of a novel enzyme required for starch metabolism in Arabidopsis leaves. the phosphoglucan, water dikinase. *Plant Physiol.* 137, 242–252. doi: 10.1104/pp.104.055954
- Leterrier, M., Holappa, L. D., Broglie, K. E., and Beckles, D. M. (2008). Cloning, characterisation and comparative analysis of a starch synthase IV gene in wheat: functional and evolutionary implications. *BMC Plant Biol.* 8:98. doi: 10.1186/1471-2229-8-98
- Liu, F., Makhmoudova, A., Lee, E. A., Wait, R., Emes, M. J., and Tetlow, I. J. (2009). The amylose extender mutant of maize conditions novel protein-protein interactions between starch biosynthetic enzymes in amyloplasts. *J. Exp. Bot.* 60, 4423–4440. doi: 10.1093/jxb/erp297
- Liu, H., Yu, G., Wei, B., Wang, Y., Zhang, J., Hu, Y., et al. (2015). Identification and phylogenetic analysis of a novel starch synthase in maize. *Front. Plant Sci.* 6:1013. doi: 10.3389/fpls.2015.01013
- Lu, K.-J., Pfister, B., Jenny, C., Eicke, S., and Zeeman, S. C. (2017). Distinct functions of STARCH SYNTHASE 4 Domains in starch granule formation. *Plant Physiol.* 176, 566–581. doi: 10.1104/pp.17.01008
- Mu-Forster, C., Huang, R., Powers, J. R., Harriman, R. W., Knight, M., Singletary, G. W., et al. (1996). Physical association of starch biosynthetic enzymes with starch granules of maize endosperm (Granule-Associated forms of starch

- synthase I and starch branching enzyme II). *Plant Physiol.* 111, 821–829. doi: 10.1104/pp.111.3.821
- Noctor, G., Mhamdi, A., Chaouch, S., Han, Y. I., Neukermans, J., Marquez-Garcia, B., et al. (2012). Glutathione in plants: an integrated overview. *Plant Cell Environ.* 35, 454–484. doi: 10.1111/j.1365-3040.2011.02400.x
- Peng, C., Wang, Y., Liu, F., Ren, Y., Zhou, K., Lv, J., et al. (2014). FLOURY ENDOSPERM6 encodes a CBM48 domain-containing protein involved in compound granule formation and starch synthesis in rice endosperm. *Plant J.* 77, 917–930. doi: 10.1111/tpj.12444
- Peng, M., Gao, M., Båga, M., Hucl, P., and Chibbar, R. N. (2000). Starch-Branching enzymes preferentially associated with a-type starch granules in wheat endosperm. *Plant Physiol.* 124, 265–272. doi: 10.1104/pp.124.1.265
- Poltronieri, P., Liu, S., Cimaglia, F., Santino, A., and Wang, Y. (2012). Characterization of Kunitz-type inhibitor B1 performance using protein chips and AFM. *Sens. Actuators B* 168, 231–237. doi: 10.1016/j.snb.2012.04.013
- Raynaud, S., Ragel, P., Rojas, T., and Mérida, Á. (2016). The N-terminal Part of *Arabidopsis thaliana* starch synthase 4 determines the localization and activity of the enzyme. *J. Biol. Chem.* 291, 10759–10771. doi: 10.1074/jbc.M115.698332
- Regina, A., Kosar-Hashemi, B., Li, Z., Pedler, A., Mukai, Y., Yamamoto, M., et al. (2005). Starch branching enzyme IIb in wheat is expressed at low levels in the endosperm compared to other cereals and encoded at a non-syntenic locus. *Planta* 222, 899–909. doi: 10.1007/s00425-005-0032-z
- Ritte, G., Heydenreich, M., Mahlow, S., Haebel, S., Kötting, O., and Steup, M. (2006). Phosphorylation of C6- and C3-positions of glucosyl residues in starch is catalysed by distinct dikinases. *FEBS Lett.* 580, 4872–4876. doi: 10.1016/j.febslet.2006.07.085
- Ritte, G., Lorberth, R., and Steup, M. (2000). Reversible binding of the starch-related R1 protein to the surface of transitory starch granules. *Plant J.* 21, 387–391. doi: 10.1046/j.1365-313x.2000.00683.x
- Roldan, I., Wattedled, F., Mercedes Lucas, M., Delvalle, D., Planchot, V., Jimenez, S., et al. (2007). The phenotype of soluble starch synthase IV defective mutants of *Arabidopsis thaliana* suggests a novel function of elongation enzymes in the control of starch granule formation. *Plant J.* 49, 492–504. doi: 10.1111/j.1365-313X.2006.02968.x
- Rongine De Fekete, M. A., Leloir, L. F., and Cardini, C. E. (1960). Mechanism of Starch Biosynthesis. *Nature* 187, 918–919. doi: 10.1038/187918a0
- Rydberg, U., Andersson, L., Andersson, R., Åman, P., and Larsson, H. (2001). Comparison of starch branching enzyme I and II from potato. *Eur. J. Biochem.* 268, 6140–6145. doi: 10.1046/j.0014-2956.2001.02568.x
- Santelia, D., Kötting, O., Seung, D., Schubert, M., Thalmann, M., Bischof, S., et al. (2011). The phosphoglucan phosphatase like sex Four2 dephosphorylates starch at the C3-position in *Arabidopsis*. *Plant Cell* 23, 4096–4111. doi: 10.1105/tpc.111.092155
- Schwall, G. P., Safford, R., Westcott, R. J., Jeffcoat, R., Tayal, A., Shi, Y.-C., et al. (2000). Production of very-high-amylose potato starch by inhibition of SBE A and B. *Nat. Biotechnol.* 18, 551–554. doi: 10.1038/75427
- Seung, D., Boudet, J., Monroe, J., Schreier, T. B., David, L. C., Abt, M., et al. (2017). Homologs of PROTEIN TARGETING TO STARCH control starch granule initiation in *Arabidopsis* Leaves. *Plant Cell* 29, 1657–1677. doi: 10.1105/tpc.17.00222
- Seung, D., Lu, K.-J., Stettler, M., Streb, S., and Zeeman, S. C. (2016). Degradation of glucan primers in the absence of starch synthase 4 disrupts starch granule initiation in *Arabidopsis*. *J. Biol. Chem.* 291, 20718–20728. doi: 10.1074/jbc.M116.730648
- Seung, D., Soyk, S., Coiro, M., Maier, B. A., Eicke, S., and Zeeman, S. C. (2015). PROTEIN TARGETING TO STARCH is required for localising GRANULE-BOUND STARCH SYNTHASE to starch granules and for normal amylose synthesis in *Arabidopsis*. *PLoS Biol.* 13:e1002080. doi: 10.1371/journal.pbio.1002080
- Silver, D. M., Kötting, O., and Moorhead, G. B. G. (2014). Phosphoglucan phosphatase function sheds light on starch degradation. *Trends Plant Sci.* 19, 471–478. doi: 10.1016/j.tplants.2014.01.008
- Skryhan, K., Cuesta-Seijo, J. A., Nielsen, M. M., Marri, L., Mellor, S. B., Glaring, M. A., et al. (2015). The role of cysteine residues in redox regulation and protein stability of *Arabidopsis thaliana* starch synthase 1. *PLoS One* 10:e0136997. doi: 10.1371/journal.pone.0136997
- Stensballe, A., Hald, S., Bauw, G., Blennow, A., and Welinder, K. G. (2008). The amyloplast proteome of potato tuber. *FEBS J.* 275, 1723–1741. doi: 10.1111/j.1742-4658.2008.06332.x
- Subasinghe, R. M., Liu, F., Polack, U. C., Lee, E. A., Emes, M. J., and Tetlow, I. J. (2014). Multimeric states of starch phosphorylase determine protein–protein interactions with starch biosynthetic enzymes in amyloplasts. *Plant Physiol. Biochem.* 83, 168–179. doi: 10.1016/j.plaphy.2014.07.016
- Szydlowski, N., Ragel, P., Raynaud, S., Lucas, M. M., Roldan, I., Montero, M., et al. (2009). Starch granule initiation in *Arabidopsis* requires the presence of either class IV or class III starch synthases. *Plant Cell* 21, 2443–2457. doi: 10.1105/tpc.109.066522
- Tetlow, I. J., Beisel, K. G., Cameron, S., Makhmoudova, A., Bresolin, N., Wait, R., et al. (2008a). Analysis of protein complexes in wheat amyloplast reveals functional interactions among starch biosynthetic enzymes. *Plant Physiol.* 146, 1878–1891. doi: 10.1104/pp.108.116244
- Tetlow, I. J., Beisel, K. G., Cameron, S., Makhmoudova, A., Liu, F., Bresolin, N. S., et al. (2008b). Analysis of protein complexes in wheat amyloplasts reveals functional interactions among starch biosynthetic enzymes. *Plant Physiol.* 146, 1878–1891. doi: 10.1104/pp.108.116244
- Tetlow, I. J., Wait, R., Lu, Z., Akkasaeng, R., Bowsher, C. G., Esposito, S., et al. (2004). Protein phosphorylation in amyloplasts regulates starch branching enzyme activity and protein–protein interactions. *Plant Cell* 16, 694–708. doi: 10.1105/tpc.017400
- Turra, D., and Lorito, M. (2011). Potato type I and II proteinase inhibitors: modulating plant physiology and host resistance. *Curr. Protein Pept. Sci.* 12, 374–385. doi: 10.2174/138920311796391151
- Tyanova, S., Temu, T., and Cox, J. (2016a). The MaxQuant computational platform for mass spectrometry-based shotgun proteomics. *Nat. Protoc.* 11:2301. doi: 10.1038/nprot.2016.136
- Tyanova, S., Temu, T., Sinitcyn, P., Carlson, A., Hein, M. Y., Geiger, T., et al. (2016b). The Perseus computational platform for comprehensive analysis of (prote)omics data. *Nat. Methods* 13, 731–740. doi: 10.1038/nmeth.3901
- Umemoto, T., and Aoki, N. (2005). Single-nucleotide polymorphisms in rice starch synthase IIa that alter starch gelatinisation and starch association of the enzyme. *Funct. Plant Biol.* 32, 763–768. doi: 10.1071/FP04214
- Valerio, C., Costa, A., Marri, L., Issakidis-Bourguet, E., Pupillo, P., Trost, P., et al. (2010). Thioresoxin-regulated -amylase (BAM1) triggers diurnal starch degradation in guard cells, and in mesophyll cells under osmotic stress. *Tree Physiol.* 38, 1–19. doi: 10.1093/jxb/erq288
- Van Harsseelaar, J. K., Lorenz, J., Senning, M., Sonnewald, U., and Sonnewald, S. (2017). Genome-wide analysis of starch metabolism genes in potato (*Solanum tuberosum* L.). *BMC Genomics* 18:37. doi: 10.1186/s12864-016-3381-z
- Wattedled, F., Dong, Y., Dumez, S., Delvallé, D., Planchot, V., Berbezy, P., et al. (2005). Mutants of *Arabidopsis* lacking a chloroplastic isoamylase accumulate phytylglucan and an abnormal form of amylopectin. *Plant Physiol.* 138, 184–195. doi: 10.1104/pp.105.059295
- Wattedled, F., Planchot, V., Dong, Y., Szydlowski, N., Pontoire, B., Devin, A., et al. (2008). Further evidence for the mandatory nature of polysaccharide debranching for the aggregation of semicrystalline starch and for overlapping functions of debranching enzymes in *Arabidopsis* leaves. *Plant Physiol.* 148, 1309–1323. doi: 10.1104/pp.108.129379
- Xing, S., Meng, X., Zhou, L., Mujahid, H., Zhao, C., Zhang, Y., et al. (2016). Proteome profile of starch granules purified from rice (*Oryza sativa*) endosperm. *PLoS One* 11:e0168467. doi: 10.1371/journal.pone.0168467
- Xu, X., Pan, P., Cheng, S., Zhang, B., Mu, D., Ni, P., et al. (2011). Genome sequence and analysis of the tuber crop potato. *Nature* 475, 189–195. doi: 10.1038/nature10158
- Yang, Y., and Steup, M. (1990). Polysaccharide fraction from higher plants which strongly interacts with the cytosolic phosphorylase isozyme. I. Isolation and characterization. *Plant Physiol.* 94, 960–969. doi: 10.1104/pp.94.3.960
- Yu, H., and Wang, T. (2016). Proteomic dissection of endosperm starch granule associated proteins reveals a network coordinating starch biosynthesis and amino acid metabolism and glycolysis in rice endosperms. *Front. Plant Sci.* 7:707. doi: 10.3389/fpls.2016.00707
- Yu, J. W., Choi, J.-S., Upadhyaya, C. P., Kwon, S. O., Gururani, M. A., Nookaraju, A., et al. (2012). Dynamic proteomic profile of potato tuber during its in vitro development. *Plant Sci.* 195, 1–9. doi: 10.1016/j.plantsci.2012.06.007

- Yun, M.-S., Umemoto, T., and Kawagoe, Y. (2011). Rice debranching enzyme Isoamylase3 facilitates starch metabolism and affects plastid morphogenesis. *Plant Cell Physiol.* 52, 1068–1082. doi: 10.1093/pcp/pcr058
- Zhang, J., Xin, L., Shan, B., Chen, W., Xie, M., Yuen, D., et al. (2012). PEAKS DB: De Novo sequencing assisted database search for sensitive and accurate peptide identification. *Mol. Cell. Proteomics* 11:M111.010587. doi: 10.1074/mcp.M111.010587
- Zybilov, B., Rutschow, H., Friso, G., Rudella, A., Emanuelsson, O., Sun, Q., et al. (2008). Sorting signals, N-Terminal modifications and abundance of the chloroplast proteome. *PLoS One* 3:e1994. doi: 10.1371/journal.pone.0001994

**Conflict of Interest Statement:** The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Copyright © 2018 Helle, Bray, Verbeke, Devassine, Courseaux, Facon, Tokarski, Rolando and Szydłowski. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.

## Annexe 2

N: Abundance (molecules/total molecules) [\*10<sup>-6</sup>]

Brut1	Brut2	Brut3	SDS1	SDS2	SDS3	Thermo1	Thermo2	Thermo3	T: Name	T: PGSC+uniprot+solanum
341273	245804	297491	368493	402892	385950	372824	389835	384182	Granule-bound starch synthase 1	PGSC0003DMP400021398
20982.9	25065.8	19382.5	28963.8	17899.3	25386.4	35433.1	29378.6	29612.4	Granule-bound starch synthase 2	PGSC0003DMP400002384
635.007	709.045	687.546	1113.18	735.858	838.833	1048.25	948.647	684.916	Starch-granule-bound R1 protein	PGSC0003DMP400013565
397030	309045	352327	504046	540378	522764	453957	464026	480612	Granule-bound starch synthase 1	PGSC0003DMP400021399
216.927	231.83	272.489	373.546	486.621	277.438	325.447	310.494	188.061	Soluble starch synthase 3	Q43846
904.794	1057.11	1079.37	1426.17	953.781	1403.11	1668.65	1314.98	821.838	Conserved gene of unknown function	PGSC0003DMP400028510
1131.37	2074.79	1867.68	2653.59	1651.77	1757.5	3089.59	2638.49	2056.01	Starch branching enzyme II	Sotub09g011090.1.1
768.967	1136.33	955.467	1407.37	742.623	1364.47	1716.23	1380.85	1031.32	1,4-alpha-glucan branching enzyme	PGSC0003DMP400017627
763.381	1089.39	949.543	1303.48	713.987	1320.08	1634.05	1320.42	930.262	1,4-alpha-glucan-branching enzyme	P30924
63777.9	82157.8	74195.2	8863.14	4388.58	6367.93	19906.6	13602.1	15512.7	Cysteine protease inhibitor 1	PGSC0003DMP400017944
225.654	397.816	199.283	458.33	381.144	485.814	355.806	464.215	394.592	Soluble starch synthase 1	PGSC0003DMP400032328
24251.6	65906.8	50114.5	7835.33	3265.02	4394.85	4656.84	3529.19	1475.23	Kunitz-type tuber invertase inhibitor	PGSC0003DMP400017953
9.1306	18.5612	14.3241	11.9644	14.8086	13.8863	29.0458	8.98749	15.6549	Glucan/water dikinase	PGSC0003DMP400029040
15480.1	29628.4	23085.2	8013.83	1747.58	2479.02	7396.38	6213.08	2134.85	Proteinase inhibitor type-2 P303.51	PGSC0003DMP400008029
162.448	727.982	90.1847	241.678	276.079	287.312	1199.93	395.061	188.758	Peptidyl-prolyl cis-trans isomerase	PGSC0003DMP400009292
12.4384	9.81778	8.36602	9.73442	1.56946	13.9152	14.7532	12.2115	3.36462	Alpha-1,4 glucan phosphorylase L-1 isozyme	P04045;PGSC0003DMP400013702
760.31	566.89	741.547	1128.39	512.682	1138.33	956.58	1031.05	788.938	Conserved gene of unknown function	PGSC0003DMP400053308
48378.1	106338	89259.9	9496.86	3639.8	7136.68	25778.4	16987.9	17850.4	Cysteine protease inhibitor 9	PGSC0003DMP400017952
128.689	117.158	80.9254	176.661	39.5202	103.726	151.133	117.841	85.752	Conserved gene of unknown function	PGSC0003DMP400017605
6324.65	6779.84	5919.29	10256.4	4305.72	7035.73	10160.9	7979.26	8723.44	Dual specificity protein phosphatase	PGSC0003DMP400050640
6448	5495.75	5594.52	670.699	200.695	556.262	1844.83	1741.27	1257.95	Aspartic protease inhibitor 8	PGSC0003DMP400016822
235.77	268.501	321.533	530.796	91.2938	423.373	631.451	252.424	296.499	Protein tyrosine phosphatase	PGSC0003DMP400047519
224.236	247.897	224.792	349.187	132.877	257.784	508.671	172.021	149.12	Protein tyrosine phosphatase	PGSC0003DMP400026742
13944.7	14790.6	9174.53	1535.03	404.122	1029.5	1646.2	2947.68	2506.6	Aspartic protease inhibitor 5	PGSC0003DMP400016824
419.435	265.431	719.25	427.404	817.521	413.102	269.255	406.119	386.601	Starch synthase VI	PGSC0003DMP400023948
707.483	476.533	532.66	982.402	574.818	641.462	921.976	762.032	685.685	Starch synthase VI	PGSC0003DMP400023947
809.422	1013.17	791.512	1074.77	1780.26	1385.78	1763.63	1051.4	1152.42	Thioredoxin	PGSC0003DMP400021964
17142.2	20274.8	9419.86	1902.27	848.678	1066.66	57.924	2659.48	2578.51	Serine protease inhibitor 7	PGSC0003DMP400017933
2779.32	24288.6	18211.9	3106.35	1092.1	4333.44	2270.14	1832.68	396.774	Proteinase inhibitor type-2 CEVI57	PGSC0003DMP400008030
2736.19	2289.12	2200.73	957.35	234.651	660.479	511.189	552.58	396.774	Proteinase inhibitor type-2 CEVI57	PGSC0003DMP400026811
161.779	99.6892	161.032	176.755	185.128	295.269	121.49	146.465	87.9774	Glutathione peroxidase	PGSC0003DMP400045625
123.163	261.139	212.517	184.662	107.46	247.055	393.972	402.095	76.7744	Isoamylase isoform 3	PGSC0003DMP400012889

49.3878	61.7575	81.7915	155.983	26.5788	44.7608	95.2705	131.476	15.5489	Glutaredoxin S12	PGSC0003DMP400034819
9.69367	7.79	3.73725	7.16067	2.79022	9.80503	7.66051	7.71993	0	Alpha-1,4 glucan phosphorylase L-2 isozyme	PGSC0003DMP400049371
296.395	1984.93	0	0	216.119	0	0	0	0	Ci21A protein	PGSC0003DMP400011775
4822.57	4661.92	4429.53	6925.36	2081.49	5186.58	9106.48	6736.23	6453.42	Starch synthase VI	PGSC0003DMP400023950
6840.98	17341.9	13038.2	1830.89	456.076	631.179	815.742	1777.93	259.764	Cysteine protease inhibitor 1	PGSC0003DMP400017942
12.5385	0	0	0	0	0	0	0	0	Basic 7S globulin 2 small subunit	PGSC0003DMP400015797
0	5.37251	0	0	4.2155	0	270.336	149.249	110.125	Actin	PGSC0003DMP400052766
9.4507	98.0823	0	48.4924	57.9912	0	218.855	89.4429	152.007	Grx_S14-glutaredoxin subgroup II	PGSC0003DMP400024119
0	0	0	0	0	0	0	0	0	Chaperonin Cpn60	PGSC0003DMP400016110
7529.22	11427.6	5706.15	1294.15	461.201	1070.06	13178.5	9910.05	10233.3	Cysteine protease inhibitor 1	PGSC0003DMP400017939
99.6976	211.21	134.39	207.995	21.7489	159.586	370.147	175.94	29.5517	Isoamylase isoform 3	PGSC0003DMP400012890
99.8616	64.9313	67.2614	0	3.4918	0	127.265	54.6614	33.9855	Chloroplast methionine sulfoxide reductase B2	PGSC0003DMP400046532
0	856.489	538.981	1067.99	489.248	0	1387.28	370.478	0	Histone H2A	PGSC0003DMP400015710
50.3317	39.4029	51.9787	0	0	0	0	0	0	Stem 28 kDa glycoprotein	PGSC0003DMP400028845
214.819	244.609	110.592	0	126.026	395.846	510.444	204.352	0	Thioredoxin M3, chloroplastic	PGSC0003DMP400044905
0	0	74.2597	0	0	0	0	0	0	Osmotin	PGSC0003DMP400005467
0	0	84.868	0	0	0	0	0	0	Osmotin	PGSC0003DMP400023388
15.214	10.3333	9.62655	0	0	0	0	0	0	Serine/threonine-protein kinase PBS1	PGSC0003DMP400037016
0	263.482	0	0	0	0	0	0	0	Histone H4	PGSC0003DMP400052913
0	0	0	0	0	0	487.853	0	0	Oligopeptide transporter OPT family	PGSC0003DMP400006514
0	0	0	0	0	0	0	0	0	Thaumatococcus	PGSC0003DMP400007584
0	0	127.75	0	72.4597	0	0	0	22.7247	P-type trefoil	PGSC0003DMP400019537
0	0	0	0	0	0	0	0	0	Cytochrome c	PGSC0003DMP400043115
0	0	0	0	0	0	2.93604	189.004	0	Ankyrin repeat-containing protein	PGSC0003DMP400052900
0	1.30515	0	0	0	5.79476	0	0	1.40413	ATP binding protein	PGSC0003DMP400065036
0	0	0	0	0	0	0	0	0	Rrm/rnp domain	PGSC0003DMP400022315
0	0	0	0	0	0	0	0	0	Vesicle-associated membrane protein	PGSC0003DMP400035594
0	0	0	0	0	0	0	0	0	N-acetyl-glutamate synthase	PGSC0003DMP400037112
0	0	0	0	0	0	0	0	0	Eukaryotic translation initiation factor 3 subunit	PGSC0003DMP400049689
121.897	74.5085	80.8734	243.506	63.6064	286.045	236.233	136.248	116.858	F-box protein	PGSC0003DMP400006529
9596.59	12420.2	7756.39	17903.1	3654.76	9425.6	18063.1	24680	24439.6	Conserved gene of unknown function	PGSC0003DMP400025621
127.978	65.1953	135.36	0	37.8742	0	124.423	247.411	204.799	Acyl carrier protein	PGSC0003DMP400053570
31.3195	20.6618	17.862	48.8289	22.9671	32.1954	45.9521	41.895	9.16349	Uncharacterized aarF domain-containing protein kinase	PGSC0003DMP400044471
158.588	536.445	737.225	1845.78	663.181	733.34	673.101	0	0	Hydroxyproline-rich glycoprotein family protein	PGSC0003DMP400023262
0	3.81761	41.2635	0	0	0	0	10.8781	1.79706	Thioredoxin peroxidase	PGSC0003DMP400049086

