École Doctorale Biologie Santé de Lille

THÈSE DE DOCTORAT

Pour l'obtention du grade de
DOCTEUR DE L'UNIVERSITE DE LILLE

Discipline
Aspects moléculaires et cellulaires de la biologie

Structural and Functional Study of ESV1 and LESV Proteins from *Arabidopsis thaliana* reveals a Novel Starch-Binding Module

L'Étude structurale et fonctionnelle des protéines ESV1 et LESV d'*Arabidopsis thaliana* révèle un nouveau module de liaison à l'amidon

Présentée et soutenue publiquement par

Rayan OSMAN

Le 7 Juillet 2023

Directrice de thèse : Dr. Coralie BOMPARD

Membres du jury

Président du Jury :	
M. Jean-Louis HILBERT	PU, INRAE UMR 1158, Université de Lille, Villeneuve d'Ascq
Rapportrices :	
Mme Mirjam CZJZEK	DR, CNRS, LBI2M, UMR 8227, Université Sorbonne, Roscoff
Mme Véronique RECEVEUR-BRECHOT	DR, CNRS, BIP UMR7281 AMU, Université de Marseille
Examinateurs :	
M. Christophe D'HULST	PU, CNRS, UMR 8576, Université de Lille, Villeneuve d'Ascq
M. Jérôme PELLOUX	PU, INRAE UMR 1158, Université de Picardie Jules, Amiens
Directrice : Mme Coralie Bompard	CR, CNRS UMR 8576, Université de Lille, Villeneuve d'Ascq

Acknowledgements

Two and a half years of discovery and immersion in the world of proteins, of meeting new people, of experimentation and happiness. My Ph.D. journey was not easy, but it was very educational and helped me to improve my knowledge in the fields of biochemistry, biophysics and shaped me personally. I would like to thank every person in the lab or in my personal life who directly or indirectly contributed to the success of this project.

There are no words that can express what I feel and what I want to say. Words are never enough to thank my supervisor, Dr. Coralie Bompard, for her guidance, her support, her trust, her availability at all times and her care for me during all these years, it would not have been possible without her. She gave me the opportunity to realize one of my dreams. Our discussions and her help with any question I had helped me develop my scientific, management, and critical thinking skills throughout this thesis. I am also grateful for her corrections, suggestions, and discussions on all my writing attempts and especially on this thesis. She is more than a supervisor to me personally. I will never forget such an amazing person in my life.

I would like to sincerely thank Dr. Mirjam Czjzek and Dr. Jean-Philippe Ral, who were members of the jury of the Individual Monitoring Committee (CSI), and whose questions and suggestions at each of our annual meetings give me more ideas and support to continue this project.

I am also grateful to Dr. Yann Guerardel, Director of the Structural and Functional Glycobiology Unit, who gave me the opportunity to work in the UGSF laboratory. My respectful acknowledgement goes to the CNRS who funded this research and supported me during these years of my research.

I cannot forget my dear friend Mélanie Bossu for the quality time we spent together and for her support whenever I needed it, I wish you all the best and opportunity in your Ph.D. defense next year and in your personal life. I would also like to thank all the members of the PlaStoPol team, Christophe D'hulst for having welcomed me in the team, to Adeline, Xavier, Océane, Camille, Fabrice, Nicolas, Claudine who have created a pleasant working environment, and especially Dr. David Dauville, for his numerous help with many experiments when I needed it most, his suggestions and help are always very helpful and kind. Thanks to Dr. Julie Bouckaert for her help with the ITC and SPR experiments. Thanks to my dear friend Shubham Semwal for his help with the ITC and SPR experiments, as well as for his daily support, our late tea breaks, for

supporting each other and staying late together in the lab to finish our experiments. A big thank you to all the members of the UGSF, the administrative team for always being present in my paprasseresque steps. Apart from the members of the laboratory, I would like to thank the people at DISCO, Swing, Proxima at the Soleil synchrotron who helped me during my experiments there, thank you for your help. My thanks also to all the Ph.D. students in the laboratory.