98.0742	64.996	79.4881	0	25.8593	150.122	160.713	108.454	0	Acyl carrier protein	PGSC0003DMP400018843
6.93963	7.20195	11.2984	0	0.966623	0	0	0	0	NAD-dependent glyceraldehyde 3-P dehydrogenase	PGSC0003DMP400046827
214.979	292.562	131.28	0	0	0	0	0	0	Glucan endo-1,3-beta-D-glucosidase	PGSC0003DMP400051976
0	0	0.512782	0	0	0	0	0	0	Starch synthase IV	PGSC0003DMP400014590
448.56	304.674	72.6148	0	0	0	0	0	0	Conserved gene of unknown function	PGSC0003DMP400050414
4.06816	0	2.23615	0	0	0	0	0	0	Catalase isozyme 3	PGSC0003DMP400017481
4.04606	148.754	33.3055	0	0	0	24.8434	0	0	MIS12 homologue	PGSC0003DMP400042399
182.248	47.5558	104.33	185.509	12.97	5.14545	0	0	0.419615	Exostosin family protein	PGSC0003DMP400003994
0	0	0	0	0	0	737.039	478.494	609.301	Conserved gene of unknown function	PGSC0003DMP400007875
0	62.2117	29.9965	0	0	0	0	0	0	Patatin-2-Kuras 4	PGSC0003DMP400024808
0	0	0	0	0	0	0	0	0	Peroxiredoxin	PGSC0003DMP400015817
0	3.79546	3.70882	9.54524	1.72033	0	9.20063	9.19538	0	U-box domain-containing protein 42	PGSC0003DMP400034136
0	0	0	0	0	0	0	0	0	Pentatricopeptide repeat-containing protein	PGSC0003DMP400049973
5.70243	21.0375	16.6112	55.7836	0	35.5006	102.371	69.6935	41.8505	Conserved gene of unknown function	PGSC0003DMP400053070
0	0	0	0	0	0	0	0	0	Conserved gene of unknown function	PGSC0003DMP400056800
0	11.1498	0	0	0	0	0	0	0	Gene of unknown function	PGSC0003DMP400068277
0	0	0	0	0	0	0	0	0	Isoform 2 of Probable galacturonosyltransferase 13	PGSC0003DMP400002613
0	0	0	0	0	0	0	0	0	Serine/threonine kinase	PGSC0003DMP400010170
0	0	0	0	0	0	0	0	0	Fimbrin	PGSC0003DMP400011407
0	0	0	0	0	0	0	0	0	Fe-S metabolism associated domain-containing protein	PGSC0003DMP400020796
0	0	0	0	0	0	0	0	0	Proteinase inhibitor type-2 CEVI57	PGSC0003DMP400026806
0	0	0	0	0	0	0	0	0	Gene of unknown function	PGSC0003DMP400027395
0	0	0	0	0	0	0	0	0	Transcription factor bHLH49	PGSC0003DMP400034158
0	0	0	0	0	0	0	0	0	Gene of unknown function	PGSC0003DMP400034456
0	0	0	0	0	0	0	0	0	Phosphoglycerate/bisphosphoglycerate mutase	PGSC0003DMP400049759

Article

# Intra-Sample Heterogeneity of Potato Starch Reveals Fluctuation of Starch-Binding Proteins According to Granule Morphology

Stanislas Helle <sup>1,2,3</sup>, Fabrice Bray <sup>1</sup> , Jean-Luc Putaux <sup>4</sup> , Jérémy Verbeke <sup>1</sup>, Stéphanie Flament <sup>1</sup>, Christian Rolando <sup>1</sup>, Christophe D'Hulst <sup>2</sup> , and Nicolas Szydlowski <sup>1,2,\*</sup>

<sup>1</sup> Univ. Lille, CNRS, USR 3290—MSAP—Miniaturisation pour la Synthèse, l'Analyse et la Protéomique, F-59000 Lille, France

<sup>2</sup> Univ. Lille, CNRS, UMR 8576—UGSF—Unité de Glycobiologie Structurale et Fonctionnelle, F-59000 Lille, France

<sup>3</sup> Florimond Desprez Veuve-et-Fils Ind., Section Biotechnologies, BP 41, 59242 Cappelle en Pévèle, France

<sup>4</sup> Univ. Grenoble Alpes, CNRS, CERMAV, F-38000 Grenoble, France

\* Correspondence: nicolas.szydlowski@univ-lille.fr; Tel.: +33-320434718

Received: 15 August 2019; Accepted: 29 August 2019; Published: 4 September 2019



**Abstract:** Starch granule morphology is highly variable depending on the botanical origin. Moreover, all investigated plant species display intra-tissular variability of granule size. In potato tubers, the size distribution of starch granules follows a unimodal pattern with diameters ranging from 5 to 100  $\mu\text{m}$ . Several evidences indicate that granule morphology in plants is related to the complex starch metabolic pathway. However, the intra-sample variability of starch-binding metabolic proteins remains unknown. Here, we report on the molecular characterization of size-fractionated potato starch granules with average diameters of  $14.2 \pm 3.7 \mu\text{m}$ ,  $24.5 \pm 6.5 \mu\text{m}$ ,  $47.7 \pm 12.8 \mu\text{m}$ , and  $61.8 \pm 17.4 \mu\text{m}$ . In addition to changes in the phosphate contents as well as small differences in the amylopectin structure, we found that the starch-binding protein stoichiometry varies significantly according to granule size. Label-free quantitative proteomics of each granule fraction revealed that individual proteins can be grouped according to four distinct abundance patterns. This study corroborates that the starch proteome may influence starch granule growth and architecture and opens up new perspectives in understanding the dynamics of starch biosynthesis.

**Keywords:** starch granule; starch; potato; *Solanum tuberosum*; proteomics; amylopectin; amylose

## 1. Introduction

Starch is the main source of calories in human and livestock diets. It is also used in glue, bioplastic, and biofuel industries as well as a food additive. Starch is composed of two glucose polymers, amylose and amylopectin, and accumulates in the form of cold water-insoluble granules. Amylose and amylopectin are both constituted by  $\alpha$ -1,4 glucan linear segments ramified through  $\alpha$ -1,6 O-glycosidic bonds (also called branching points) [1]. However, while amylose is mainly linear with less than 1% of  $\alpha$ -1,6 linkages, amylopectin is moderately branched and contains 5 to 6% of  $\alpha$ -1,6 bonds [2,3]. Starch also contains minor compounds such as proteins [4], most of which being starch metabolizing enzymes [5].

Starch synthesis involves several families of enzymes. Starch synthases (SSs) are responsible for the elongation of  $\alpha$ -1,4 glucans [6]. On the one hand, the SS1 isoform synthesizes the shortest amylopectin chains [7]. On the other hand, SS2 synthesizes medium-size chains and SS3 elongates the longest chains that compose amylopectin molecules [8,9]. A fourth isoform, SS4, is involved in the initiation of starch synthesis as illustrated by the mutant phenotypes of Arabidopsis plants [10–14].

Two other isoforms of SSs, SS5 and SS6 were found in the potato genome [15]. Moreover, in a previous study, we identified these proteins in association with potato starch by proteomics based on mass spectrometry [16]. However, their putative functions in starch synthesis remain unknown. The granule-bound starch synthase (GBSS) elongates amylose glucans and is noticeably the most abundant protein bound to the starch granule. The starch branching enzymes (SBE1 and SBE2) catalyze the formation of  $\alpha$ -1,6 linkages in starch polymers after rearranging a pre-existing  $\alpha$ -1,4 glucan [17]. Both of them were found in association with the potato starch granule in our previous study [16]. Moreover, we showed the presence of a third isoform annotated as BE3. Further sequence analysis revealed that this enzyme is a variant of BE1. Both enzymes were thus renamed BE1.1 and BE1.2, respectively. Debranching enzymes (DBEs) hydrolyze  $\alpha$ -1,6 linkages. Among them, isoamylase 1 and 2 (ISA1 and ISA2) are involved in amylopectin synthesis, whereas isoamylase 3 (ISA3) participates in starch degradation [18,19]. Among starch-synthesizing enzymes, GBSS, SS1, SS2, SS3, BE1, BE2, ISA1, and ISA2 were found in association with amyloplastic starch in several species such as potato, maize, rice, pea, barley, and wheat [20–27] and ISA 3 was recently identified in potato starch [16].

Other proteins observed in association with the starch granule are involved in starch degradation and its regulation. The glucan water dikinase (GWD) and the phosphoglucan water dikinase (PWD) add phosphate groups at C6 or C3 position of the glucose residues, respectively, to initiate starch degradation [28,29]. The phosphatases starch excess 4 (SEX4) and like SEX four 2 (LSF2) were also identified as chloroplastic proteins and possess a carbohydrate binding module (CBM) [30,31]. SEX4 is able to dephosphorylate phosphoglucans in C3 and C6 positions, whereas LSF2 is only active on C3-phosphoesters [31]. Moreover, SEX4 was observed in the surrounding of starch granules by fluorescent microscopy and LSF2 peptides were identified by SDS-PAGE LC-MS/MS analysis of Arabidopsis starch-bound proteins [31,32]. Recently, other proteins were found associated to the starch granule in Arabidopsis leaves [33–35]. These include protein targeting to starch 1 (PTST1), which addresses GBSS to the granule, early starvation protein 1 (ESV1), and its homologue LESV [33,34]. The two latter proteins are involved in the control of starch degradation. It was proposed that both proteins modulate the molecular organization of starch and consequently affect the glucan accessibility to catabolic enzymes. Noteworthy, we recently confirmed the presence of these proteins in potato starch with the addition of the starch phosphorylases PHS1a, PHS1b, and PHS2 [16]. Furthermore, our study revealed the presence of a thioredoxin (TRx), a glutathione peroxidase (GPx) and a peptidyl-prolyl-cis-trans isomerase (CYP20-20). On the one hand, the Arabidopsis counterpart of TRx is involved in the ferredoxin/thioredoxin pathway that regulates metabolic processes in function of light and is known to activate SS1 and the beta-amylase BAM1 [36,37]. On the other hand, glutathione peroxidases are thought to contribute to cellular redox homeostasis during plant-cell stress responses [38]. Moreover, the relative abundance of starch-associated proteins is likely under the control of regulatory mechanisms related to the initiation and biogenesis of the granules [39].

Starch granules display a broad structural and morphological heterogeneity. This disparity is observed at the level of: (i) the botanical origin; (ii) the organelle and tissue localization within a given species; (iii) the individual cells of a given tissue; (iv) within the same plastid, where distinct granule populations can be observed [40,41]. In potato starch, a high diversity of granule morphology can be observed with diameters ranging from 5 to 100  $\mu$ m [42]. A previous study indicated intra-sample variations of potato starch structure and composition with large granules displaying higher amylose contents and an increased proportion of long amylopectin chains ( $\geq$  DP 37) [40]. Moreover, in the latter study, the smallest granules displayed a higher proportion of small amylopectin chains (DP 6-13) and an A-type X-ray diffraction pattern [43]. In order to evaluate the potential heterogeneity of starch-binding proteins in link with starch structure and morphology, we investigated the size and shape repartition as well as the composition and structure of size-fractionated potato starch granule populations with average diameters of  $14.2 \pm 3.7 \mu$ m,  $24.5 \pm 6.5 \mu$ m,  $47.7 \pm 12.8 \mu$ m, and  $61.8 \pm 17.4 \mu$ m (fractions A, B, C, and D, respectively). Amylose content, starch crystallinity, amylopectin chain length distribution, and phosphate contents were determined as well as the protein content and stoichiometry.

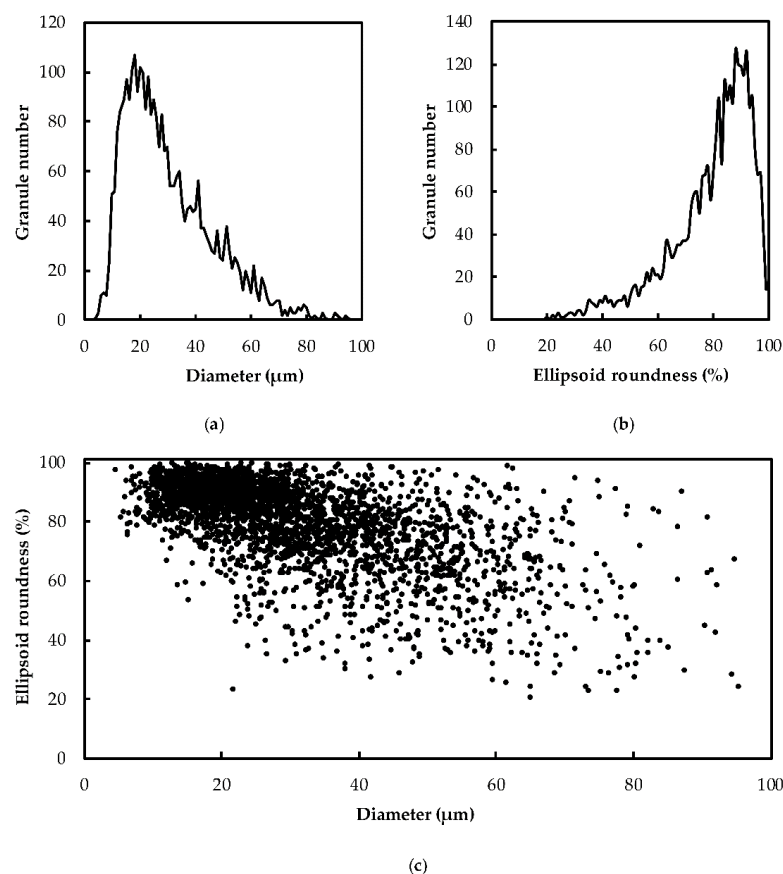


In addition to changes in phosphate content as well as small differences of the amylopectin structure, we found that the starch-binding protein stoichiometry significantly varies according to granule diameter. Quantitative proteomics revealed that individual proteins can be grouped according to four distinct abundance patterns indicating that the association between glucans and individual proteins might be specifically regulated.

## 2. Results

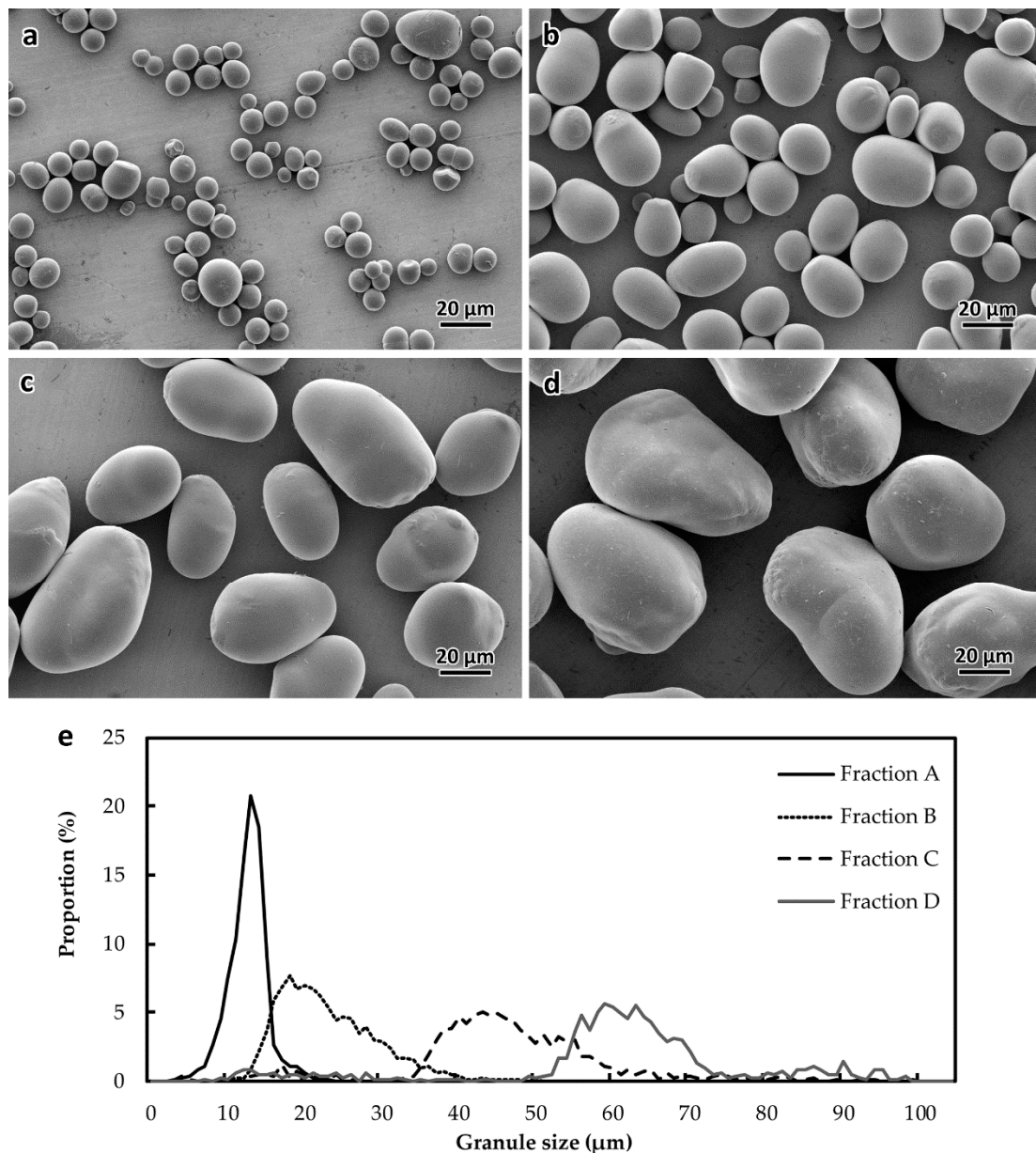
### 2.1. Heterogeneity of Starch Morphology and Size-Fractionation

Starch granules isolated from potato (cv. Monalisa) tubers were analyzed for size and shape distribution by granulomorphometry (Figure 1). The granule shape was estimated by measuring the ellipsoid roundness corresponding to the ratio between the average radius of peripheral circles and the radius of the maximal circle included in the particle. The granule size was characterized by the area diameter that corresponds to the diameter of a perfect circle of the same area than the particle. A high heterogeneity of granule size was observed, with diameters ranging from 4 to 100  $\mu\text{m}$  and most of the granules displayed a diameter between 15 and 40  $\mu\text{m}$  (Figure 1a). Starch granules were mostly roundish with an ellipsoid roundness above 60% (Figure 1b). Moreover, the smaller the granule, the higher the roundness. Indeed, most starch granules below 40  $\mu\text{m}$  diameter displayed an ellipsoid roundness above 80%, while bigger particles showed important shape disparity (Figure 1c).



**Figure 1.** Size and shape analysis of unfractionated starch granules from potato cv. Monalisa. About 3000 starch granules were used for granulomorphometry. (a) Size distribution of potato starch. The number of analyzed granules was plotted versus their corresponding area diameter. (b) Shape distribution of Monalisa potato starch. Same as in (a) with the use of the ellipsoid roundness parameter. (c) Scatter plot of the ellipsoid roundness versus granule diameter.

In order to investigate the disparity of starch structure and composition according to the granule morphology, starch was fractionated sequentially with the use of nylon filters (11, 31, and 50  $\mu\text{m}$  mesh widths). Each fraction was then analyzed by granulomorphometry and scanning electron microscopy (Figure 2 and Table 1). Average granule diameters of  $14.2 \pm 3.7 \mu\text{m}$  (fraction A),  $24.5 \pm 6.5 \mu\text{m}$  (fraction B),  $47.7 \pm 12.8 \mu\text{m}$  (fraction C), and  $61.8 \pm 17.4 \mu\text{m}$  (fraction D) were measured in the different fractions (Figure 2e and Table 1). As expected from the analysis of unfractionated samples, smaller granules were more spherical with an average ellipsoid roundness of about 80% for fraction A and B, while fraction C and D displayed an ellipsoid roundness about 70% (Figure 2a–d, and Table 1).



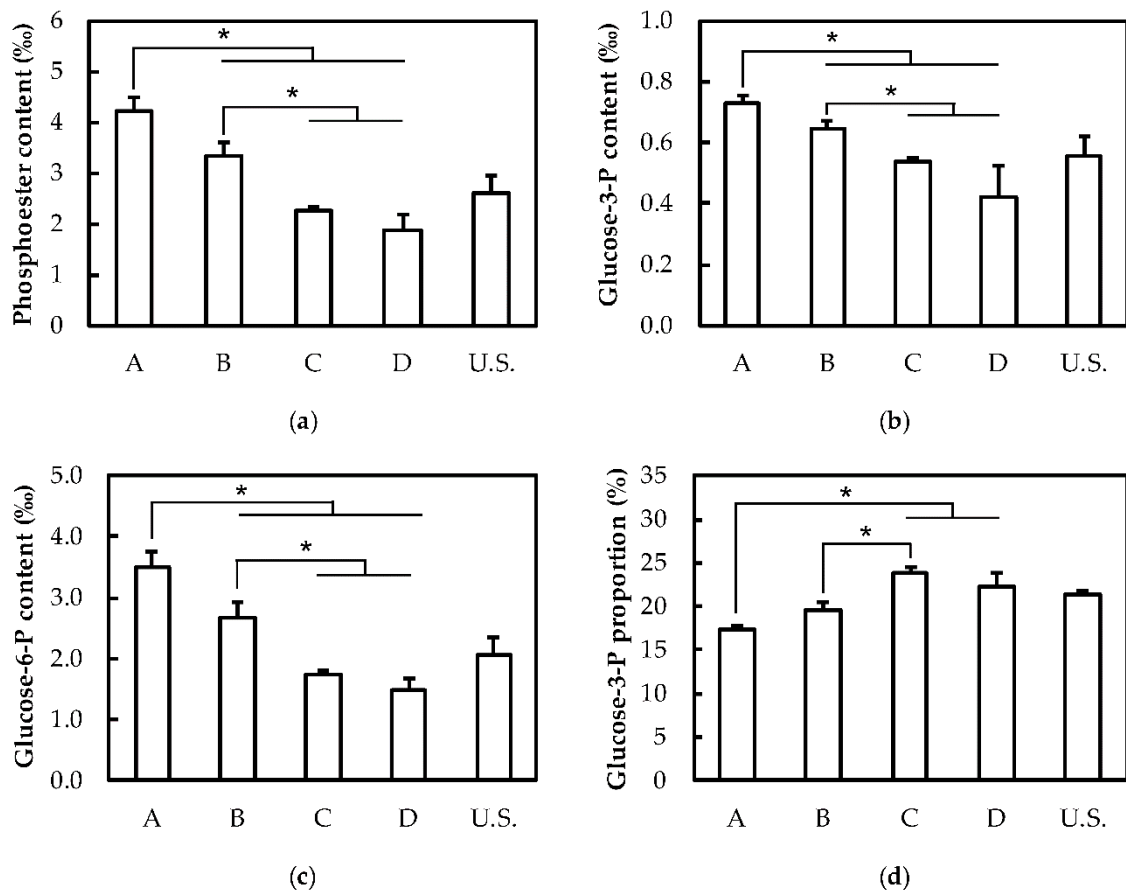
**Figure 2.** Morphology of size-fractionated starch granules. Scanning electron microscopy images of fractions A (a), B (b), C (c), and D (d). (e) Size distribution of each fraction. 3035 granules from fraction A, 3436 granules from fraction B, 1198 granules from fraction C, and 784 granules from fraction D were used for granulomorphometry.

**Table 1.** Size, shape, protein and amylose contents and  $\lambda_{max}$  of the amylopectin-I<sub>2</sub> complex. Statistics were performed by Student t-test compared to unfractionated starch (U.S.). \* corresponds to  $P < 0.05$  and \*\* to  $P < 0.001$ .

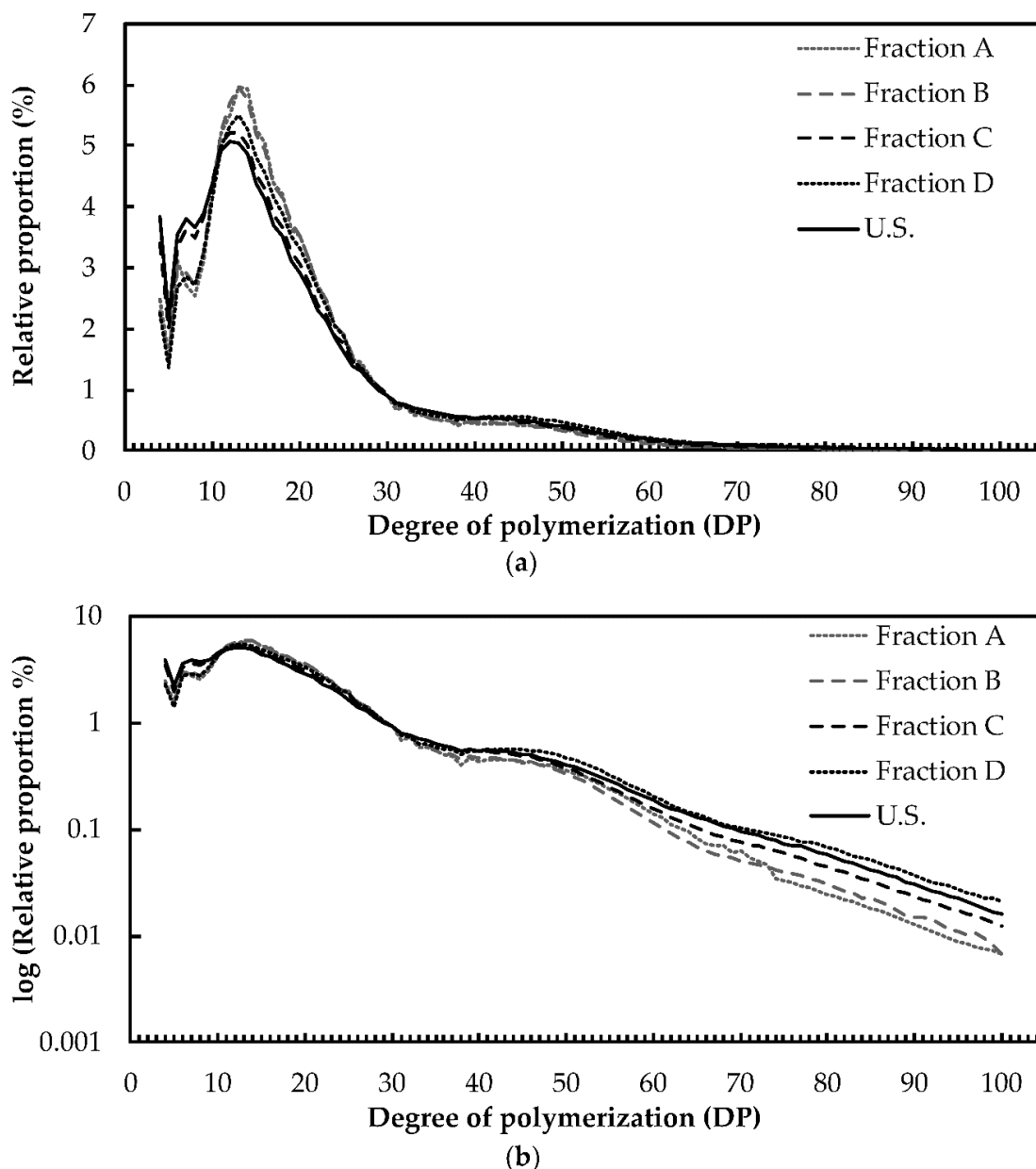
Fraction	Average Diameter ( $\mu\text{m}$ )	Ellipsoid Roundness (%)	Protein Content ( $\mu\text{g}\cdot\text{mg}^{-1}$ )	Amylose Content (%)	$\lambda_{max}$ of the Amylopectin-I <sub>2</sub> Complex (nm)
A	14.2 $\pm$ 3.7 **	85.9 $\pm$ 12.5 **	2.6 $\pm$ 0.2 *	22.6 $\pm$ 5.1	548 $\pm$ 2
B	24.5 $\pm$ 6.5 **	82.5 $\pm$ 13.6 **	2.8 $\pm$ 0.1 **	20.7 $\pm$ 2.1	546 $\pm$ 4
C	47.7 $\pm$ 12.8 **	72.0 $\pm$ 14.8 **	2.1 $\pm$ 0.4 *	20.8 $\pm$ 2.9	552 $\pm$ 2
D	61.8 $\pm$ 17.4 **	73.2 $\pm$ 15.0 **	2.4 $\pm$ 0.1 *	20.6 $\pm$ 1.9	549 $\pm$ 2
U.S.	30.7 $\pm$ 15.9	79.9 $\pm$ 14.6	2.5 $\pm$ 0.2	21.5 $\pm$ 1.0	554 $\pm$ 4

## 2.2. Structure and Composition of Starch

A previous study reported on the structural heterogeneity of size-fractionated potato starch granules [43]. This work highlighted differences in the starch X-ray diffraction (XRD) patterns, amylose contents, and the amylopectin chain length distribution (CLD) among three size fractions [43]. To further investigate these disparities, we determined protein, amylose, and phosphate contents as well as the amylopectin structure in fractions A, B, C and D. On the one hand, no significant differences were observed regarding the amylose content, starch crystallinity and the wavelength at the maximum of absorption ( $\lambda_{max}$ ) of the amylopectin-I<sub>2</sub> complex (Table 1 and Figure S1). On the other hand, protein contents significantly varied in the different granule fractions with a maximum of  $2.8 \pm 0.12 \mu\text{g}\cdot\text{mg}^{-1}$  in fraction B (Table 1). However, these little changes did not follow any specific correlation with the average granule diameter of the corresponding fractions. Conversely, phosphate contents were significantly reduced with the increase of the average granule diameter (Figure 3a). A two-fold decrease of the total phosphate (P) content was observed in fraction D compared to fraction A with  $1.9 \pm 0.3$  and  $4.2 \pm 0.3\%$ , respectively (Figure 3a). Although both C3- and C6-P contents followed the same pattern, the decrease in total P was mainly the consequence of the reduction of C6-P contents (Figure 3b,c). Hence, the proportion of P at the C3 position increased by 1.13- and -1.38 fold in fractions B and C, respectively, compared to fraction A (Figure 3d). In order to investigate the structural heterogeneity of amylopectin from the different fractions, total starch was debranched and chain length distribution (CLD) was determined by fluorescence-assisted capillary electrophoresis (FACE) (Figure 4). A subtle decrease of glucans from DP 10 to 24 was observed with the increase of granule diameter among the fractions, in the order fractions A, B, C, and D (Figure 4a). This was correlated with an increase in the relative abundance of longer chains, from DP 45 to 100 (Figure 4b).



**Figure 3.** Phosphate contents. One milligram of starch from fractions A, B, C, and D and unfractionated starch (U.S.) was hydrolyzed prior to APTS-labelling and capillary electrophoresis analysis. (a) Total phosphate contents. (b) Glucose-3-P contents. (c) Glucose-6-P contents. (d) Proportion of glucose-3-P expressed in % of total phosphate. Error bars represent SDs calculated from three experimental replicates. The asterisks indicate significant differences according to one-way ANOVA with a P-value  $\leq 0.01$ .



**Figure 4.** Amylopectin chain length distribution. One milligram of starch from each fraction and unfractionated starch (U.S.) was derivatized with APTS after enzymatic digestion by pullulanase and isoamylase. Peaks corresponding to DP 4 to DP 100 were integrated and their relative proportion in % is displayed. (a) Chain length distribution in linear scale. (b) Chain length distribution in log<sub>10</sub> scale. Glucan chain proportion was averaged from analyses of two independent fractionations.

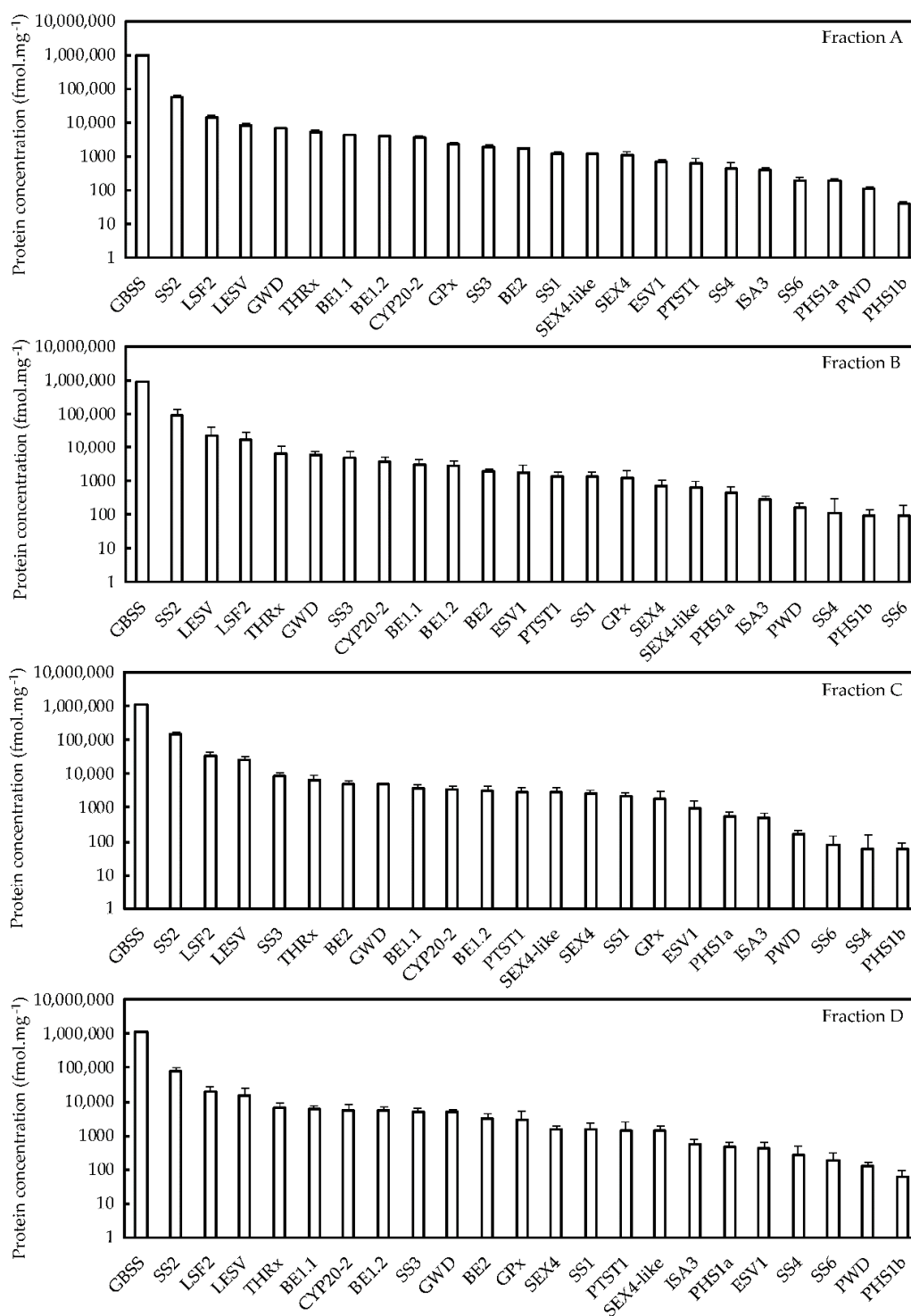
### 2.3. Quantitative Proteomics and Correlation Analysis

Each fraction was analyzed by shotgun mass spectrometry using 300 mg of starch. Starch granules were washed with 2% SDS to remove unspecifically bound proteins prior to gelatinization and isolation of internally trapped proteins. Peptides were obtained by the eFASP (enhanced filter-aided sample preparation) method and analyzed by UPLC-nano-ESI-MS/MS. Protein quantification was performed with the Maxquant and Perseus 1.6.0.2 software and a custom database to avoid quantification biases because of the redundant gene fragments present in the PGSC (potato genome sequencing consortium) databank (Supplementary File 1). This database included proteins that we previously identified in association with potato starch and comprised most of the starch metabolizing enzymes as well

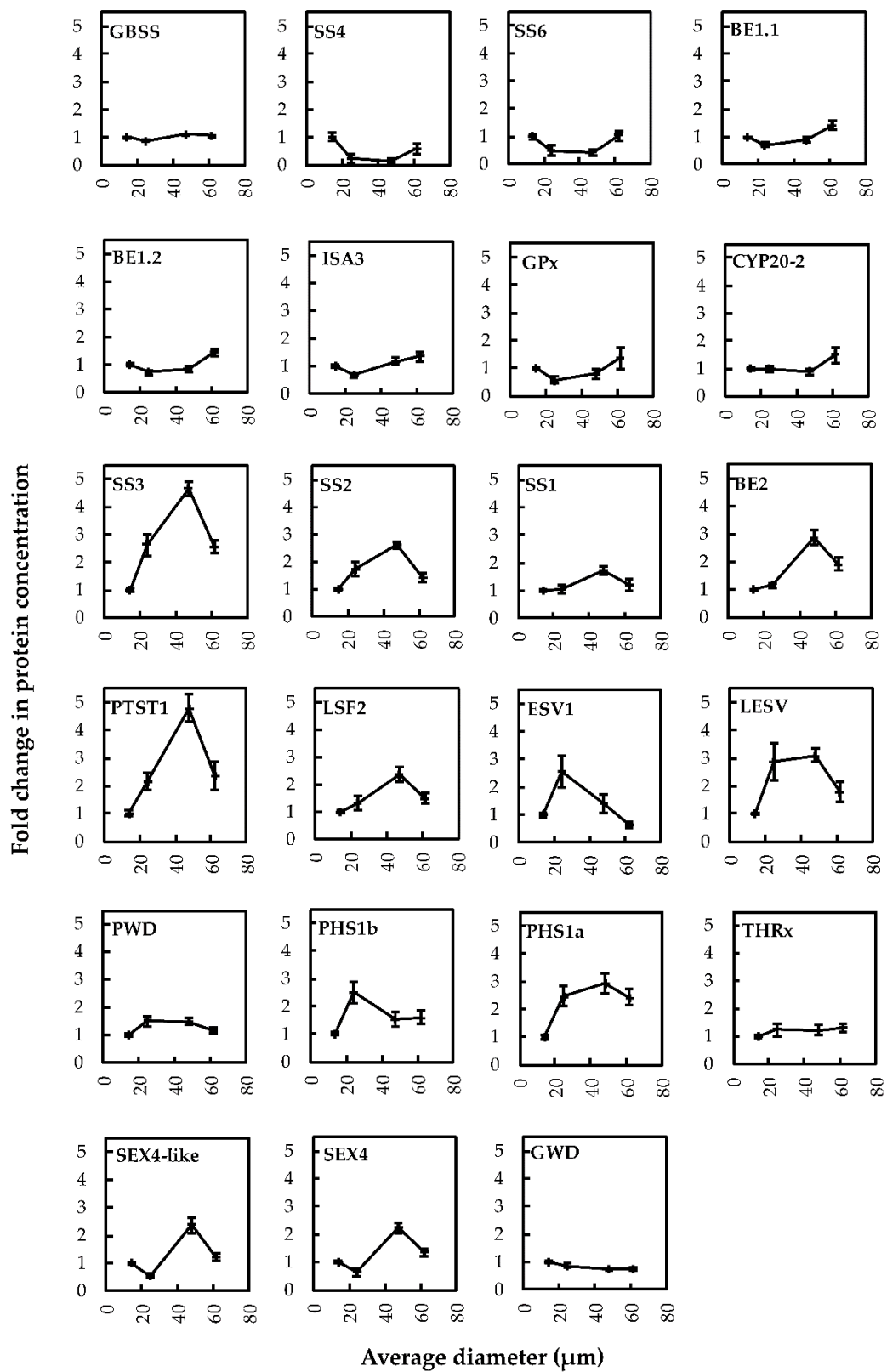
as a thioredoxin (THR<sub>x</sub>), a glutathione peroxidase (GP<sub>x</sub>), and a peptidyl-propyl-*cis-trans*-isomerase (CYP20-20) [16].

In order to evaluate protein stoichiometry within starches of the different fractions, the abundance score (iBAQ, intensity based absolute quantification) of each protein was expressed in fmol per mg of starch (Supplementary File 2, Table 1, Figure 5). Regardless of the fraction, GBSS and SS2 remained the two major proteins, directly followed by LSF2 and LESV that were about 10 times less abundant than SS2 (Figure 5). Interestingly, the relative amount of GWD continuously decreased with the increase of granule average diameter, which was consistent with the observed decrease in glucose-6-P content. The relative abundance of the other proteins also displayed significant variations between the four fractions (Figure 5). This was still the case concerning the less abundant proteins including ISA3, SS4, SS6, PHS1a, and PHS1b. To better illustrate the changes in protein abundance, the concentration of individual proteins in fraction A was set to 1 and fold changes were plotted versus the average granule diameter of the different fractions (Figure 6). SS3 and PTST1 displayed fold changes above 4 while other protein concentrations such as those of GBSS or SS4 varied faintly according to granule diameter with less than 0.2-fold variations (Figure 6). Furthermore, SS1, 2 and 3, BE2, LESV, PWD, PTST1, PHS1a and b, and THR<sub>x</sub> harbored fold changes above 1 between fraction A and the other fractions. Overall, starch-bound proteins could be grouped according to four distinct abundance patterns. SS4, SS6, BE1.1, BE1.2, ISA3, GP<sub>x</sub>, and CYP20-2 showed similar patterns with a decrease in protein amounts in fraction B followed by a continuous increase in fractions C and D (Figure 6). SS1, 2 and 3, BE2, PTST1, LSF2, ESV1, LESV, PWD, PHS1a and b grouped together, displaying an inverse pattern with an increase in fraction B and C followed by a decrease in fraction D (Figure 6). SEX4 and SEX4-like harbored a unique pattern while GBSS, THR<sub>x</sub>, and GWD showed faint variations among the fractions (Figure 6). Paired samples Pearson correlation analysis confirmed these protein behaviors with the strongest positive correlation observed between PTST1 and LSF2 (Table S1). Noteworthy, SS4 and SS6 were strongly anticorrelated with PWD and LESV, illustrating the opposition between their corresponding patterns (Figure 6 and Table S1).

A correlation analysis was also performed between individual protein concentrations and the other parameters measured in this study (Tables S2 and S3). SS1, 2 and 3 showed negative correlations with the relative abundance of amylopectin chains of DP 13 to 28 while their abundance was positively correlated with chains of DP 4 to 10 and 31 to 40 (Table S2). As expected from the positive correlations between the concentrations of LSF2, BE2, PTST1, SEX4, or SEX4-like and the concentrations of SS1, 2 and 3, the former proteins displayed similar correlations with the abundance of these amylopectin chains (Tables S1 and S2). On the other hand, BE1.1, BE1.2, CYP20.2, GP<sub>x</sub>, and ISA3 showed similar positive correlations with the proportion of the longest amylopectin glucans (above DP 42) (Table S2). Most strikingly, the abundance of GWD did not correlate with that of any other starch-bound protein (Table S1). However, it was significantly correlated with the amounts of starch phosphoesters, especially at C6 position, as well as with granule morphology and the amylopectin chain length distribution (Tables S1–S3).



**Figure 5.** Label-free quantitative proteomics of each fraction. Proteins were isolated from 300 mg of starch. They were then digested into peptides and analyzed by MS/MS. Interrogation was realized with Maxquant and Perseus software. Protein concentration in nM was calculated with the Perseus algorithm using the iBAQ score from Maxquant. These concentrations were then expressed in fmol and normalized to starch quantity (mg) using the protein concentrations displayed in Table 1. Error bars represent SDs calculated from three experimental replicates, with three technical replicates for each. Statistical analysis was performed by ANOVA using the LFQ scores from the Maxquant analysis (Supplementary File 3).



**Figure 6.** Fold changes in protein concentrations according to average granule diameter. Fold changes were calculated by normalizing the protein content in fraction A to 1. Error bars represent standard errors (SEs) calculated from three experimental replicates, with three technical replicates for each.