I would like to thank Dr. Mirjam Czjzek and Dr. Véronique Receveur-Brechot for agreeing to review my thesis. Thanks also to the other members of the jury, Professor Christophe D'hulst, Professor Jean-Louis Hilbert and Professor Jérôme Pelloux.

I would like to sincerely thank Dr. Henri Happy of the IEMN laboratory who helped me with his advice, his encouragement to be where I am today. Thank you, Dr. Jean-Luc Mession, for your help, as well as Mrs. Nora Benbahlouli and Mr. Andy Ledent. A big thank you to Mr. Maxime Flamant, in charge of human resources at the CNRS, who has always been available to listen to me and to be where I am today.

I would like to thank the first person I discovered biochemistry with him, Dr. Roland Bouraad the person who support me during my bachelor degree. Also, a big thank you to Dr. Jamilah Borjac who encouraged me a lot during my master's degree.

Finally, I would like to express my heartfelt gratitude to my family who have supported me emotionally and morally throughout the years of my studies. My mother, Oumayda, my father, Mohamad, whose parental guidance, encouragement and financial support to ensure that I have this wonderful education throughout my school years to this moment. Thank you for always being there for me. Your willingness and desire to see me through has been a challenging factor that draws inspiration to focus on my goals and career objectives. I hope I have made you proud. Thank you to my sister Souraya for always being there and to my brothers Abdulkader, Hassan and Hussein. Thank you Laurice for your care and support during the writing period. I dedicate my success to my family and the soul of my dear grandmother.

أمي وأبي، أحبكم كثيرًا بحجم هذه الدنيا وأكثر ♡ أود أن أشكركم على كل ما فعلتموه وقدمتموه لي، وأتمنى أن تكونون فخورين بي بكل ما وصلت إليه حتى الآن. أود أن أهدي نجاحي لكم، وأسأل الله أن تبقوا دائمًا بصحة جيدة وفوق رؤوسنا ♥ الله يحميكم، أحبكم بحجم الدنيا وأكثر ♥ أهدي نجاحي لروح جدتي الغالية جدًا على قلبي، أم خزعل حبيبة قلبي♥

I would like to express my sincere appreciation to my old overseas friends Rahme Osman, Sireen Miari, Rawan Kawach and Abbas Mansour for their friendship and many happy times.

I will always remember the precious time I spent with you. Also, people I met in Lille and whom I consider close friends, Baker Shalak, Abdelghafour Sid, Ali Mourad, Warda Raiah and Rayane Hijazi, who supported me daily with their precious friendship. I could not have done it without them. Thanks for my support system Ihab, I wish you all the success in your trading life.

To my friends in Nancy, Ziad Hassan, the friend of beautiful travels, Sarah El Hajj, Racha Zgheib, Rami Bechara, Youssef Ali, Rawan Kawach, Abbas Mansour, Christelle Ghazaly, Elio El Kahi, Abdelsamad Choukri, Bipasha Barua and many others, thank you for your generosity that cannot be found elsewhere, thank you for the unforgettable memories.

Thank you all,

Rayan Osman

Abstract

Starch metabolism is a critical process for plant growth and development, regulated by various enzymatic and non-enzymatic machinery. Starch accumulates in plants as semi-crystalline granules composed of two glucan polymers: amylose and amylopectin. The physicochemical properties of starch depend on the organization of these polymers within the granule. It has been widely assumed that the amylopectin chains assembled into double helices are able to self-organize by physical processes to form the starch granule matrix. Recent studies have identified two conserved non-enzymatic proteins, Early Starvation 1 (ESV1) and Like Early Starvation 1 (LESV), which have been proposed to be involved in the organization of the polysaccharides within the granule.