### 3. Discussion

We investigated the heterogeneity of structure, composition, and morphology existing within individual starch samples isolated from whole potato tubers. While the morphological heterogeneity of starch is well-known and often relates to the botanical, tissular, and cellular origins [40], variations in starch composition within a starch extract were scarcely investigated. A recent study highlighted that small potato starch granules harbored a faceted morphology with an A-type crystalline structure contrary to the bigger ovoid B-type starch granules [43]. In our study, such differences of crystalline structure could not be observed by XRD analysis. Moreover, while small granules were significantly rounder than bigger ones, a vast majority displayed a spherical or ovoid smooth morphology as observed by SEM and confirmed by granulomorphometric analysis. These discrepancies might originate from the use of different potato cultivars in both studies or from the use of commercial starch that could contain traces of B-type starch from other species. However, the origin of the starch used in the work from Wang and collaborators was not indicated [43]. On the other hand, we observed similar variations of the amylopectin chain length distribution according to the average granule diameter. The relative proportion of glucans from DP 10 to 24 decreased with the increase of granule diameter and was accompanied by an increase of longer amylopectin chains ranging from DP 45 to 100. Differences in amylopectin CLD among starches from different species and organs have already been described [44]. Such differences within the same starch extract could be explained by several parameters. First, the starch metabolism may be differently regulated in different cell types thus leading to amylopectin structural differences. The drastic disparity between stomatal guard cell and mesophyll starch metabolisms in *Arabidopsis* leaves illustrates this phenomenon [45]. Second, differential expression and regulation of the starch biosynthetic enzymes, especially SSs, Bes, and DBEs, during storage starch accumulation may modulate the amylopectin structure at the time of sampling. If the size of starch granules is related to their maturity state and tissue aging (i.e., small and large granules originating from immature and mature cells, respectively), one could, for example, speculate that temporal expression patterns of starch synthesizing enzymes induce amylopectin structural variations during granule growth. While the dynamics of starch synthesis and degradation were investigated at the granule level in *Arabidopsis* leaves [46], such analysis remains to be performed for storage starch accumulation. Furthermore, a significantly marked decrease in phosphate concentration was observed with the increase of the average granule diameter in the different starch fractions. A previous study mentioned that small potato starch granules are richer in phosphate than larger ones [47]. Here, we show that the amount of starch phosphoesters decreases with increasing granule size at both C3- and C6-position. Moreover, the decrease of total phosphate content mostly comes from the variation of C6-phosphoester concentration.

In order to relate the structural and morphological heterogeneity of starch to the corresponding metabolic pathways, we investigated the starch-bound proteome in the fractionated granule populations. Bottom-up proteomics allowed to identify all the proteins that we reported previously [16]. Label-free quantitation was performed to evaluate the concentration changes of each protein in the different fractions. Even though it is not clear whether all starch-binding enzymes are active when entrapped in the granule, such analysis provides a snapshot of the amount of each protein at the time of harvesting. Our study shows that the protein stoichiometry within starch granules significantly changes among the four fractions. Taken individually, the protein concentrations changed from less than 0.2-fold to more than 4-fold according to the average granule diameter. The most stable concentrations were those of GBSS and THR $\alpha$ . Conversely, PTST1 and SS3 harbored the most drastic variations in concentration. Proteins could be classified into four groups depending on their fold-change profiles which was confirmed by paired-correlation analysis. While a group of proteins comprising SS4, SS6, BE1.1, BE1.2, ISA3, GPx, and CYP20-2 showed similar patterns with a decrease followed by a continuous increase with the increase of granule diameter, SS1, 2 and 3, BE2, PTST1, LSF2, ESV1, LESV, PWD, PHS1a and b grouped together and harbored an inverse profile. These observations may illustrate the existence of protein complexes inside the starch granules such as those described in wheat, maize,

and rice [48–50]. As expected, a strong positive correlation was observed between phosphate contents and the concentration of GWD. Although the decrease in GWD concentration was low relatively to changes of other proteins, it continuously decreased with the increase of granule size similar to the concentration of glucose-6-P. It is intuitive that the decrease in phosphate concentration results from the decreased concentration of GWD. However, correlation analysis do not allow to conclude on the causal links between the correlated parameters. This is especially important concerning the multiple potential relationships between the starch-bound proteins themselves or between the latter and starch structure or morphology. For example, the strong correlation between phosphate content and granule morphology could either result from an impact of starch phosphoesters on granule size or from a temporal variation of GWD expression during granule growth. Further investigation including single cell transcriptomics and isotopic protein labelling in pulse-chase proteomic experiments will allow to address these questions. Overall, this study corroborates that the starch-bound proteome may influence the starch granule growth and architecture thus opening new perspectives in understanding the dynamics of starch biosynthesis.

#### 4. Materials and Methods

All solutions were prepared using ultrapure water purified with a MilliQ™ Academic system (Merck Millipore, Burlington, MA, USA). All chemicals, biochemicals and solvents were purchased from Sigma-Aldrich (Saint-Louis, MO, USA) and used without purification.

##### 4.1. Starch Extraction

Starch granules were isolated from tubers of *Solanum tuberosum* (Monalisa) cultivated in the field in Villeneuve d'Ascq (50.607735, 3.143431), France, between March and July 2014. Potato tubers (approximately 70 g) were washed with tap water and peeled prior extraction. The tubers were ground with a blender in 200 mL of ultrapure water. The tuber extracts were then filtered through a nylon net (100 µm mesh) and left for starch granule sedimentation during 3 h. The supernatant was then removed and sedimented starch was resuspended in 500 mL of ultrapure water. The starch suspensions were subsequently washed three times with 1 L of ultrapure water and stored in 20% ethanol at 4 °C.

##### 4.2. Granulomorphometry

The starch granule size and shape were measured with an Ipac 1 granulomorphometer (Occhio, Angleur, Belgium). The retained parameters for granule size were the ISO area diameter and granule shape with ellipsoid roundness. Before measurement, a numerical filter was applied. It excluded particles with a diameter smaller than 4 µm or higher than 100 µm. In addition, particles with an ellipsoid roundness lower than 46% and a diameter between 4 and 18 µm were also excluded. An average number of 3000 granules was analyzed.

##### 4.3. Filter-Fractionation of Starch Granules

Granules were separated using nylon filters with a mesh width of 100, 50, 31, and 11 µm. The first filtration was realized with the 11 µm filter in order to isolate fraction A. The retained fraction was refiltered with a 31 µm filter to obtain fraction B. Then, the new retained fraction was filtered with a 50 µm filter. The eluted fraction was fraction C, and the retained fraction was fraction D. The fractionation quality was estimated by size measurement with an Ipac 1 granulomorphometer (Occhio, Angleur, Belgium). The parameters used were the same as described above.

##### 4.4. Scanning Electron Microscopy (SEM)

Drops of dilute aqueous suspensions of starch granules were deposited on metallic stubs covered with C/Ni tape and allowed to dry. The specimens were coated with Au/Pd in a Baltec MED 020 sputter

coater (Leica, Wetzlar, Allemagne) and observed in secondary electron mode in a FEI Quanta FEG 250 microscope (Hillsboro, OR, USA) operating at an accelerating voltage of 2.5 kV.

#### 4.5. X-Ray Diffraction (XRD)

The starch granule fractions were placed in a chamber maintaining a relative humidity (R.H.) of 93% for 5 days. The powders were then poured into 0.7-mm (outer diameter) glass capillaries that were flame-sealed. All specimens were X-rayed by transmission in a Warhus vacuum chamber with a Ni-filtered CuK $\alpha$  radiation ( $\lambda = 0.1542$  nm) using a Philips PW3830 generator (Eindhoven, The Netherlands) operating at 30 kV and 20 mA. Two-dimensional XRD patterns were recorded on Fujifilm imaging plates exposed for 2 h and read off-line with a Fujifilm BAS 1800-II bioimaging analyzer (Minato, Tokyo, Japan). Diffraction profiles were calculated as rotational averages of the 2D patterns.

#### 4.6. Amylose Content

The protocol for amylose assay was adapted for Megazyme (Wicklow, Ireland) K-AMYL kit [51,52]. Briefly, 1.5 mg of starch from each fraction was gelatinized in 100% dimethylsulfoxide (DMSO), and precipitated with 95% ethanol in order to remove lipids. Starch was then resolubilized in 100% DMSO. Amylopectin was precipitated with concavalin A (ConA) following the manual instructions. After centrifugation, the supernatant containing amylose was recovered, and the pellet containing amylopectin was resolubilized in 100% DMSO. ConA was then inactivated by heating at 100 °C. Amylose, amylopectin, and total starch were then digested using the glucose determination reagent (GOPOD) available in the Megazyme K-AMYL kit. The absorbance was read at 510 nm using a SPECTROstar Nano spectrometer (BMG LABTECH's, Allmendgrün, Germany). At the same time, an iodine staining was performed on amylose, and amylopectin fractions, and total starch. The absorbance was read between 500 and 700 nm in order to determine the wavelength at the maximum of absorbance ( $\lambda_{max}$ ) for each glucan-I<sub>2</sub> complex.

#### 4.7. Amylopectin Chain Length Distribution

The amylopectin chain length distribution was measured by fluorescence-assisted capillary electrophoresis (FACE). Two milligram of starch from each fraction, in suspension in 250  $\mu$ L of ultrapure water, was gelatinized by heating at 100 °C, with intermittent high-speed stirring in a vortex mixer, until the solution became clear. Then, 20 U of isoamylase and 1 U of pullulanase were added and the solution was incubated at 42 °C overnight in order to debranch amylopectin. Proteins were then precipitated by heating at 95 °C for 15 min, and the supernatant was recovered after centrifugation. The samples were desalted on Alltech™ Carbograph (Grace, Columbia, MD, USA) solid-phase extraction (SPE) column, then freeze-dried. The dry samples were resolubilized in 2  $\mu$ L of 1 M sodium cyanoborohydride in tetrahydrofuran (THF), and 2  $\mu$ L 0.2 M 8-aminopyrene-1,3,6-trisulfonic acid trisodium salt (APTS) in 15% acetic acid (*v/v*). The samples were incubated overnight at 42 °C, and diluted with 46  $\mu$ L ultrapure water. After the first dilution, the sample was then rediluted 50 times with ultrapure water prior to injection in a Beckman Coulter PA800-plus Pharmaceutical Analysis System (Beckman Coulter, Brea, CA, USA) equipped with a silicon capillary column (inner diameter: 50  $\mu$ m; outer diameter: 360  $\mu$ m; length: 60 cm). The capillary was rinsed and coated with Beckman Coulter N-linked carbohydrate separation gel buffer one-third diluted with ultrapure water prior to injection (15 s at 10 kV). The migration was performed at 10 kV during 1 h at 20 °C. The detection was made with a 488 nm solid-state laser module, and a laser-induced fluorescence (LIF) detector. The detection window was located 10 cm above the outlet of the capillary column.

#### 4.8. Phosphate Content

The phosphate content for each fraction was determined by FACE. The protocol was the same as that described in Verbeke et al. 2016. Briefly, 1 mg of starch from each fraction was submitted to acid

hydrolysis with 100  $\mu\text{L}$  of 2 M trifluoroacetic acid (TFA) at 95 °C during 1 h and 20 min. A total of 20  $\mu\text{L}$  of the hydrolysate was then diluted with 200  $\mu\text{L}$  of ultrapure water to stop the reaction. The sample was then dried with a SpeedVac™ Concentrator (Eppendorf™ Concentrator Plus, Eppendorf, Hamburg, Germany). The dry samples were then derivatized with APTS as described before. The derivatized samples were diluted with 96  $\mu\text{L}$  of ultrapure water. A total of 15  $\mu\text{L}$  of this first dilution was rediluted with 185  $\mu\text{L}$  of ultrapure water prior to injection in a Beckman Coulter PA800-plus Pharmaceutical Analysis System with the same silicon capillary column as that described above. The capillary was rinsed and coated with the Beckman Coulter *N*-linked carbohydrate separation gel buffer prior to injection (10 s at 10 kV). The migration was performed at 10 kV during 10 min. The detection was made at 488 nm as described above.

#### 4.9. Starch Granule Associated Protein Extraction

The starch proteins were extracted from 300 mg of dry fractionated starch. Ten aliquots of 30 mg starch were then suspended in 500  $\mu\text{L}$  of extraction buffer (0.2 M Tris at pH 6.8; 2% sodium dodecyl sulfate (SDS); 50 mM dithiothreitol (DTT)). Aliquots are then heated to 100 °C until gelatinization of the granules. After gelatinization, the samples were kept for an additional 10 min at 100 °C and centrifuged for 10 min at 13,000 rounds per minute (rpm). Supernatants containing proteins were rescued, pooled, and dropped on an Amicon® (Millipore, Billerica, MA, USA) membrane filter with a 50 kDa cutoff inside a 15 mL vial. The vial was centrifuged at 7500 rpm for 30 min in order to concentrate the samples. A total of 5  $\mu\text{L}$  of each sample, diluted 10 times, was used to carry out a Bicinchoninic assay (BCA; Thermo Fisher Scientific, Waltham, MA, USA) following the manual instructions.

#### 4.10. Label-Free Quantitative Proteomics, Sample Preparation

A total of 50  $\mu\text{g}$  of proteins are reduced, alkylated, and digested with eFASP Digestion protocol. Ultra-filtration filters from Amicon® units (10 kDa cutoff limit; Millipore, Billerica, MA, USA) were incubated overnight with 5% *v/v* TWEEN-20 (T20, Sigma-Aldrich, Saint-Louis, MO, USA). Following incubation, the filter units were rinsed thoroughly by three immersions in ultrapure water. eFASP digestion is realized with 50  $\mu\text{g}$  of samples estimated from the BCA assay. These samples were mixed in 50  $\mu\text{L}$  of solubilization and reducing buffer A (4% SDS, 0.2% deoxycholic acid (DCA), 50 mM DTT, 200 mM ammonium bicarbonate (ABC)) overnight at 4 °C. Samples were centrifuged at 13,000 *g* at 20 °C for 15 min. Supernatants were mixed with 200  $\mu\text{L}$  of SDS buffer exchange (8 M urea, 0.2% DCA, 100 mM ABC, pH 8) and were transferred into Amicon® units. After 30 min of 13,000 *g* centrifugation, the filtrate was discarded, an additional 200  $\mu\text{L}$  of exchange buffer A was deposited in each unit, and centrifugation resumed for 30 more minutes. This buffer addition/centrifugation step was repeated twice more. The reduced proteins were alkylated within the filter unit by adding alkylation buffer A (8 M urea, 50 mM iodoacetamide, and 100 mM ABC, pH 8) and incubating at 37 °C for 1 h with shaking in the dark. After 30 min of 13,000 *g* centrifugation, the filtrate was discarded, an additional 200  $\mu\text{L}$  of exchange buffer A was deposited in each unit, and centrifugation resumed for 30 more minutes. This buffer addition/centrifugation step was repeated once more. After buffer exchange was performed with three exchanges with eFASP digestion buffer (50 mM ABC, 0.2% DCA pH 8), 100  $\mu\text{L}$  of eFASP digestion buffer was added to each Amicon® units, followed by 1  $\mu\text{g}$  of trypsin (1:50 *w/w*). Digestion proceeded for 16 h at 37 °C. Peptides were recovered by transferring the UF filter to a new collection tube and spinning at 13 000 *g* for 20 min. To complete peptide recovery, filters were rinsed twice with 50  $\mu\text{L}$  of 50 mM ABC that was collected by centrifugation. A total of 200  $\mu\text{L}$  of ethyl acetate was added to the peptide containing filtrate and was transferred to a 2 mL tube. After that 2.5  $\mu\text{L}$  TFA was added and was quickly vortex. White thread like precipitate may be visible for large quantities of peptides. Peptides precipitate were mixed with 800  $\mu\text{L}$  of ethyl acetate and were centrifuged at 13,000 *g* for 10 min. The organic supernatant was discarded and this step was repeated twice more. The upper organic layer was carefully removed and discarded as much as possible without disturbing the boundary layer. Uncovered sample tubes were placed in a thermomixer at 60 °C, in a fume hood,

for 5 min to remove residual ethyl acetate. Residual organic and volatile salts were removed by vacuum drying in a SpeedVac™ Concentrator (Eppendorf™ Concentrator Plus, Eppendorf, Hamburg, Germany). This step was repeated two times with 50% methanol. Ten micrograms of dry peptides are then resuspended in 0.1% (*v/v*) formic acid.

#### 4.11. Mass Spectrometry Analysis

A nanoflow HPLC instrument (U3000 RSLC, Thermo Fisher Scientific, Waltham, MA, USA) was coupled on-line to a Q Exactive plus Orbitrap mass spectrometer (Thermo Fisher Scientific) fitted with a nanoelectrospray ion source. One microliter of peptide mixture (corresponding to 500 ng of proteins) was loaded onto the preconcentration trap (Thermo Fisher Scientific, Acclaim PepMap100 C18, 5  $\mu\text{m}$ , 300  $\mu\text{m}$  i.d.  $\times$  5 mm) using a partial loop injection, for 5 min at a flow rate of 10  $\mu\text{L}\cdot\text{min}^{-1}$  with buffer A (5% acetonitrile and 0.1% formic acid) then separated on column (Acclaim PepMap100 C18, 3  $\mu\text{m}$ , 75 mm i.d.  $\times$  500 mm) and separated with a linear gradient of 5–50% buffer B (75% acetonitrile and 0.1% formic acid) at a flow rate of 250  $\text{nL}\cdot\text{min}^{-1}$  and temperature of 45 °C. The total time for an LC MS/MS run was 270 min. MS data was acquired on Q Exactive plus Orbitrap mass spectrometer (Thermo Fisher Scientific) using a data-dependent top 20 method, dynamically choosing the most abundant precursor ions from the survey scan (350–1600  $m/z$ ) for HCD fragmentation. Dynamic exclusion duration was 60 s. Isolation of precursors was performed with a 1.6  $m/z$  window and MS/MS scans were acquired with a starting mass of 80  $m/z$ . Survey scans were acquired at a resolution of 70,000 at  $m/z$  400 (AGC set to  $10^6$  ions with a maximum fill time of 100 ms). Resolution for HCD spectra was set to 70,000 at  $m/z$  200 (AGC set to  $10^5$  ions with a maximum fill time of 200 ms). Normalized collision energy was 28 eV. The underfill ratio, which specifies the minimum percentage of the target value likely to be reached at maximum fill time, was defined as 0.4%. The instruments were run with peptide recognition mode (i.e., from 2 to 8 charge), exclusion of singly charged and of unassigned precursor ions enabled.

#### 4.12. Proteins Label Free Quantification with Maxquant and Perseus

Proteins were quantified with Maxquant software 1.5.3.30 with custom starch bound protein databank (Supplementary file 1) regrouping proteins identified as bound to starch in our previous study [16], in order to obtain iBAQ and LFQ (label-free quantitation) score for each identified proteins. iBAQ score is used on Perseus 1.6.0.2 software to obtain the abundance of each protein. Retained parameter for the quantification is “Concentration [nM] iBAQ” (Supplementary file 2). The concentration score is then divided by the total concentration of protein in  $\mu\text{g}\cdot\text{mg}$  of starch<sup>-1</sup> in order to obtain a result in  $\text{fmol}\cdot\text{mg}$  of starch<sup>-1</sup>. Fold change values are obtained by normalizing the concentration of one protein on the concentration of this protein in the fraction A. LFQ values are transformed in log2 then filtered and used on Perseus 1.6.0.2 to realize statistical testing. For multiple comparisons, ANOVA analysis with a permuted based FDR correction (FDR = 0.05) was performed. For the comparison between two groups, *t*-tests were performed with P-values less than 0.05 were considered to be statistically significant (Supplementary file 3). The correlations between variables were defined by the Pearson coefficient and performed using R version 3.6.0.

**Supplementary Materials:** The following are available online at <http://www.mdpi.com/2223-7747/8/9/324/s1>, Figure S1: X-ray diffraction (XRD) profiles of size-fractionated Monalisa starch granules, Table S1: Pearson paired-samples correlation analysis of starch-bound protein concentrations, Table S2: Pearson paired-samples correlation analysis between starch-bound protein concentrations and the relative proportion of each DP from chain length distribution (CLD) analysis, Table S3: Pearson paired-samples correlation analysis between starch-bound protein concentrations and phosphate contents or granule morphology, Supplementary file 1: Custom protein database, Supplementary file 2: Intensity based absolute quantification of proteins, Supplementary file 3: Statistical analysis of protein quantification.

**Author Contributions:** Conceptualization: N.S., C.R. and C.D.; sample collection: N.S and J.V.; sample fractionation: S.H. and J.V.; physical, biochemical, and morphological analyses of starch: S.H., N.S., J.V., and J.-L.P.; mass spectrometry analysis: S.H., F.B., and S.F.; data analysis: S.H., F.B., and N.S.; writing-original draft preparation: S.H. and N.S.; writing—review and editing: J.-L.P., C.R., C.D., F.B., J.V., S.F., and N.S.

**Funding:** This research was funded by an ANR grant through the JCJC-2017 program (ANR-17-CE20-0012) as well as the Hauts de France region through the CPER “Alibiotech”.

**Acknowledgments:** We thank the NanoBio-ICMG Platform (FR 2607, Grenoble) for granting access to the electron microscopy facility and Christine Lancelon-Pin (CERMAV) for the SEM observations. The proteomic platform was supported by the IBISA network. The mass spectrometers were funded by the University of Lille, the CNRS, the Région Hauts-de-France, and the European Regional Development Fund.

**Conflicts of Interest:** The authors declare no conflict of interest.

## References

1. Manners, D.J. The structure and metabolism of starch. *Essays Biochem.* **1974**, *10*, 37–71. [[PubMed](#)]
2. Curá, J.A.; Jansson, P.-E.; Krisman, C.R. Amylose is not strictly linear. *Starch Stärke* **1995**, *47*, 207–209. [[CrossRef](#)]
3. Bertoft, E. Understanding starch structure: Recent Progress. *Agronomy* **2017**, *7*, 56. [[CrossRef](#)]
4. Swinkels, J.J.M. Composition and properties of commercial native starches. *Starch Stärke* **1985**, *37*, 1–5. [[CrossRef](#)]
5. Baldwin, P.M. Starch granule-associated proteins and polypeptides: A Review. *Starch Stärke* **2001**, *53*, 475–503. [[CrossRef](#)]
6. Denyer, K.; Clarke, B.; Hylton, C.; Tatge, H.; Smith, A.M. The elongation of amylose and amylopectin chains in isolated starch granules. *Plant J.* **1996**, *10*, 1135–1143. [[CrossRef](#)]
7. Commuri, P.D.; Keeling, P.L. Chain-length specificities of maize starch synthase I enzyme: Studies of glucan affinity and catalytic properties. *Plant J.* **2001**, *25*, 475–486. [[CrossRef](#)]
8. Craig, J.; Lloyd, J.R.; Tomlinson, K.; Barber, L.; Edwards, A.; Wang, T.L.; Martin, C.; Hedley, C.L.; Smith, A.M. Mutations in the gene encoding starch synthase II profoundly alter amylopectin structure in pea embryos. *Plant Cell* **1998**, *10*, 413–426. [[CrossRef](#)]
9. Lin, Q.; Huang, B.; Zhang, M.; Zhang, X.; Rivenbark, J.; Lappe, R.L.; James, M.G.; Myers, A.M.; Hennen-Bierwagen, T.A. Functional interactions between starch synthase III and isoamylase-type starch-debranching enzyme in maize endosperm. *Plant Physiol.* **2012**, *158*, 679–692. [[CrossRef](#)]
10. Ragel, P.; Streb, S.; Feil, R.; Sahrawy, M.; Annunziata, M.G.; Lunn, J.E.; Zeeman, S.; Mérida, Á. Loss of starch granule initiation has a deleterious effect on the growth of arabidopsis plants due to an accumulation of ADP-glucose. *Plant Physiol.* **2013**, *163*, 75–85. [[CrossRef](#)]
11. Raynaud, S.; Ragel, P.; Rojas, T.; Mérida, Á. The N-terminal part of *Arabidopsis thaliana* starch synthase 4 determines the localization and activity of the enzyme. *J. Biol. Chem.* **2016**, *291*, 10759–10771. [[CrossRef](#)] [[PubMed](#)]
12. Roldán, I.; Wattedled, F.; Mercedes Lucas, M.; Delvallé, D.; Planchot, V.; Jiménez, S.; Pérez, R.; Ball, S.; D’Hulst, C.; Mérida, A. The phenotype of soluble starch synthase IV defective mutants of *Arabidopsis thaliana* suggests a novel function of elongation enzymes in the control of starch granule formation. *Plant J. Cell Mol. Biol.* **2007**, *49*, 492–504. [[CrossRef](#)] [[PubMed](#)]
13. Seung, D.; Lu, K.-J.; Stettler, M.; Streb, S.; Zeeman, S.C. Degradation of glucan primers in the absence of starch synthase 4 disrupts starch granule initiation in Arabidopsis. *J. Biol. Chem.* **2016**, *291*, 20718–20728. [[CrossRef](#)] [[PubMed](#)]
14. Szydlowski, N.; Ragel, P.; Raynaud, S.; Lucas, M.M.; Roldán, I.; Montero, M.; Muñoz, F.J.; Ovecka, M.; Bahaji, A.; Planchot, V.; et al. Starch granule initiation in Arabidopsis requires the presence of either class IV or class III starch synthases. *Plant Cell* **2009**, *21*, 2443–2457. [[CrossRef](#)] [[PubMed](#)]
15. Van Harsselaar, J.K.; Lorenz, J.; Senning, M.; Sonnewald, U.; Sonnewald, S. Genome-wide analysis of starch metabolism genes in potato (*Solanum tuberosum* L.). *BMC Genomics* **2017**, *18*, 37. [[CrossRef](#)] [[PubMed](#)]
16. Helle, S.; Bray, F.; Verbeke, J.; Devassine, S.; Courseaux, A.; Facon, M.; Tokarski, C.; Rolando, C.; Szydlowski, N. Proteome analysis of potato starch reveals the presence of new starch metabolic proteins as well as multiple protease inhibitors. *Front. Plant Sci.* **2018**, *9*, 746. [[CrossRef](#)] [[PubMed](#)]
17. Li, C.; Wu, A.C.; Go, R.M.; Malouf, J.; Turner, M.S.; Malde, A.K.; Mark, A.E.; Gilbert, R.G. The characterization of modified starch branching enzymes: Toward the control of starch chain-length distributions. *PLoS ONE* **2015**, *10*, e0125507. [[CrossRef](#)] [[PubMed](#)]

18. Ferreira, S.J.; Senning, M.; Fischer-Stettler, M.; Streb, S.; Ast, M.; Neuhaus, H.E.; Zeeman, S.C.; Sonnewald, S.; Sonnewald, U. Simultaneous silencing of isoamylases ISA1, ISA2 and ISA3 by multi-target RNAi in potato tubers leads to decreased starch content and an early sprouting phenotype. *PLoS ONE* **2017**, *12*, e0181444. [[CrossRef](#)] [[PubMed](#)]
19. Streb, S.; Zeeman, S.C. Replacement of the endogenous starch debranching enzymes ISA1 and ISA2 of Arabidopsis with the rice orthologs reveals a degree of functional conservation during starch synthesis. *PLoS ONE* **2014**, *9*, e92174. [[CrossRef](#)]
20. Borén, M.; Larsson, H.; Falk, A.; Jansson, C. The barley starch granule proteome - Internalized granule polypeptides of the mature endosperm. *Plant Sci.* **2004**, *166*, 617–626. [[CrossRef](#)]
21. Denyer, K.; Sidebottom, C.; Hylton, C.M.; Smith, A.M. Soluble isoforms of starch synthase and starch-branching enzyme also occur within starch granules in developing pea embryos. *Plant J.* **1993**, *4*, 191–198. [[CrossRef](#)] [[PubMed](#)]
22. Grimaud, F.; Rogniaux, H.; James, M.G.; Myers, A.M.; Planchot, V. Proteome and phosphoproteome analysis of starch granule-associated proteins from normal maize and mutants affected in starch biosynthesis. *J. Exp. Bot.* **2008**, *59*, 3395–3406. [[CrossRef](#)] [[PubMed](#)]
23. Peng, M.; Gao, M.; Båga, M.; Hucl, P.; Chibbar, R.N. Starch-branching enzymes preferentially associated with A-type starch granules in wheat endosperm. *Plant Physiol.* **2000**, *124*, 265–272. [[CrossRef](#)] [[PubMed](#)]
24. Rahman, S.; Kosar-Hashemi, B.; Samuel, M.S.; Hill, A.; Abbott, D.C.; Skerritt, J.H.; Preiss, J.; Appels, R.; Morell, M.K. The major proteins of wheat endosperm starch granules. *Funct. Plant Biol.* **1995**, *22*, 793–803. [[CrossRef](#)]
25. Regina, A.; Kosar-Hashemi, B.; Li, Z.; Pedler, A.; Mukai, Y.; Yamamoto, M.; Gale, K.; Sharp, P.J.; Morell, M.K.; Rahman, S. Starch branching enzyme IIb in wheat is expressed at low levels in the endosperm compared to other cereals and encoded at a non-syntenic locus. *Planta* **2005**, *222*, 899–909. [[CrossRef](#)] [[PubMed](#)]
26. Stensballe, A.; Hald, S.; Bauw, G.; Blennow, A.; Welinder, K.G. The amyloplast proteome of potato tuber. *FEBS J.* **2008**, *275*, 1723–1741. [[CrossRef](#)] [[PubMed](#)]
27. Xing, S.; Meng, X.; Zhou, L.; Mujahid, H.; Zhao, C.; Zhang, Y.; Wang, C.; Peng, Z. Proteome profile of starch granules purified from rice (*Oryza sativa*) endosperm. *PLoS ONE* **2016**, *11*, e0168467. [[CrossRef](#)] [[PubMed](#)]
28. Ritte, G.; Lloyd, J.R.; Eckermann, N.; Rottmann, A.; Kossmann, J.; Steup, M. The starch-related R1 protein is an alpha -glucan, water dikinase. *Proc. Natl. Acad. Sci. USA* **2002**, *99*, 7166–7171. [[CrossRef](#)] [[PubMed](#)]
29. Kötting, O.; Pusch, K.; Tiessen, A.; Geigenberger, P.; Steup, M.; Ritte, G. Identification of a novel enzyme required for starch metabolism in Arabidopsis leaves. The phosphoglucan, water dikinase. *Plant Physiol.* **2005**, *137*, 242–252. [[CrossRef](#)]
30. Niittylä, T.; Comparot-Moss, S.; Lue, W.-L.; Messerli, G.; Trevisan, M.; Seymour, M.D.J.; Gatehouse, J.A.; Villadsen, D.; Smith, S.M.; Chen, J.; et al. Similar protein phosphatases control starch metabolism in plants and glycogen metabolism in mammals. *J. Biol. Chem.* **2006**, *281*, 11815–11818. [[CrossRef](#)] [[PubMed](#)]
31. Santelia, D.; Kötting, O.; Seung, D.; Schubert, M.; Thalmann, M.; Bischof, S.; Meekins, D.A.; Lutz, A.; Patron, N.; Gentry, M.S.; et al. The phosphoglucan phosphatase like sex Four2 dephosphorylates starch at the C3-position in Arabidopsis. *Plant Cell* **2011**, *23*, 4096–4111. [[CrossRef](#)] [[PubMed](#)]
32. Sokolov, L.N.; Dominguez-Solis, J.R.; Allary, A.-L.; Buchanan, B.B.; Luan, S. A redox-regulated chloroplast protein phosphatase binds to starch diurnally and functions in its accumulation. *Proc. Natl. Acad. Sci. USA* **2006**, *103*, 9732–9737. [[CrossRef](#)] [[PubMed](#)]
33. Seung, D.; Soyk, S.; Coiro, M.; Maier, B.A.; Eicke, S.; Zeeman, S.C. PROTEIN TARGETING TO STARCH is required for localising GRANULE-BOUND STARCH SYNTHASE to starch granules and for normal amylose synthesis in Arabidopsis. *PLoS Biol.* **2015**, *13*, e1002080. [[CrossRef](#)] [[PubMed](#)]
34. Feike, D.; Seung, D.; Graf, A.; Bischof, S.; Ellick, T.; Coiro, M.; Soyk, S.; Eicke, S.; Mettler-Altmann, T.; Lu, K.J.; et al. The starch granule-associated protein EARLY STARVATION1 is required for the control of starch degradation in *Arabidopsis thaliana* leaves. *Plant Cell* **2016**, *28*, 1472–1489. [[CrossRef](#)] [[PubMed](#)]
35. Seung, D.; Boudet, J.; Monroe, J.; Schreier, T.B.; David, L.C.; Abt, M.; Lu, K.-J.; Zanella, M.; Zeeman, S.C. Homologs of PROTEIN TARGETING TO STARCH control starch granule initiation in Arabidopsis leaves. *Plant Cell* **2017**, *29*, 1657–1677. [[CrossRef](#)] [[PubMed](#)]
36. Valerio, C.; Costa, A.; Marri, L.; Issakidis-Bourguet, E.; Pupillo, P.; Trost, P.; Sparla, F. Thioredoxin-regulated  $\beta$ -amylase (BAM1) triggers diurnal starch degradation in guard cells, and in mesophyll cells under osmotic stress. *J. Exp. Bot.* **2011**, *62*, 545–555. [[CrossRef](#)] [[PubMed](#)]

37. Skryhan, K.; Cuesta-Sejjo, J.A.; Nielsen, M.M.; Marri, L.; Mellor, S.B.; Glaring, M.A.; Jensen, P.E.; Palcic, M.M.; Blennow, A. The role of cysteine residues in redox regulation and protein stability of *Arabidopsis thaliana* starch synthase 1. *PLoS ONE* **2015**, *10*, e0136997. [[CrossRef](#)] [[PubMed](#)]
38. Noctor, G.; Mhamdi, A.; Chaouch, S.; Han, Y.; Neukermans, J.; Marquez-Garcia, B.; Queval, G.; Foyer, C.H. Glutathione in plants: An integrated overview. *Plant Cell Environ.* **2012**, *35*, 454–484. [[CrossRef](#)]
39. Pfister, B.; Zeeman, S.C. Formation of starch in plant cells. *Cell. Mol. Life Sci. CMLS* **2016**, *73*, 2781–2807. [[CrossRef](#)]
40. Jane, J.-L.; Kasemsuwan, T.; Leas, S.; Zobel, H.; Robyt, J.F. Anthology of starch granule morphology by scanning electron microscopy. *Starch Stärke* **1994**, *46*, 121–129. [[CrossRef](#)]
41. Tanaka, E.; Ral, J.-P.F.; Li, S.; Gaire, R.; Cavanagh, C.R.; Cullis, B.R.; Whan, A. Increased accuracy of starch granule type quantification using mixture distributions. *Plant Methods* **2017**, *13*, 107. [[CrossRef](#)] [[PubMed](#)]
42. Tester, R.F.; Karkalas, J.; Qi, X. Starch structure and digestibility Enzyme-Substrate relationship. *Worlds Poult. Sci. J.* **2004**, *60*, 186–195. [[CrossRef](#)]
43. Wang, C.; Tang, C.-H.; Fu, X.; Huang, Q.; Zhang, B. Granular size of potato starch affects structural properties, octenylsuccinic anhydride modification and flowability. *Food Chem.* **2016**, *212*, 453–459. [[CrossRef](#)] [[PubMed](#)]
44. BeMiller, J.N.; Whistler, R.L. *Starch: Chemistry and Technology*; Elsevier Science: Amsterdam, The Netherlands, 2009; pp. 213–217. ISBN 978-0-12-746275-2.
45. Santelia, D.; Lunn, J.E. Transitory starch metabolism in guard cells: Unique features for a unique function. *Plant Physiol.* **2017**, *174*, 539–549. [[CrossRef](#)] [[PubMed](#)]
46. Crumpton-Taylor, M.; Grandison, S.; Png, K.M.Y.; Bushby, A.J.; Smith, A.M. Control of starch granule numbers in *Arabidopsis* chloroplasts. *Plant Physiol.* **2012**, *158*, 905–916. [[CrossRef](#)] [[PubMed](#)]
47. Jane, J.; Shen, J.J. Internal structure of the potato starch granule revealed by chemical gelatinization. *Carbohydr. Res.* **1993**, *247*, 279–290. [[CrossRef](#)]
48. Tetlow, I.J.; Beisel, K.G.; Cameron, S.; Makhmoudova, A.; Liu, F.; Bresolin, N.S.; Wait, R.; Morell, M.K.; Emes, M.J. Analysis of protein complexes in wheat amyloplasts reveals functional interactions among starch biosynthetic enzymes. *Plant Physiol.* **2008**, *146*, 1878–1891. [[CrossRef](#)] [[PubMed](#)]
49. Liu, F.; Ahmed, Z.; Lee, E.A.; Donner, E.; Liu, Q.; Ahmed, R.; Morell, M.K.; Emes, M.J.; Tetlow, I.J. Allelic variants of the amylose extender mutation of maize demonstrate phenotypic variation in starch structure resulting from modified protein–protein interactions. *J. Exp. Bot.* **2012**, *63*, 1167–1183. [[CrossRef](#)] [[PubMed](#)]
50. Hayashi, M.; Crofts, N.; Oitome, N.F.; Fujita, N. Analyses of starch biosynthetic protein complexes and starch properties from developing mutant rice seeds with minimal starch synthase activities. *BMC Plant Biol.* **2018**, *18*, 59. [[CrossRef](#)] [[PubMed](#)]
51. Yun, S.-H.; Matheson, N.K. Estimation of amylose content of starches after precipitation of amylopectin by concanavalin-A. *Starch Stärke* **1990**, *42*, 302–305. [[CrossRef](#)]
52. Morrison, W.R.; Laignelet, B. An improved colorimetric procedure for determining apparent and total amylose in cereal and other starches. *J. Cereal Sci.* **1983**, *1*, 9–20. [[CrossRef](#)]





>PGSC0003DMP400021398 GBSS Protein

MASITASHHFVSRQSSTLDTKSTLSQIGLRNHTLTHNGLRAVNKLDGLQSRNTTKVTPKM  
ASRTETKRPGCSATIVCGKGMNLI FVGTEVGPWSKTGGLGDVLGGLPPALAARGHRVMTI  
SPRYDQYKDAWDTSVAVEVKVGDSEIEIVRFHCHYKRGVDRVFDHPMFLEKVGKTGSKI  
YGPKAGLDYLDNELRFSLLCQAAL EAPKVLNLSNYSFSGPYGEDVLF IANDWHTALIPC  
YLKSMYQSRGIYLNKAVAFCIHNIAYQGRFSFSDFP LLNLPDEF RGSFDFIDGYEKPVKG  
RKINWMKAGILESHRVTVSPYYAQELVSAVDKGVLEDSVLRKTCITGIVNGMDTQEWNP  
ATDKYTDVKYDITVMDAKPLLKEALQAAVGLPVDKKIPLIGFIGRLEEQKGS DILVAAI  
HKFIGLDVQIVVLGTGKKEFEQEIEQLEVL YPNKAKGVAKFNVPLAHMITAGADFMLVPS  
RFEPICGLIQLHAMRYGTVPICASTGGLVDTVKEGYTGFHMGAFNVECDVDPADV LKIVT  
TVARALAVYGTLAF AEMIKNCMSEELSWKEPAKKWETLLLGLGASGSEPGVEGEEIAPLA  
KENVATP

>PGSC0003DMP400002384 SS2 Protein

MENSILLHSGNQFHPNLP LLALRPKKLSLIHGSSREQMWRIKRVKATGENSGEAA SADES  
NDALQVTIEKSKKVLAMQQDLLQQAERRKVVSSIKSSLANAKGTYDGGSGSLSDVDIPD  
VDKDYNTVTPSTAATPITDVKNTPPAISQDFVESKREIKRDLADERAPPLSRSSITASS  
QISSTVSSKRTLNVPPETPKSSQETLLDVNSRKS LVDVPGKKIQSYMPSLRKESSASHVE  
QRNENLEGSSAEANEETEDPVNIDEKPPPLAGTNVMNIILVASECAPWSKTGGLGDVAGA  
LPKALARRGHRVMVAPRYDNYPEPQDSGVRKIYKVDGQDVEVTFQAFIDGVDFV FIDS  
HMF RHIGNNIYGGNRVDILKRMVLFCKAAIEVPWHVPCGGV CYGDGNLVIANDWHTALL  
PVY LKAYYRDNGIMNYTRSVLVIHNI AHQGRGPLEDFS YVDLPPHYMDPFKLYDPVGGEH  
FNIFAAGLKTADRVTVSHGYSWELKTSQGGWGLHQIINENDWKLQGIVNGIDTKEWNPE  
LDVHLQSDGYMNYSLDTLQTGK PQCKAALQKELGLPVRDDVPLIGFIGRLDPQKGVDLIA  
EAVPWMMGQDVQLVMLGTGRRDLEQMLRQFECQHNDKIRGWVGF SVKTSHRITAGADILL  
MPSRFEPICGLNQLYAMKYGTIPVHAVGGLRDTVQPFDPFNESGLGWTFSRAEASQLIHA  
LGNCLLTYREYKKSWEGIQTRCMTQDLSWDNAAQNYEEVLI AAKYQW

>PGSC0003DMP400013565 GWD Protein

MSNSLGNLLYQGF LTSTVLEHKSRI SPKCVGGNSLFQQQVISK SPLSTEFRGNRLKVQK  
KKIPMEKKRAFSSSPHAVLTTDTSSELA EKFSLEGNIELQVDVRPPTSGDVSFVDFQATN  
GSDKLFLHWGAVKFGKETWSLPNDRPDGTVKYKNKALRTPFVKSGSNSILRLEIRGTAIE  
AIEFLIYDEAYDKWIKNNGGNFRVKLSRKEIRGPDVSVPEELVQIQSYLRWERK GKQNYT  
PEKEKEEYEAARTELQEEIARGASIQDIRARLTKTNDKSQSKEEPLHVTKSEIPDDLAQA  
QAYIRWEKAGKNYPPEKQIEELEEARELQLELEKGITLDEL RKKITKGEIKTKAEKHV  
KRSSFAVERIQRKKRDFGQLINKYPSSPAVQVQK VLEPPALSKIKLYAKEKEEQIDDPI  
LNEKIFKVDDGELLVLVAKSSGKTKVHLATDLNQPITLHWALSKSRGEWMVPPSSILPPG  
SII LDKAAETPFSASSDGLTSKVQSLDIVIEDGNFVGM PFVLLSGEKWIKNQGSDFYVD  
FSAASKLALKAAGDGS GTAKSLDKIADMESEAQSFMHRFNIAADLIEDATSAGELGFA  
GILVWMRFMATRQLIWNKNYNVKPREISKAQDRLTDL LQNAFTSHPQYREILRMIMSTVG  
RGGEDVVGQIRDEILVIQRKNDCKGGMEEWHQKLHNNTSPDDVVICQALIDYIKSDFD  
LGVYKTLNENGITKERLLSYDRAIRSEPNFRGDQKNGLLRDLGHYMR TLKAVHSGADLE  
SAIANCMGYKTEGEGFMVGVQINPVSGLPSGFQGLLHFVLDHVEDKNVETLLEGLLEARE  
ELRPLLLKPNR LKDLLFLDIALDSTVRTAVERGYEELNNANPEKIMYFISLVLENLALS  
VDDNEDLVYCLKGWNQALSMSNGGDNHWALFAKAVLDRTRLALASKAEWYHLLQPSAEY  
LGSILGVDQWALNIFTEEIRAGSAA SSSLLNRLDPVLRKTANLGSWQIISPVEAVGYV  
VVVDELLSVQNEIYEKPTILVANSVKGEE EIPDGAVALITPDMPDVLSHVSVRARNGKVC  
FATCFDPN ILADLQAKEGRILLKPTPSDIYSEVNEIEFQSSSNLVEAETSATLRLVKK  
QFGGCYAI SADEFTSEMVGAKSRNIAYLKGKVPSSVGIPTSVALPFGVFEKVLSDDINQG  
VAKELQILTKKLSEGNFSALGEIRTTILDLSAPAQLVKE LKEKMQSGMPWPGEDEGPKRW  
DQAWMAIKKVWASKWNERAYFSTRKVKLDHDYLCMAVLVQE IINADYAFVIHTTNPSSGD  
DSEIYAEVVRGLGETLVGAYPGRALS FICKKKDLNSTQVLGYPSKPIGLFIKRSIIFRSD  
SNGEDLEGYAGAGLYDSVPMDEEEKVVIDYSSDPLITDGNFRQTI LSSIARAGHAIEELY  
GSPQDIEGVVRDGIYVVQTRPQM

>PGSC0003DMP400029040 PWD Protein

MDSMHLSHCYSALNVKKRQPQIKISKQFNVVHQISHRSVRNLTLLPTRNLGFFMDRRVKG  
IVCGVSSVETRENQNGKGNKNSGSSTEKVLRFRLDHQVEYGEHIAVLGSAKELGSWKK  
NIMMDWTENGWIGELEVRSGETLEYKFVIVGKDKKMLWENGSNRILKLPEGGGFELVCQW  
NVTDEPVNLLPLDPFEVEKVEETSNGAKIISQAAPDVVTSPFVEQWQGRAASFVRSN  
DQLDSDKNRKWDTSGLTGISLKLVEGDKNARNWWRKLEVVRELVENMDSHRLEALTYA  
AVYLKWINTGQIPCLEDDGGHHRPNRHAEISRLIFREVEKVLRRDRTLQEILVIRKMQPC  
LPSFKAEFTASVPLTRIRDIAHRNDIPHDLKQEIKHTIQNKLHRNAGPEDLVSTEAMLER  
ITKQPGQYSEAFVEQFKIFHNELKDFFNAGSLDEQLESMRESLDGSSLSMLSSFLESKKE  
LVRLDEKHNVSERTERTGILVRTINSLNALREVIKAGLESGLRNDAPDASIAMRQKWLCE  
IGLEDYAFVLLSRFVNAVEALGGADWLAENVTVKNISSWNDPIGALTVGIQQLGISGWKP  
EECKAVGNELLSWKERGISEIEGSEDGKTIWALRLKATLDRSRLTEEYSETLLQIFPEK  
VQILGKSLGIPENTVRTFTEAEIRAGVVFQVSKLATLLKAVRRTIGSSGWDVLPVGDFAF  
GELIQVDRIIPGTLPSSATGPVILLVNKADGDEEVTAAAGSNISGVLLQELPHLSHLGVR  
ARQEKVVFVTCDDDDKVSQVDRQLLGKYVRLASSTGVKLTASSEKAGGVSPNKLPSNA  
SSAGATSSDSSASSIAVKSSQVKEVGPTRGVIPLVDADIQTSGAKAASCAQLASLAISST  
KVYSDQGAPASFNVPAGAVIPFGSMETALMKNLMEFTLLVEQIETAEIDGGELDKHCE  
DLQKLISLLPGQDVIESLGEVFPGNARLIVRSSANVEDLAGMSAAGLYDSIPNVSPSDP  
IRFGHAVARVWASLYTRRAVLSRRAAGVSQKDATMAVLVQEMLSPDLSFVLHTLSPTDNN  
HNFIEAEIAPGLGETLASGTRGTPWRLLSSGKFDDTVRTLAFANFSEEMVVGGNPADGEV  
ILLTVDYSKKPLTIDPIFRRQLGQRLGAVGFYLERKFGSPQDVEGCLVGNEIFIVQSRPQ  
PQ

>PGSC0003DMP400014590 SS4 Protein

MEMKISNCFPTQKWCAGAPFSRQVNVRFYPLPSQRLLPASSKMRQRNFSLQNKKKQTKT  
INIERPPVDLQSSDDVSDTEKMSKRSLSASNQEVSIENVDLTLTETKSSDEITYSSVD  
SNEEGRPSSVHLKDLIGMIRNAEKNIHLLNEARVHALEELQKVLGEKEDLHGKINILEMK  
LAETDARLRVASQEIKHVELLEDQLEKIKNELSSRGSEENVLHVNSVPLSHSDSVNSL  
SEEFDSLKENILLKQDLQAMKSGLSLVKETDERILMLEKERSVLESSLSELESKLAASQ  
EDVSELSLKLCKKLYEKVEHLQALLAKATKQADQALSVLQQNQELRKKVDRLEESLEE  
ASIYKLSSEKLQYNEQMQQKIKLLDERLQRSDEEQSYVQLHQDSVKEFQDTLDNLKNE  
TKKKALDEPVNEMPSEFWSWLLMIEGWSMEKKISKDDAKLLRELVWKRDSICDAYMSC  
KEKNDREILAAFLRFTSSPTRPGLHIIHIAEMAPVAKVGGGLGDVVAGLGKALQEKGHLV  
EIVLPKYDCMQYESIKDMKVLVVVESYFDGRLYKNNIWTGIVEGLPVYFIEPQHPGKFF  
CRGQLYGEHDDFKRFSFFSRVALELLQAEKRPDIHCHDWQTAFAVPLYWEIYVPGKLD  
SARICFTCHNFYQGTAAASELTSCGLDAYHLNKPDRMQDNSANDRINPVKGAIVFSNIV  
TTVSPTYAQEVRSAGGKGLHATINSHSKKFAGILNGIDTDAWNPASDNFLKVQYSASDI  
DGKIQNKEALRRLLGLSSDFRQPLVSCITRLVPQKGVHLIRHAIYRTELEGGQFVLLGS  
SPVPHIQREFEDIGNHFQNHARLVLYDEALSHLIYAASDMLIIPSI FEP CGLTQMIA  
MRYGSIPIARKTGGLNDSVFDVDDDSIPLQFRNGFTFVTADEQGFNNALERAFNYMNA  
ETWKELVQKMSIDFSWDSSASQYEELYEKAVLRARAASRS

>PGSC0003DMP400032329 SS1 Protein

MGSLQTPTNLSNKSCLCVSGRVVRLRVERQVGLGFSWLLKGRNRKQVSLCVTSSVSDG  
SSIAENKVVSEGLLLGAERDGSVGVGFQLIPHSVAGDATMVESHDIVANDRDDLSEDETE  
EMEETPIKLT FNII FVTAEAAPYSKTGGLGDVCGSLPMALAARGHRVMVSPRYLNGGPS  
DEKYANAVDLVDRATVHCFGDAQEVAFYHEYRAGVDWVFDHSSYCRPGTPYGDYGAFG  
DNQFRFTLLSHAACEAPLVLPGLGGFTYGEKCLFLANDWHAALVPLLAAKYRPGVYKDA  
RSIVAIHNIHQGVPAVTYNNLGLPPQWYGAVEWIFPTWARAHALDTGETVNVKGAIA  
VADRILTVSQGYSWEITTPEGGYGLHELLSSRQSVLNGITNGIDVNDWNPSTDEHIASHY  
SINDLSGKVQCKTDLQKELGLPIRDCPLIGF IGR LDYQKGVDIILSAIPELMQNDVQVV  
MLGSGEKQYEDWMRHTENLFKDKFRAWVGFNVPVSHRITAGCDILLMPSRFEP CGLNQLY  
AMRYGTIPIVHSTGGLRDTVKDFNPYAQEGNGEGTGWTF SPLTSEKLLDTLKLAIGTYTE

HKSSWEGLMKRGMGRDYSWENAAVQYEQVFTWAFIDPPYVR

>PGSC0003DMP400028510 LESV Protein

MATRFPATTHSRVVAASSGEPKVEYVPSVSTRKWQVVTERRRKNRSRIKATDSDFSLEMWK  
RAMERERKSAEFKRIAENIAPPEVEESPEILEKKTEEFNKILQVSPEERDKVQSMQIIDR  
AAAALAAAKALIKENPLPRKDDDEADKNNKQGATGNVISIVPRSGTMMGTPGPSFWSWTP  
PSDSSFDDIQMKSDVSLSPDPPSPVIEKERSPDFLSIPFQSATIDRKHSPLPPLQSHLE  
VENLEDSSSTPEIPHQVEERELGILFSANAAEAAYALHQENEASSEGISPDGSRWWKETG  
TERRPDGVVCKWTLTRGISADKTVEWEDKYWEAADEFHGKELGSEKSGRDAAGNVWHEFW  
KESMWQSGGLVHMEKTADKWGRNNKGEWEHEKWWEHYGAGGQAEKWAHKWCSIDPNTPLD  
AGHAHVWHERWGEKYDGKGGSIKYTDKWAERFEVDGWSKWGDKWDENFDLNGHGKQGET  
WWAGKHGERWNRWTWEGHNGSGVWHKYGKSSDGEHWDTHVNEETWYERFPHFGFHCFQN  
SVQLREVKRPSDWP

>PGSC0003DMP400023948 SS6 Protein

MDFTSGLSFPSSLAPNGNCSSNNCCSSFNHRSCYNNISSKANLLKKQSYVFSKYQLRNF  
KIRRVPRIEAQLKAEQLESSEAGCESDGLQWSPNDEIPFWKGFPSWDLSSDASAGVEQ  
DSDLLHIVHVTAEMAPIAKVGGGLGDVVTGLGRACLNRGHKVDVMIPFYECIPKHCINELA  
LMKTYNSYHDGNWVACKAYRGEVSGVPVILIEPSNHFFKGNIIYGGSYNELDAYLFFSRA  
CLEWMQVNGTQPDIIHVHEWQTGALPLFYWDMYHFLSLQKPRIVLTIHNMEHYGECRQEQ  
LSKFGLDGSAYATEDKAVDDRTVGHNPERLSLLKGGIVYSNAIVTVSPTYLKETLCSGWL  
SGALMRNRDKYSGILNGIDTEMWNPATDIYLPKFDGNYFCDVVEVLRHLL

>PGSC0003DMP400053324 SSV Protein

MVLTSLIGSKEASEFRWLVMDSIGSIADYFSEKMHKQDTELLVELRHFPRKSRKSGYHII  
HICTEMAPVSVGSLAPYVTGLSRALQRKGNLVEVILPKYASLNLNEVHGLREVEAEFHS  
YFNGQLHGNRIWTGVVCGIGVTFIEPLYSAFFGCENIYGYSNDFERFTYFSRASLDYII  
KAGKHPDVLHIHNWETSIVGPLFWDVFNQGLGGTRIMLTCSQSFESQCVEQPEKLALCGL  
DPYGLHCSDRLQDNNKSHLVNVLKAGVVYSNNVIIMSSMQTKGQIIHATSHGLEPTLTIH  
KDKLVVAPPGFSSAWDPSVDMFLPQNYADLKGKSVCKVSLQQHLGLQEKASIVLVGFT  
LVEDLPAKVYNQTLIMGISQRLSLLLAFSL

>PGSC0003DMP400017605 ESV1 Protein

MAASASSRGFTAPIGATVEVFPTSHNGFVEFRKKRSSIGGFFSIQCCSDSVMPIRGGG  
GSGNGVDKGGDGRDLSKKITANYMRTQASSPMPFASPQSRFVSKPEKFFSRCTPRNSGPK  
SRDSPPKRDTGIANEKDWGISMLNDNVNETGVNEDGSTWYRESGEDLDNGYRCRWTRMG  
GQSSDGTLEWKETWWEKSDWTGYKELGVEKSGKNVEGDAWETWREVLHQDEWSNLARIE  
RSAQKQAKSGTENAGWYENWWEKYDAKGWTEKGAHKYGRLEQSWWEKWGEHYDGRGSVL  
KWTDKWAETELGTKWGDKWEKFFAGIGSRQGETWHVSPVGDRWSRTWGEEHFGNGKVHK  
YGKSTTGESWDIVDEGTYYEAEPHYGWADVVDSSQLLSIQPRDRPPGVFSPMDFGPSL  
PPEDEFPPTLSQ

>PGSC0003DMP400047519 SEX4-Like Protein

MNCLQNLPRSSGLPLRSFTGNSRKPSTVVS LGMTKFADQRLSIVAQVVS GPESSTEKDE  
EKSDTYSHDMTAAAGAVLTYRHDLMNMFIRPDLIVGSLQTPEDVDKLRISIGVKTI  
LQQNPDLEYFGVDINAIREYANKCGDIEHLRAEIRDFDAFDLRLRLPAVISILNAINRN  
GGVTYIHTAGIGRAVALTYMFVWQSYKLEAFD LMSKRSCFPKLD AIKSATADVLT  
GLKKMPVTLTWHGDNCTTVEISGLDIGWQRTPLKFDEERGLWTLQKDLHEGIYKYIV  
DGEWICNEFEPITSPNKDGHVNNYVEVLDPDNITSVAVRKRLTGDDPDLTSDERLIE  
QFLEAYADVE

>PGSC0003DMP400026742 SEX4 Protein

MNCLQNLPRSSGLPLRSFTGNSRKPSTVVS LGMTKFAEQRLSIVAQVVS GPESSTAKDK  
DAGKSDTYSHDMTAAAGAVLTYRHELMNMFILPDLIVGSLQTPEDVDKLRISIGVKTI

FCLQQNPDLEYFGVDIEAICEYANKCGDIEHLRAEIRDFDAFDLRLRLPAVIRKLNNAIN  
RNGGVTYIHCTAGLGRAPAVALTYMFVWQCYKLTDAVSLMSKRPCFPNLDAIKSATADI  
LTGLKKTPTLWRGDNCTTVELSGLDIGWWRIPKFDKEQGLWTLQRELPEGHYKEY  
IVDGEWICNEFEPITSPNKDGYVNNYVEVLDETPDNITSAAIRKRFSGDDPDLTSYEGLI  
IGQFLEGYADVE

>PGSC0003DMP400050640 LSF2 Protein

MEEYNLAMKRMMRNPYEHHELGMNYTLITEDLIVGSQPQKVEDIDYLKEEENVAFILNL  
QQDKDIEFWGIDLQSIITRCSELGIHMRPARDFDPDSLRSVLPKAVSSLEWAISEGG  
RVVYHCTAGLGRAPAVSIAYMFWFCGMDLNTAYDTLVSKRCPGPNKRSIQGATYDLAKND  
QWKEPFENLPDYAFADVADWERKLIQDRVRLRDT

>PGSC0003DMP400053308 PTST1 Protein

MASYNRKGFRMHKSFNSNIPCTSWKIFCAPENLEKRFSVVSEKIADAGLSDPEQPLR  
SEELRLLADAERSKLLKLLSEANRYNRLKRELQAKEDALVNFKSELSVTELEIQVLAR  
LAEEIAKSAIPAGSRKIKGRYIQSHLLSRLEAIREKLKEQIKGVEAVQAKEVPLSWVGVA  
ESVQVMGSFDGWSQGEHLSPEYTGSYMNFSAFLRPGRYEIKFMVDDEWKLSPELPTTG  
EGLTKNNLLVVE\*

>PGSC0003DMP400016823 Kunitz-type proteinase inhibitor Protein

MKCLFLLCLCLVPIVFSSTFTSQNPINLPSDATPVLDVTGKELDSRLSYRIISTFWGAL  
GGDVYLGKSPNSDAPCANGIFRYNSDVGPSGTPVRFIGSSSHFGQGFENELLNIQFAIS  
TSKLCVSYTIWKVGDYDASLGTMLLETGGTIGQADSSWFKIVKSSQLGYNLLYCPVTSTM  
IGILQN

>PGSC0003DMP400017933 Serine protease inhibitor 7 Protein

MKCLFLLCLCLVPIVFSSTFTSKNPINLPSDATPVLDVAGKELDSRLSYRIISTFWGAL  
GGDVYLGKSPNSDAPCANGIFRYNSDVGPSGTPVRFIGSSSHFGQGFENELLNIQFAIS  
TSKLCVSYTIWKVGDYDASLGTMLLETGGTIGQADSSWFKIVKSSQFGYNLLYCPVTSTM  
SCPFSSDDQFCLKVGVVHQNGKRRLLALVKDNPLDVSFKQVQ

>PGSC0003DMP400017950 Cysteine protease inhibitor 1 Protein

MKSINILSFLLSSTLSLVAFARSFTSENPIVLPTTCHDDNLVLPEVYDQDGDPLRIGE  
RYIINNPLIGAGAVLYNIGNLQCPNAVLQHMSIPQFLGKGTVPVVRKSESDYGDVVRQ  
MTAVYIKFFVKTTKLCVDQTVWKVNDDEQLVVTGGKVGNEIDFKIRKIDLVTGSSKNVY  
KLLHCPSHLQCKNIGGNFKNGYPRLVTVDDDKDFIPVFVFIKA

>PGSC0003DMP400017953 Kunitz-type tuber invertase inhibitor Protein

MKSINILSFLLSSTLSLVAFARSFTSENPIVLPTTCHDNLVLPEVYDQDGNPLRIGERY  
IINKPLLIGAGAVLYNIGNLQCPNAVLQHMSIPQFLGKGTVPVVRKSESDDGDVVRMT  
AVYIKFFVKTTKLCVDETVWKVNDDEQLVVTGGNVGNENDIFKIKKDLVIRGMKNVYKLL  
HCPSHLQCKNIGSNFKNGYPRLVTVDDDKDFIPVFVFIKA\*

>PGSC0003DMP400017939 Cysteine protease inhibitor 1 Protein start\_err

TSHDNLVLPEVYDQDGHPLRIGERYIINNPLTGGGAVYLANIGNLQCPNAVLHHSIPQF  
MGKGTVPVVRKSESDYGDVVRVMTGVYIKFFVKTTKLCVNDDEGLVVTGGNV  
GNENGIFKIMKTDLVMPGGMKNVYKLLHCPAHLCKDIGGTFKDGYPRLVTVDNDKDFIP  
FVFVFIKAKANISFYRSM

>PGSC0003DMP400016824 Aspartic protease inhibitor 5 Protein

MKCLFLLCLCLVPIVFSSTFTSQNLIDLPSESLPKPVLDTNGKELNPSSYRIISIGR  
GALGGDVYLGKSPNSDAPCPDGVFRYNSDVGPSGTPVRFIPLSTNIFEDQLLNIQFNIP  
VKLCVSYTIWKVGNLNAFRMTMLLETGGTIGQADNSYFKIVKSSKIGYNLLSCPFTSIIC  
LRCPEDQFCAKVGVIQNGKRRLLALVNNENPLDVLVQEV\*

>PGSC0003DMP400008029 Proteinase inhibitor type-2 P303.51 Protein  
MDVHKEVNFVAYLLIVLGLLVLSAMEHVDAKACTLECGNLGFGICPRSEGSPEENRICTN  
CCAGYKGCNYYANGAFICEGQSDPKPKACPLNCDPHIAYSCKPRSEKSLIYPTGCTT  
CCTGYKGCYYFGKNGKFVCEGESDEPKANMYPAM\*

>PGSC0003DMP400017942 Cysteine protease inhibitor 1 Protein  
MKSINILSFLLSSTLSLVAFAQSFTSENPIVLPTTSHDDNLVLPEVYDDEGHPLRIGE  
EYIINNPLLGAGAVYLDNIGNLTCPNAVLQHISTFNLSGDGTPVKFVRKSESDDGDVVRV  
MTVVYIKFFPNIPNLLCVNENVWKNDEELVVTGGNVGNENDIFKIMKTDLVTPEGSKYV  
YKLLHCPAHECKNIGVNFKDGYPRLTTVDDDKNFAPFVFSKAKNAN\*

>PGSC0003DMP400017952 Cysteine protease inhibitor 9 Protein  
MLQHMSIPQFLGEGTPVMFVRKSESDYGDVVRIMTGVIKFFVKTTKLCVDQTGGKVGNE  
NDIFKIRKTDLVTGGSKNVYKLLHCPSHLQCKNIGGNFKNGYPRLVTVDDDKFIPFVF  
IKA\*

>PGSC0003DMP400016822 Aspartic protease inhibitor 8 Protein  
MKCLFLLCLLPIVVFSSFTSQNLIDLPSPLPKPVLDTNGKELNPSSYRIISIGR  
GALGGDVYLGKSPNSDAPCPDGFRYNSDVGPSVTPVRFIPLSGGIFEDQLLNIFNIAT  
VKLCVSYTIWKVGNLNAYFRMTLLETGGTIGQADSSYFKIVKLSNFGYNLLYCPITPPFL  
CPFCRDDNFCAKVGVVIQNGKRRRLALVNEPLDVLQEV\*

>PGSC0003DMP400021964 Thioredoxin Protein  
MAAVLETITVPRASALPSSSLAPVAGYSFGPRSSVRFSSQFSGLKIQPIRSSVSTSSCSK  
IIPRGGRRVCEAQNATAVEVAASDKTWKTLVVESTVPVLVEFWAPWCGPCRMIHPVIDEL  
AKEVVGKLTCLKLNDESPIATEFGIRSIPTVMIFKNGEKDAIIGAVPKSTLTTSIEK  
FL

>PGSC0003DMP400028845 Stem 28 kDa glycoprotein Protein  
MRCFTILFFFSIVVALASNVEEIVSQVVEIHRLRPQTGSAGYVTPQLDCLSWRLAVET  
NNLQNWKLVPKECTNYVGHYMLGKQYRRDCEYVAKQAIYAKTLKLGKGDGKDVWVFDIDE  
TTLSNSPPYARSDVAFGAIYNNTKFSAWLAEGKAPAIPSILGVYKTVLSLGIKPVFITG  
TRENFKQVRIVNLKKGYSNWAALILKGENDSGSAVQFKSSKRTLVKAGYRIVGNIGDQ  
WTDLIGENVGARTFKVPDPMYYIS\*

>PGSC0003DMP400045625 Glutathione peroxidase Protein  
MASLSSSSVLSPICTFFSQPKLNLYCNSSSSSSSYKPTCSSSKSSIFQNGLSLLTTKSCGS  
SLKSKNMGNTIYARAATEKTIYDFTVKDIDDKDVPLSTFKGKVLIVNVASRCGLTTSNY  
TELSRVYEKYKNQGLEILAFPCNQFGAQEPGSNPEIKQFACTRFKAEPFIFDKVDVNGPN  
TAPVYQFLKSSAGGFLGDLVKWNFEKFLVDKNGKVVERYPPTTSPLQIEKDIQKLLAA\*

>PGSC0003DMP400009292 CYP20-2 PPIase chloroplastic Protein  
MAATFATVSNLGSISVPRVANPTYAASPCLLKSSFSSSSLFGGSLRLETSSNRSVHRKTG  
GSSGSIQAAVEAVDLQSKVTHKVYFDISIGNPVGKLAGRIVIGLYGDDVPQTAENFRALC  
TGEKGFQKDSAFHRVIKDFMIQGGDFDKGNGTGGKSIYGRTEFKDENFKLVHTGPGIVSM  
ANAGPNTNGSQFFICTVKTPLDQKHVVFQVLEGM DIVKLI ECQETDRGDRPRKRVVIS  
DCGELPIA

>PGSC0003DMP400011774 Ci21A Protein  
MEEKHHLHLLFHHKDKAE EGPVDY EKETKHHKHL EQIGKLGTV AAGAYALHEKHEAKKDP  
EHAHKHKIEEEIAAAA AVGAGGFALHEHHEKDKAKKEQKKAEGGHHHLF

>PGSC0003DMP400051081 Beta-tubulin Protein

MREILHIQGGQCGNQIGSKFWEVICDEHGIDPTGRYNAGNAGASDLQLERINVFNEATA  
GRFVPRAVLMDLEPGTMSIRSGPSGQIFRPDNFVFGQSGAGNNWAKGHYTEGAELIDSV  
LDVVRKEAEYCDCMQGFQVCHSLGGGTGSGMGTLLISKIREEYPDRMMLTFSVFPSPKVS  
DTVVEPYNATLSVHQLVENADECMVLDNEALYDICFRTLKLTNPTFGDLNHLISATMSG  
TCCLRFPGQLNSDLRKLAVNLI PFPRLHFFMVGFAPLTSRGSQQYSSLTVPELTQQMWDA  
KNMMCAADPRHGRYL TASAMFRGKMSTKEVDEQMINVQNKSSYFVEWIPNNVKSSVCDI  
PPTGLKMASTFIGNSTSIQEMFRRVSEQFTAMYRRKAF LHWYTGEGMDEMEFTEAESNMN  
DLVAEYQQYQDATAEVEEDFDDGVDEHYEGQ

>sp|Q43846|SS3

MDVPFPLHRSLSCTSVSNAITHLKIPI LGFVSHGTTLSLVQSSSWRKDGMVTGVSF  
SICANFSGRRRRKVVSTPRSQSS PKGFVPRKPSGMSTQRKVQKSNQDKESKST  
STSKESISEISNQTVEARVETSDDDTKGVVRDHKFL EDEDEINGSTKSISM  
SPVRVSSQFVESEETGGDDKDAVKLNKSKRSEESGFIIDSVIREQSGSQGETNASS  
KGSHAVGTKLYEILQVDVEPQQLK ENNAGNVEYKGPVASKLLEITKASDVEHTES  
NEIDDLTNSFFKSDLIEEDEPLAAGTVETGDSSLNLRLEMEANLRRQA IERLAEEN  
LQGIRLFCFPEVVKPDEDVEIFLNRGLSTLKNESDVLIMGAFNEWRYRSFTTRLTETH  
LNGDWWSCKIHVP KEAYRADVFVFNQDVYDNDGNDFSITVKGGMQIIDFENFLLEEK  
WREQEKLAKQAERERLAEEQRRIEAEKAEIEAD RAQAKKEAAKKKKVLRRELMV  
KATKTRDITWYIEPSEFKCEDKVRLYNKSSGPLSHAKDLWIHGGYNNWKDGLSIVK  
KLV KSERIDGDMWYEVVIPDQALFDWV FADGPPKHAIAYDNNHRQDFHAI VPNHIPE  
ELYWVEEHQIFKTLQEERLREA AMRAKVEKTALLKTETKERTMKSFLLSQKHVYTE  
PLDIQAGSSVTVYYPANTVNLNGKPEIWFRCFNRWTHRLGPLPP QKMSPAENGTHV  
RATVKVPLDAYMMDFVFSEREDGGIFDNKSGMDYHIPVFGGVAKEPPMHIVHIAVEM  
APIAKVGGLD VVTSLSRAVQDLNHNVDIILPKYDCLKMNNVKDFRFHKNYFWGGTEI  
KVWFGKVEGLSVYFLEPQNGLFSKGCVYGCSD GERFGFFCHAALEFL LQGGFSPDI  
IHCHDWSSAPVAWL FKEQYTHYGLSKSRIVFTIHNLEFGADLIGRAMTNADKATT  
V SPTYSQEVSGNPVIAPHLHKFHGIVNGIDPDIWDLNDKFIPIPYTSENVVEGKTA  
AKEALQRKLGKQADLPLVGIITR LTHQGIHLIKHAIWRTLERNQVVLG SAPDPRV  
QNNFVNLANQLH SKYNDRARLCLTYDEPLSHLIYAGADFILVPSI FEPGGLTQL  
TAMRYGSI PVVRKTGGLYDTVFDVDHDKERAQQCLEPNGFSFDGADAGGVYALN  
RALSAWYDGRDWFNS LCKQVMEQDWSWNRPALDYLELYHAARKLE

>sp|P04045|PHS1a\_Protein

MATANGAHLFNHYSSNSRFIHFSTRNTSSKFLTKTSHFRPKRCFHVNNTLSEKIHHP  
ITEQGGESDLSSFAPDAASIT SSIKYHAEFTPVFSPELFPAFFATAQSVRDSLLIN  
WNATYDIYEKLNMKQAYYLSMEFLQGRALLNAIGNLELTGAF AEALKNLGHNLNVA  
SQEPDAALGNGGLGRLASCF LDSLATLNYPAWGYGLRYKYGLFKQRITKDGQEEVA  
EDWLEIGSP WEVVRNDVSYPKIFYGKVGSTGSDGKRYWIGGEDIKAVAYDVPIPGY  
KTRTTISLRLWSTQVPSADFDLSAFNAGEHTKAC EAQANA EKICYILYPGDESEEG  
KILRLKQYTLCSASLQDIISR FERRSGDRIKWEEFPEKVAVQMNDTHPTLCIPELM  
R ILIDLKGLNWN EAWNITQRTVAYTNHTVLP EALEKWSYELMQKLLPRHVEIEAID  
EELVHEIVLKYGSMDLNKLEKLT TMRILENFDLPSSVAELFIKPEISVDDDTETVE  
VHDKVEASDKVVTNDEDDTGKTSVKIEAAA EKIDKKTVPSP EAV IPPKKVRMANL  
CVVGGHAVNGVAEIHSEIVKEEVFNDFYELWPEK FQNKTNVTPRRWIRFCNPPLSA  
IITKWTGTEDWV LKTEKLAELQKFADNEDLQNEWREAKRSNKIKVVSFLKEKTGYS  
VVPDAMFDIQVKRIHEYKRQLLNIFGIVYRYKKMKE MTAERKTNFVPRVCFGGKA  
FATYVQAKRIVKFITDVGATINHDP EIGDLLKVVFPDYNVSVAEL LIPASDLSEH  
IST AGMEASGTSNMKFAMNGCIQIGTLDGANVEIREEVEENFFLFGAQAHEIAGL  
RKERADGKFPDERFEVKEFVRS GAF GSYNYDDLIGSLEGNFGFRADYFLVGKDF  
PSYIECQEKVDEAYRDQKRWT TMSILNTAGSYKFS SDRITHEYAKDIWNI EAVE  
IA\*>Sotub09g011090.1.1MVYTLSGVRFPTVPSVYKSN GFSNGDRRANISV  
FLKKHSLSRKILAEKSSYN SESRPSTVAASGKVLVPGTQSDSSSSSTQFEFTET  
SPENSPASTDVSSTMEHASQIKTENDDVGPSSDLTGSVEELDF ASSLQLQEGGK  
LEESKTLNTSEETIIDESDRIRERGI PPPGLGQKIYEIDPLL TNYRQHLDYRYSQ  
YKCLR EADIKYEGG LEAFSRGYEKMGFTRSATGITYREWAPGAQSAALIGDFNWDAN  
ADIMTRNEFGVWEIFLPNNVDGSPAIPHGSRVKIRM DTPSGVKDSIPAWINYSLQ  
LPDEIPYNGIYYDPPEEERYIFQHPRPKPKSLRIYESHIGMSSPEPKINSYVNF  
RDEVLP RIKKLGYNVQIMAIQEHSYASFGYHVTNFFAPSSRFGT PDDLKSLIDKA  
HELGI VVLMDIVHSHASNTLDGLNMF DG TDSCYFHSGARGYHMMWDSRLFN  
YGNWEVLRYL LSNARWWLDEFKFDGFRFDGVTSM MYTHHGLSVGFTGN YEEY  
FGLST DVDAVVYLMLVNDLIHGLFPDAITIGEDVSGMPTFCIPVQDGGVGF  
DYRLHMAIADKWIELLKKRDEDWRVGDIVHTLTN RRWSEKCVSYAESHDQALV  
GDKTIAFWLMDKMYDFMALDRPSTSLIDRGIALHKMIRLVTMGLGGEGYLNFMG  
NEFGHP EWIDFPRAEQHLSGGSVIPGNQFSYDKCRRRFDLGD AEYLRYRGLQE  
FDRAMQYLEDKYE FMTSEHQFISRKNEGDRMIV FERGNLVFVFNHWT  
KSYSDYRIGCLKPGKYKVALDSD DPLFGGFGRIDHNAEYFTFEGWYDDRPRS  
IMVYAPSRTAVVY ALVDKEEEEEEVAVVEEPIEE

>sp|P32811|PHS2\_Protein

MEGGAKSNDVSAAPIAQPLSEDPDIASNIKYHAQYTPHFSPFKFEPLQAYYAATADSVRDRLIKQWNDTYLHYDKVNPK  
QTYYLSMEYLQGRALTNAVGNLDIHNAYADALNKLGGQLEEVVEQEKDAALGNGGLGRLASCFLDSMATLNLPAWGYGLR  
YRYGLFKQLITKAGQEVPEDWLEKFSWEIVRHDVVFPIRFFGHVEVLPSGSRKWVGGEVLQALAYDVPIPGYRTKNTN  
SLRLWEAKASSEDFNLFLFNDGQYDAAAQLHSRAQQICAVLYPGDATENGKLLRLKQOFFLCSASLQDIIARFKEREDGK  
GSHQWSEFPKKVAIQLNDTHPTLTIPELMRLMLDDEGLGWDESWNITTRTIAYTNTHTVLEALEKWSQAVMWKLLPRHME  
IEEIDKRFVATIMSERPDLENKMPMRILDHNATKPVVHMANLCVSSHTVNGVAQLHSDILKAELFADYVSVWPTKFQ  
NKTNGITPRRWIRFCSPESLHIITKWLKTDQWVTNLELLANLREFADNSELHAEWESAKMANKQRLAQYILHVTGVSIDP  
NSLFDIQVKRIHEYKRQLLNILGVIYRYKCLKGMSPEERKNTTPRTVMIGGKAFATYTNKRIVKLVTDVGDVVNSDPDV  
NDYLKVVFPNPNVSAEMLIPGSELSQHISTAGMEASGTSNMKFALNGCLIIIGTLDGANVEIREEIGEDNFFLFGATAD  
EVPQLRKDRENGLFKPDPRFEEAKQFIRSGAFGTYDYNPLLESLEGNSGYGRGDYFLVGHDFPSYMDAQRVDEAYKDRK  
RWIKMSILSTSGSGKFSSDRTISQYAKEIWNIAECRVP

>tr|Q84YG5|\_SOLTU\_ISA3

MIRGPPQIVQKCPDITVNRNIVPRTHRHALQDLRQLRRRDSLRLFSSDHRIKFKCTSEEFQPRLVAAAKLQEEAPQ  
MLDTFSPFKVSPGLAHLGVSETESGINFAIFSQHASAVTLCIILPKSVHDGMIELALDPQKNRTGDIWHICIKELPQGG  
VLYGYRIDGPRNWHGHRFDDSIILVDPYAKLIEGRRVFGDESNKMCRRFFGTDFNSLPFDWGENYKLPNIPKDLVIYE  
MNVRAFTADETSSLDQDQGRGSLGLIEKIPHLLLELGVNAVELLPVFEFDELELQRRPNPRDHMINTWGYSTINFAPMSR  
YASCGGGPVRASWEFKEMVKALHGAGIEVILDVVYNHTNEADDENPYTTFRGIDNKVYYMVDLNNNAQLLNFAAGCNTF  
NCNHPTVMELILESLRHWVTEYHVDGFRFDLASVLCRGTDGTPINAPPLVKAIKDSVLSRCKIIAEPWDCGGLYLVGKF  
PNWDRWAEWNGKYRDDIRRFIKGDAGMKNFATRIAGSADLYRVNKRKPYHSVNFVIAHDGFTLYDLVSYNNKHNDANGE  
GGNDGCNDNFSWNCGIEGETSDANINALRSRQMKNFHLALMVSQGTMMMLMGDEYGHTRYGNNSYGHDTAINNFQWGQL  
EARKNDHFRFFSKMIKFRLSHNVLKENFIEKNDITWLEDNWNYESRFLAFMLHDGNGGDIYLAFAHHSIKTAIPSP  
PRNRSWYRVVDTNLKSPDDFVTEGVSGISKTYDVAPYSAILLEAKQ\*

## Annexe 5

iBAQ N: Concentration [nM] échantillon fraction A										T: Protein IDs	T: Gene
11-1-1	11-1-2	11-1-3	11-2-1	11-2-2	11-2-3	11-3-1	11-3-2	11-3-3			
457.9	465.152	440.253	521.519	519.572	670.8	421.468	418.532	421.373	P04045		PHS1a
11404	10347.9	10836.3	10089.5	10545.4	12895.4	11363.3	10920.7	8918.91	P30924		BE1.1
156541	147206	153707	136098	146557	177282	127735	129363	114783	PGSC0003DMP400002384		SS2
11208	9939.54	11136.6	8744.14	9135.74	11710.5	8700.48	8484.22	8744.08	PGSC0003DMP400009292		CYP20-2
19324.3	17813.7	16784.8	16662.2	17959.4	20590.2	14250.9	16743.4	14199.5	PGSC0003DMP400013565		GWD
2265.49	1664.11	2476.01	1587.07	1229.1	1855.18	1594.76	1574.54	1456.27	PGSC0003DMP400017605		ESV1
10563.6	9488.07	10012.1	9335.32	9751.56	11876.5	10382.4	10127.8	8378.62	PGSC0003DMP400017627		BE3-BE1.2
19468.6	15337.2	20651.8	16235.2	18694.7	20275.7	14767.1	17223.8	13067.9	PGSC0003DMP400017950		Cysteine-protease-inhibitor1
23478.7	21799.5	22530.6	26664.7	26244.8	26793.1	20798.9	30188.7	19382.6	PGSC0003DMP400017953		Kunitz-type-tuber-invertase-inhibitor
2620540	2654170	2638090	2668060	2650750	2580430	2692100	2678740	2718200	PGSC0003DMP400021398		GBSS
13607.7	14498	11654.7	13349.1	15278.5	17935.5	10996	14693.7	12225.6	PGSC0003DMP400021964		Thioredoxin
3550.73	2598.27	3355.03	2023.67	2216.98	2714.83	2983.77	3681.3	3547.35	PGSC0003DMP400026742		SEX4
25676.3	22160.1	21682.5	19781.2	21123.4	26400.3	20490	18450.3	19739.1	PGSC0003DMP400028510		LESV
331.664	318.953	310.095	326.75	290.817	349.75	224.234	248.723	236.997	PGSC0003DMP400029040		PWD
4035.14	3481.89	2790.98	2769.01	2969.54	3401.35	3213.16	3376	2785.71	PGSC0003DMP400032329		SS1
4563.84	6096.14	6584.07	4713.72	5308.24	5403.81	6583.71	6753.58	5496.16	PGSC0003DMP400045625		Gluthatione-peroxydase
3372.41	2690.58	3591.68	2827.72	2766.12	3436.79	2623.33	3026.32	3145.65	PGSC0003DMP400047519		SEX4-like
84.232	88.4392	89.8	94.8977	104.411	98.1034	110.733	126.715	123.09	PGSC0003DMP400049372		PHS1b
43571.9	30017.9	38313.9	45226.7	38304.1	44395.1	27209.7	32389.5	25587.4	PGSC0003DMP400050640		LSF2
2594.46	1740.96	1871.01	2168.82	1706.24	1189.5	1049.54	1114.79	1064.31	PGSC0003DMP400053308		PTST1
5469.13	4748.3	5641.98	5062.69	4943.71	6543.56	3789.89	4401.54	3905.54	Q43846		SS3
1065.28	1134.98	1119.14	1173.18	1180.66	1333.03	866.085	903.389	919.918	Q84YG5		ISA3
5148.15	4343.6	4342.75	4273.28	3885.51	4819.09	4284.99	4872.62	3966.03	Sotub09g011090.1.1		BE2
729.959	1084.22	1665.89	1099.38	1097.2	1721.48	1453.16	52.2527	1377.05	SS4-gene--translation_SS4		SS4
620.353	566.085	826.396	320.581	373.984	351.265	421.331	454.758	443.629	SS6-gene---SS6-translation_SS6		SS6



## iBAQ N: Concentration [nM] échantillon fraction B

31-1-1	31-1-2	31-1-3	31-2-1	31-2-2	31-2-3	31-3-1	31-3-2	31-3-3	T: Protein IDs	T: Gene
1347.38	1553.87	1743.2	550.953	430.412	534.15	1711.04	1869.37	1601.59	P04045	PHS1a
9036.97	10272.8	8096.24	11074.3	9963.94	12189.3	5208.88	4990.46	3159.45	P30924	BE1.1
304535	413150	390387	129601	133993	155385	357324	291988	196345	PGSC0003DMP400002384	SS2
8812.28	5759.33	6134.31	13567.6	11083.6	17028.7	8193.43	7827.25	13491.3	PGSC0003DMP400009292	CYP20-2
15445.6	23352.1	21038	12914.2	12856.3	13834.1	18691.5	15655.2	8911.28	PGSC0003DMP400013565	GWD
3941.19	5589.69	4018.46	2050.86	1051.94	1232.88	6839.91	7192.44	10770	PGSC0003DMP400017605	ESV1
8039.23	9608.86	6938.53	10520.3	9307.59	11580.7	5246.17	5033.34	3064.55	PGSC0003DMP400017627	BE3-BE1.2
1552.09	1437.53	1747.82	967.151	761.198	922.216	1298.11	1176.25	0	PGSC0003DMP400017950	Cysteine-protease-inhibitor1
1669.8	2640.01	1610.31	1542.72	1344.04	1633.63	2372.65	1457.66	0	PGSC0003DMP400017953	Kunitz-type-tuber-invertase-inhibitor
2383130	2201060	2254720	2699480	2703150	2656630	2299190	2403240	2502750	PGSC0003DMP400021398	GBSS
25249.6	23968.8	27091.6	9785.05	9830.92	7825.31	31237.8	23329.4	4948.9	PGSC0003DMP400021964	Thioredoxin
1235.44	1882.97	1172.42	3124.13	3329.88	3704.82	1047.84	765.152	1308.41	PGSC0003DMP400026742	SEX4
62107.7	80846.9	62354.8	18003.2	16610.9	17937	94926.6	90913	157961	PGSC0003DMP400028510	LESV
602.456	531.13	398.649	359.493	280.09	303.92	469.24	417.858	801.433	PGSC0003DMP400029040	PWD
4860.52	6581.25	5083.94	2909.15	2706.96	2836.59	3394.31	3196.61	398.478	PGSC0003DMP400032329	SS1
2557.24	2845.62	4958.01	5148.56	4775.06	6110.56	4831.99	106.552	0	PGSC0003DMP400045625	Gluthatione-peroxydase
1251.75	1693.88	1906.96	2499.59	2727.06	2571.82	1677.68	1636.25	41.4151	PGSC0003DMP400047519	SEX4-like
321.25	322.36	325.706	182.144	101.6	196.138	192.928	274.275	517.395	PGSC0003DMP400049372	PHS1b
80533.3	91028.1	94097.7	34925.4	33548.4	43243.5	46731	24383.8	4703.08	PGSC0003DMP400050640	LSF2
4361.69	5981.84	4475.34	2299.65	2789.96	2595.91	5259.69	4414.47	769.832	PGSC0003DMP400053308	PTST1
18244.5	15915.5	19094.6	6330.65	5637.7	5003.54	16717.8	19199.8	17654.1	Q43846	SS3
490.547	543.613	556.298	1008.55	871.795	882.483	756.545	582.482	1172.91	Q84YG5	ISA3
5469.84	6292.47	6507.12	4327.46	4400.76	5074.1	6157.82	6126	4036.78	Sotub09g011090.1.1	BE2
1337.36	25.1652	5.95255	3.53582	14.271	1184.38	83.4267	19.6759	119.591	SS4-gene--translation_SS4	SS4
91.2746	0	112.528	570.506	549.356	645.593	71.1655	77.9091	168.178	SS6-gene---SS6-translation_SS6	SS6

## iBAQ N: Concentration [nM] échantillon fraction C

50-1-1	50-1-2	50-1-3	50-2-1	50-2-2	50-2-3	50-3-1	50-3-2	50-3-3	T: Protein IDs	T: Gene
1683.7	1277.72	1594.46	1174.03	1286.76	1372.59	325.562	815.264	879.532	P04045	PHS1a
10838.3	7845.68	10911.9	8210.83	8019.55	8698.39	2499.22	4571.54	9027.3	P30924	BE1.1
249813	266118	308958	299346	286973	265467	415344	324807	326751	PGSC0003DMP400002384	SS2
5714.82	5912.39	6940.58	6365.66	6414.23	9330	2812.89	11441.5	8539.36	PGSC0003DMP400009292	CYP20-2
11830.6	12271.5	10543.5	9186.89	9937.11	10034.8	9102.63	9778.05	9870.33	PGSC0003DMP400013565	GWD
555.251	3395.35	3711.4	2596.49	2583.31	2725	0	66.6628	2419.69	PGSC0003DMP400017605	ESV1
9557.75	6932.6	9653.71	6529.88	5970.67	7342.59	2203.87	4041.31	8463.96	PGSC0003DMP400017627	BE3-BE1.2
1198.37	3585.56	4817.45	1907.47	1848.43	1967.51	0	0	0	PGSC0003DMP400017950	Cysteine-protease-inhibitor1
2080.53	2274.17	1217.63	4381.79	3125.25	5639.3	945.757	498.354	325.563	PGSC0003DMP400017953	Kunitz-type-tuber-invertase-inhibitor
2461260	2442740	2375540	2418790	2437610	2427530	2322200	2382410	2349790	PGSC0003DMP400021398	GBSS
22172.7	23034.4	13342	14015.5	15933.9	10340.3	6689.29	9992.28	9122.95	PGSC0003DMP400021964	Thioredoxin
7789.98	4661.81	5866.19	5323.35	5198.16	5447.4	2918.95	6936.78	4698.04	PGSC0003DMP400026742	SEX4
50401.8	57521.8	53759.2	54599.8	51741.7	77945.6	36091.1	50045.4	65751.1	PGSC0003DMP400028510	LESV
361.628	290.41	475.986	416.91	324.686	397.895	173.9	336.953	410.215	PGSC0003DMP400029040	PWD
4790.79	6642.56	5370.19	4375.12	3811.48	3542.03	2467.15	4190.59	5043.08	PGSC0003DMP400032329	SS1
3585.36	4700.12	4993.8	1614.6	1220.83	2052.45	1448.49	8086.61	6420.96	PGSC0003DMP400045625	Gluthatione-peroxydase
7305.85	6759.21	8952.04	5940.89	6003.07	4274.46	1393.28	7148.44	5896.42	PGSC0003DMP400047519	SEX4-like
130.495	216.422	231.959	124.666	97.5203	130.676	21.9726	87.4788	117.02	PGSC0003DMP400049372	PHS1b
55640.6	56038	74653	62964.5	63044.5	30537.4	106529	95697.1	86739.3	PGSC0003DMP400050640	LSF2
5577.97	10034.8	6097.98	5908.62	4685.87	7309.51	5584.64	8161.26	3630.97	PGSC0003DMP400053308	PTST1
19515.9	17660.9	22300.9	18133.1	17014.2	23426.9	11768.4	19220.6	20201.1	Q43846	SS3
1810.52	618.114	822.083	821.907	1126.07	1213.44	499.55	1195.77	1238.03	Q84YG5	ISA3
12120.4	11037.2	9134.47	9200.37	10614	15110.7	5713.54	8208.72	12871.7	Sotub09g011090.1.1	BE2
16.8867	398.031	18.0797	11.8753	33.0303	56.6865	17.0136	596.359	33.233	SS4-gene--translation_SS4	SS4
288.889	165.904	125.955	56.9417	135.737	37.7159	0	406.899	302.3	SS6-gene---SS6-translation_SS6	SS6

## iBAQ N: Concentration [nM] échantillon fraction D

100-1-1	100-1-2	100-1-3	100-2-1	100-2-2	100-2-3	100-3-1	100-3-2	100-3-3	T: Protein IDs	T: Gene
693.155	793.874	789.089	980.918	1096.81	887.827	1218.45	1133.19	1943.3	P04045	PHS1a
17219.4	18218.9	17403.5	15838.4	15074	15768.1	8528.61	7396.67	9016.58	P30924	BE1.1
145811	152736	149001	144828	145119	134037	247056	288775	253232	PGSC0003DMP400002384	SS2
28979.5	15294.3	15246	12067.7	13789.4	12015.1	5260.63	5536.03	9031.08	PGSC0003DMP400009292	CYP20-2
13911.2	12149.7	13299.2	11591.2	12687.9	10701.9	7249.18	8873.42	12001.5	PGSC0003DMP400013565	GWD
1285.95	1362.17	1684.83	1002.93	454.777	529.186	1191.67	1061.17	318.576	PGSC0003DMP400017605	ESV1
16203.4	17117.1	16288.2	14825.2	14432.7	14936.2	8032.46	7135.78	8071.46	PGSC0003DMP400017627	BE3-BE1.2
595.972	322.57	275.412	76.4667	0	55.0773	135.528	180.994	668.17	PGSC0003DMP400017950	Cysteine-protease-inhibitor1
845.772	694.218	565.137	1109.61	1597.02	1113.03	2595.59	3603.51	323.167	PGSC0003DMP400017953	Kunitz-type-tuber-invertase-inhibitor
2630070	2626230	2637720	2641460	2647050	2673990	2480200	2456410	2451060	PGSC0003DMP400021398	GBSS
14417.7	17359.5	12459.1	14286	15726.1	11137.3	10514.1	19944.7	29969.8	PGSC0003DMP400021964	Thioredoxin
4686.29	4884.1	4822.52	3094.98	3085.04	3310.95	2488.42	2147.73	4428.5	PGSC0003DMP400026742	SEX4
19308.7	19978.3	18639.1	30068.4	22112	19801.1	70776	42538.5	72916.1	PGSC0003DMP400028510	LESV
399.975	347.109	364.708	362.8	246.298	275.713	220.003	137.442	373.746	PGSC0003DMP400029040	PWD
3036.17	3013.3	2531.15	2011.18	2124.13	1157.76	5450.52	5127.6	6471.79	PGSC0003DMP400032329	SS1
5740.9	7537.34	5974.71	1289.35	1434.46	2348.42	14930	17983.9	6401.2	PGSC0003DMP400045625	Gluthatione-peroxydase
5248.81	4117.99	4225.43	3479.29	3317.05	3018.52	1886.56	1914.54	3029.4	PGSC0003DMP400047519	SEX4-like
170.417	159.186	119.248	271.442	186.934	143.167	65.7516	69.545	144.396	PGSC0003DMP400049372	PHS1b
48383.5	48710.2	42652.8	41764.9	37101.8	36569.5	52766.1	91357.3	39192.5	PGSC0003DMP400050640	LSF2
4418.28	3614.64	3047	2973.59	3375.21	3019	8725.86	605.533	1018.68	PGSC0003DMP400053308	PTST1
8368.83	8497.05	8515.36	12057.1	11947	10629.4	13377.4	11173.2	17925.2	Q43846	SS3
2239.03	1386.1	1172.04	1061.55	879.431	785.403	1261.56	769.337	2210.33	Q84YG5	ISA3
7670.55	7967.89	7404.81	5239.18	6369.48	6093.41	12977.2	4443.39	10856.9	Sotub09g011090.1.1	BE2
0	642.543	673.803	1270.11	1362.51	1051.4	137.083	141.652	180.892	SS4-gene--translation_SS4	SS4
763.623	588.886	808.368	163.923	292.749	188.765	310.642	309.446	581.414	SS6-gene---SS6-translation_SS6	SS6

## Annexe 6

iBAQ score										
KO6-1	KO6-2	KO6-3	P-1	P-2	P-3	WT-1	WT-2	WT-3	T: Fasta headers	Name
58112000	51081000	54407000	1358100000	1557300000	1334500000	47331000000	51177000000	48629000000	PGSC0003DMP400021398	GBSS
4711700000	4315100000	4446200000	4163500000	4579400000	4878000000	5578500000	5583300000	5458300000	PGSC0003DMP400002384	SS2
444940000	483260000	462810000	541820000	708580000	726110000	634990000	630660000	619440000	P30924	BE1.1
408460000	445180000	438030000	621640000	755340000	750920000	596520000	627700000	574040000	PGSC0003DMP400017627	BE1.2
831390000	879140000	888540000	840340000	810550000	823730000	495110000	508500000	497780000	PGSC0003DMP400013565	GWD
291860000	304600000	289310000	240450000	277380000	290830000	366390000	341780000	356000000	PGSC0003DMP400028510	LESV
287650000	282780000	287150000	535600000	514390000	501830000	217800000	215010000	219590000	Q43846	SS3
143340000	169420000	106740000	171620000	146430000	187910000	201880000	187350000	159390000	Sotub09g011090.1.1	BE2
82116000	79180000	100680000	137250000	127010000	106220000	125090000	160290000	160820000	PGSC0003DMP400032328	SS1
154360000	190870000	137400000	87672000	124900000	148630000	106730000	78711000	100670000	PGSC0003DMP400050640	LSF2
52033000	46173000	42950000	326990000	315920000	322990000	68854000	76254000	77112000	PGSC0003DMP400017950	Cysteine protease inhibitor 1
96492000	117640000	120670000	36751000	37213000	33450000	50233000	64653000	51595000	PGSC0003DMP400014283	Elongation factor 1-alpha
106300000	89512000	100180000	67460000	63910000	70677000	41362000	43860000	43394000	PGSC0003DMP400009292	CYP20-2
39537000	38337000	39336000	35452000	35834000	34486000	29041000	34889000	23910000	Q84YG5	ISA3
82643000	88077000	83851000	70211000	62938000	65121000	25904000	27312000	28997000	PGSC0003DMP400017933	Serine protease inhibitor 7
41975000	33176000	34474000	48982000	40612000	42675000	16673000	18066000	18634000	P04045	PHS1a
105090000	104260000	116650000	67951000	52635000	89272000	16623000	17520000	16402000	PGSC0003DMP400021964	Thioredoxin
16412000	24234000	20492000	59130000	75840000	61420000	16346000	18033000	13917000	PGSC0003DMP400017953	Kunitz-type tuber invertase inhibitor
7824900	7137700	6455800	49994000	40074000	38431000	16233000	14816000	15777000	PGSC0003DMP400029040	PWD
35407000	41767000	40479000	20809000	25046000	27046000	16075000	18071000	13923000	PGSC0003DMP400026742	SEX4
47453000	50370000	49478000	31753000	27678000	32556000	12923000	14803000	14444000	PGSC0003DMP400047519	SEX4-Like
36947000	35975000	28631000	15882000	0	12654000	11444000	10615000	12774000	PGSC0003DMP400002952	Peptidyl-prolyl cis-trans isomerase
17207000	13890000	19512000	7783200	10068000	13078000	10353000	8928900	10191000	PGSC0003DMP400045625	Glutathione peroxidase
0	0	0	28075000	32379000	35805000	7388600	8891600	8278500	PGSC0003DMP400017447	Alpha-amylase
19300000	19062000	20093000	10297000	11122000	11026000	6693900	7246500	8380400	PGSC0003DMP400044905	Thioredoxin M3
492490000	562130000	547110000	0	0	0	4891500	4166200	3787600	PGSC0003DMP400052766	Actin
0	0	0	0	0	0	4595800	4804000	4065000	PGSC0003DMP400053308	PTST1
2627600	2604100	0	0	0	0	4202400	4858000	8074700	PGSC0003DMP400036451	Lipoxygenase
87659000	72278000	117880000	3688600	6931100	11135000	4106900	4443500	3170000	PGSC0003DMP400048253	Hsc70
0	0	0	0	0	0	3509900	4226800	3417000	PGSC0003DMP400014012	Annexin
2519600	2344500	2739200	5384200	5300600	5549200	3232500	3302000	3782300	PGSC0003DMP400017605	ESV1
13945000	10678000	16944000	2126900	1675200	2162400	2217200	1504100	1697500	PGSC0003DMP400032298	Ubiquitin fusion protein
36335000	33492000	33129000	0	0	0	1778200	0	0	PGSC0003DMP400068683	Histone H4
0	0	0	0	0	0	1356600	1037700	1063700	PGSC0003DMP400042657	101 kDa heat shock protein
1537700	1618000	1649300	0	0	0	1094600	826520	0	PGSC0003DMP400026560	Defensin P322
14709000	14522000	14694000	16820000	14552000	17276000	917420	487940	0	PGSC0003DMP400045760	Ferredoxin-3

## Annexe 7

NAME	Incorporation rate																							
	Beginning	Beginning	Beginning	Beginning	Beginning	Beginning	Middle	Middle	Middle	Middle	Middle	Middle	End	End	End	End	End	End	End	End	End	End	End	
GBSS	41%	44%	41%	40%	45%	42%	24%	24%	24%	27%	27%	27%	42%	41%	42%	22%	19%	21%	48%	45%	46%	36%	34%	34%
SS2	40%	27%	40%	40%	38%	41%	25%	25%	32%	27%	27%	27%	45%	44%	45%	29%	34%	22%	53%	50%	49%	38%	34%	37%
GWD	44%	43%	35%	45%	44%	34%	24%	25%	22%	30%	28%	21%	46%	44%	44%	25%	19%	31%	49%	49%	44%	31%	32%	40%
SS3	43%	44%	42%	45%	46%	46%	25%	26%	27%	25%	26%	27%	41%	43%	39%	12%	17%	23%	45%	38%	47%	28%	20%	31%
LESV	40%	41%	41%	26%	44%	40%	22%	27%	29%	25%	28%	28%	41%	46%	42%	33%	12%	12%	47%	49%	48%	35%	34%	35%
SS1	40%	41%	36%	42%	41%	34%	23%	23%	25%	26%	31%	26%	42%	37%	41%	14%	14%	26%	42%	40%	48%	30%	37%	36%
ISA3	42%	49%	37%	42%	53%	77%	18%	28%	68%	17%	24%	58%	16%	51%	47%	21%	20%	12%			46%	38%	37%	27%
BE1	51%	51%	49%	52%	50%	50%	32%	31%	30%	33%	34%	32%	48%	49%	48%	15%	15%	14%	49%	36%	46%	37%	32%	34%
BE2	49%	47%	42%	53%	40%	49%	32%	24%	27%	34%	28%	30%	48%	42%	37%	2%	1%	4%	50%	50%	48%	30%	32%	32%
PHS1a	35%	38%	40%	42%	38%	43%	25%	25%	21%	25%	25%	23%	43%	36%	38%	32%	36%	20%	49%	41%	39%	32%	24%	23%
SEX4	44%	38%	39%	49%	37%	42%	22%	29%	44%	29%	26%	22%	44%	36%	39%	5%	6%	10%						
SEX4-like	42%	44%	42%	37%	46%	43%	26%	27%	24%	26%	27%	25%	40%	39%	37%	50%	26%	0%	46%			31%		
LSF2	41%	40%	38%	40%	43%	39%	21%	25%	19%	25%	26%	22%	42%	43%	32%	16%	9%	18%	40%	48%	43%	33%	37%	33%
CYP20-2	46%	44%	40%	45%	45%	42%	20%	26%		27%	27%	19%	43%	51%	37%	10%	6%	-2%						
PTST1	43%	35%	-1%	59%	28%	1%	26%	15%	17%	30%	21%	5%	42%	35%			19%							

test-student: Beginning/middle	test-student:Middle/End	test-student Beginning/Middle	
GBSS	0.0%	0.23%	2.61%
SS2	0.1%	0.03%	27.15%
GWD	0.0%	0.06%	19.63%
SS3	0.0%	5.78%	0.22%
LESV	0.2%	1.24%	29.16%
SS1	0.0%	1.44%	7.31%
ISA3	10.2%	34.98%	1.36%
BE1	0.0%	20.44%	0.14%
BE2	0.0%	36.54%	1.02%
PHS1a	0.0%	0.08%	5.19%
SEX4	0.6%	27.46%	3.01%
SEX4-like	0.0%	9.04%	8.11%
LSF2	0.0%	1.09%	3.41%
CYP20-2	0.0%	48.65%	4.19%
PTST1	21.9%	8.89%	35.29%

significatif à p<0.1%

significatif à p<1%

significatif à p<5%

test-sudent protein/protein

Begining	
GBSS-SS2	8%
SS2-GWD	20%
GWD-SS3	8%
SS3-LESV	5%
LESV-SS1	45%
SS1-ISA3	8%
ISA3-BE1	46%
BE1-BE2	3%
BE2-PHS1a	1%
PHS1a-SEX4	13%
SEX4-SEX4-like	39%
SEX4-like-LSF2	6%
LSF2-CYP-2O-2	0%
CYP20-2-PTST1	6%
GBSS-GWD	27%
SS2-SS3	2%
GWD-LESV	29%
SS3-SS1	1%
LESV-ISA3	5%
SS1-BE1	0%
ISA3-BE2	28%
BE1-PHS1a	0%
BE2-SEX4	0%
PHS1a-SEX4-like	10%
SEX4-LSF2	29%
SEX4-like/cyp20-2	22%
LSF2-PTST1	12%
GBSS-SS3	3%
SS2-LESV	41%
GWD-SS1	3%
SS3-ISA3	16%
LESV-BE1	1%
SS1-BE2	1%
ISA3-PHS1a	5%
BE1-SEX4	0%
BE2-SEX4-like	12%
PHS1a-LSF2	28%

test-sudent protein/protein

Middle	
GBSS-SS2	13%
SS2-GWD	18%
GWD-SS3	30%
SS3-LESV	27%
LESV-SS1	27%
SS1-ISA3	17%
ISA3-BE1	36%
BE1-BE2	5%
BE2-PHS1a	1%
PHS1a-SEX4	15%
SEX4-SEX4-like	23%
SEX4-like-LSF2	1%
LSF2-CYP-2O-2	49%
CYP20-2-PTST1	15%
GBSS-GWD	40%
SS2-SS3	15%
GWD-LESV	25%
SS3-SS1	39%
LESV-ISA3	16%
SS1-BE1	0%
ISA3-BE2	27%
BE1-PHS1a	0%
BE2-SEX4	43%
PHS1a-SEX4-like	0%
SEX4-LSF2	12%
SEX4-like/cyp20-2	13%
LSF2-PTST1	15%
GBSS-SS3	21%
SS2-LESV	32%
GWD-SS1	34%
SS3-ISA3	16%
LESV-BE1	1%
SS1-BE2	5%
ISA3-PHS1a	14%
BE1-SEX4	22%
BE2-SEX4-like	3%
PHS1a-LSF2	11%

test-sudent protein/protein

End	
GBSS-SS2	0%
SS2-GWD	12%
GWD-SS3	0%
SS3-LESV	5%
LESV-SS1	17%
SS1-ISA3	39%
ISA3-BE1	25%
BE1-BE2	5%
BE2-PHS1a	25%
PHS1a-SEX4	6%
SEX4-SEX4-like	18%
SEX4-like-LSF2	22%
LSF2-CYP-2O-2	28%
CYP20-2-PTST1	44%
GBSS-GWD	5%
SS2-SS3	0%
GWD-LESV	20%
SS3-SS1	14%
LESV-ISA3	23%
SS1-BE1	24%
ISA3-BE2	21%
BE1-PHS1a	39%
BE2-SEX4	28%
PHS1a-SEX4-like	31%
SEX4-LSF2	14%
SEX4-like/cyp20-	18%
LSF2-PTST1	46%
GBSS-SS3	1%
SS2-LESV	4%
GWD-SS1	1%
SS3-ISA3	36%
LESV-BE1	35%
SS1-BE2	19%
ISA3-PHS1a	42%
BE1-SEX4	0%
BE2-SEX4-like	17%
PHS1a-LSF2	33%

SEX4-CYP20-2	12%
SEX4-like-PTST1	11%
GBSS-LESV	9%
SS2-SS1	34%
GWD-ISA3	12%
SS3-BE1	0%
LESV-BE2	7%
SS1-PHS1a	48%
ISA3-SEX4	11%
BE1-SEX4-like	0%
BE2-LSF2	2%
PHS1a-CYP20.2	3%
SEX4-PTST1	9%
GBSS-SS1	3%
SS2-ISA3	5%
GWD-BE1	0%
SS3-BE2	16%
LESV-PHS1a	44%
SS1-SEX4	16%
ISA3-SEX4-like	11%
BE1-LSF2	0%
BE2-CYP20.2	11%
PHS1a-PTST1	15%
GBSS-ISA3	11%
SS2-BE1	0%
GWD-BE2	4%
SS3-PHS1a	0%
LESV-SEX4	28%
SS1-SEX4-like	8%
ISA3-LSF2	7%
BE1-CYP20.2	0%
BE2-PTST1	4%
GBSS-BE1	0%
SS2-BE2	1%
GWD-PHS1a	28%
SS3-SEX4	10%
LESV-SEX4-like	3%
SS1-LSF2	17%
ISA3-CYP20.2	17%

SEX4-CYP20-2	6%
SEX4-like-PTST1	6%
GBSS-LESV	19%
SS2-SS1	19%
GWD-ISA3	18%
SS3-BE1	0%
LESV-BE2	14%
SS1-PHS1a	15%
ISA3-SEX4	20%
BE1-SEX4-like	0%
BE2-LSF2	1%
PHS1a-CYP20.2	31%
SEX4-PTST1	5%
GBSS-SS1	40%
SS2-ISA3	18%
GWD-BE1	0%
SS3-BE2	5%
LESV-PHS1a	9%
SS1-SEX4	24%
ISA3-SEX4-like	17%
BE1-LSF2	0%
BE2-CYP20.2	5%
PHS1a-PTST1	10%
GBSS-ISA3	16%
SS2-BE1	1%
GWD-BE2	2%
SS3-PHS1a	5%
LESV-SEX4	27%
SS1-SEX4-like	46%
ISA3-LSF2	13%
BE1-CYP20.2	0%
BE2-PTST1	1%
GBSS-BE1	0%
SS2-BE2	15%
GWD-PHS1a	17%
SS3-SEX4	24%
LESV-SEX4-like	24%
SS1-LSF2	4%
ISA3-CYP20.2	29%

SEX4-CYP20-2	42%
SEX4-like-PTST1	22%
GBSS-LESV	41%
SS2-SS1	1%
GWD-ISA3	17%
SS3-BE1	4%
LESV-BE2	5%
SS1-PHS1a	44%
ISA3-SEX4	27%
BE1-SEX4-like	45%
BE2-LSF2	24%
PHS1a-CYP20.2	11%
SEX4-PTST1	28%
GBSS-SS1	7%
SS2-ISA3	5%
GWD-BE1	14%
SS3-BE2	41%
LESV-PHS1a	28%
SS1-SEX4	5%
ISA3-SEX4-like	36%
BE1-LSF2	13%
BE2-CYP20.2	25%
PHS1a-PTST1	18%
GBSS-ISA3	25%
SS2-BE1	3%
GWD-BE2	3%
SS3-PHS1a	18%
LESV-SEX4	7%
SS1-SEX4-like	32%
ISA3-LSF2	41%
BE1-CYP20.2	2%
BE2-PTST1	42%
GBSS-BE1	39%
SS2-BE2	1%
GWD-PHS1a	10%
SS3-SEX4	4%
LESV-SEX4-like	47%
SS1-LSF2	26%
ISA3-CYP20.2	29%

BE1-PTST1	3%
GBSS-BE2	7%
SS2-PHS1a	26%
GWD-SEX4	42%
SS3-SEX4-like	11%
LESV-LSF2	29%
SS1-CYP20.2	0%
ISA3-PTST1	7%
GBSS-PHS1a	6%
SS2-SEX4	7%
GWD-SEX4-like	30%
SS3-LSF2	0%
LESV-CYP20.2	8%
SS1-PTST1	12%
GBSS-SEX4	38%
SS2-SEX4-like	9%
GWD-LSF2	33%
SS3-CYP20.2	32%
LESV-PTST1	19%
GBSS-SEX4-like	44%
SS2-LSF2	18%
GWD-CYP20.2	3%
SS3-PTST1	7%
GBSS-LSF2	1%
SS2-CYP20.2	3%
GWD-PTST1	7%
GBSS-CYP20.2	10%
SS2-PTST1	18%
GBSS-PTST1	10%

BE1-PTST1	1%
GBSS-BE2	1%
SS2-PHS1a	7%
GWD-SEX4	21%
SS3-SEX4-like	27%
LESV-LSF2	3%
SS1-CYP20.2	18%
ISA3-PTST1	11%
GBSS-PHS1a	8%
SS2-SEX4	30%
GWD-SEX4-like	36%
SS3-LSF2	3%
LESV-CYP20.2	15%
SS1-PTST1	6%
GBSS-SEX4	22%
SS2-SEX4-like	17%
GWD-LSF2	4%
SS3-CYP20.2	18%
LESV-PTST1	7%
GBSS-SEX4-like	46%
SS2-LSF2	5%
GWD-CYP20.2	5%
SS3-PTST1	6%
GBSS-LSF2	3%
SS2-CYP20.2	14%
GWD-PTST1	4%
GBSS-CYP20.2	19%
SS2-PTST1	4%
GBSS-PTST1	7%

BE1-PTST1	22%
GBSS-BE2	6%
SS2-PHS1a	0%
GWD-SEX4	1%
SS3-SEX4-like	33%
LESV-LSF2	4%
SS1-CYP20.2	22%
ISA3-PTST1	41%
GBSS-PHS1a	28%
SS2-SEX4	2%
GWD-SEX4-like	34%
SS3-LSF2	33%
LESV-CYP20.2	9%
SS1-PTST1	32%
GBSS-SEX4	2%
SS2-SEX4-like	14%
GWD-LSF2	1%
SS3-CYP20.2	18%
LESV-PTST1	45%
GBSS-SEX4-like	49%
SS2-LSF2	0%
GWD-CYP20.2	6%
SS3-PTST1	36%
GBSS-LSF2	3%
SS2-CYP20.2	4%
GWD-PTST1	15%
GBSS-CYP20.2	10%
SS2-PTST1	6%
GBSS-PTST1	22%