Forward genetics studies of these proteins highlight their importance in influencing the structure and organization of starch granules as well as starch metabolism in plants. The mechanisms by which ESV1 and LESV function are not yet understood, but they appear to have different roles. ESV1 may promote a high level of organization of glucans in the granule matrix, whereas LESV may reduce this level of organization, as its overexpression in plants increases the susceptibility of starch to degradation during the day. In *Arabidopsis*, ESV1 (49 kDa) and LESV (66 kDa) are proteins with no previously annotated regions, such as catalytic domains or carbohydrate-binding domains. On the other hand, although their N-terminal regions differ, they share a common C-terminal tryptophan-rich domain containing conserved motifs, which has been proposed to constitute a novel and specific starch-binding module.

In this study, we have carried out a structural characterization of both proteins using model modelling and small-angle X-ray scattering (SAXS). Our results indicate that the conserved C-terminal domain adopts an unusual β-sheet fold in which conserved residues form parallel lines whose positions are consistent with interactions with starch molecules. We have shown that ESV1 and LESV interact differently and have different affinities for starch polymers, explaining their distinct functions.

Keywords: *Starch metabolism, Early Starvation 1 (ESV1), Like Early Starvation 1 (LESV), amylopectin, glucans, Protein biochemistry, Structural biology*.

Résumé

Le métabolisme de l'amidon est un processus critique pour la croissance et le développement des plantes, régulé par divers mécanismes enzymatiques et non enzymatiques. L'amidon s'accumule dans les plantes sous forme de grains semi-cristallins composés de deux polymères de glucane : l'amylose et l'amylopectine. Les propriétés physicochimiques de l'amidon dépendent de l'organisation de ces polymères à l'intérieur des grains. On suppose généralement que les chaînes d'amylopectine assemblées en doubles hélices sont capables de s'auto-organiser par des processus physiques pour former la matrice du grain d'amidon. Des études récentes ont identifié deux protéines non enzymatiques conservées, Early Starvation 1 (ESV1) et Like Early Starvation 1 (LESV), dont on a proposé qu'elles soient impliquées dans l'organisation des polysaccharides à l'intérieur du grain.

Les études de génétique inverse de ces protéines soulignent leur importance dans l'influence de la structure et de l'organisation des grains ainsi que du métabolisme de l'amidon chez les plantes. Les mécanismes par lesquels ESV1 et LESV fonctionnent ne sont pas encore compris, mais ils semblent avoir des rôles différents. ESV1 pourrait favoriser un niveau élevé d'organisation des glucanes dans la matrice des granules, tandis que LESV pourrait réduire ce niveau d'organisation, car sa surexpression dans les plantes augmente la susceptibilité de l'amidon à la dégradation au cours de la journée. Chez *Arabidopsis*, ESV1 (49 kDa) et LESV (66 kDa) sont des protéines dépourvues de régions précédemment annotées, telles que des domaines catalytiques ou des domaines de liaison aux hydrates de carbone. D'autre part, bien que leurs régions N-terminales diffèrent, elles partagent un domaine C-terminal riche en tryptophane contenant des motifs conservés, qui a été proposé pour constituer un nouveau module spécifique de liaison à l'amidon.

Dans cette étude, nous avons procédé à la caractérisation structurelle des deux protéines en utilisant la modélisation et la diffusion des rayons X aux petits angles (SAXS). Nos résultats indiquent que le domaine C-terminal conservé adopte un pliage β-sheet inhabituel dans lequel les résidus conservés forment des lignes parallèles dont les positions sont cohérentes avec les interactions avec les molécules d'amidon. Nous avons montré que ESV1 et LESV interagissent différemment et ont des affinités différentes pour les polymères d'amidon, expliquant leurs fonctions distinctes au sein de la plante.

Mots clés : *métabolisme de l'amidon, Early Starvation 1 (ESV1), Like Early Starvation 1 (LESV), amylopectine, glucanes, Biochimie des protéines, Biologie structurale.*

Table of Content

List of Figures

Figure 5.18: SR-CD spectra for the proteins secondary structure content of LESV alone (violet color) and ESV1 alone (blue color).

Figure 5.19: Superposition of the *SR-CD* spectra of LESV alone (Pink) and when mixed with amylopectin (violet) or amylose (blue) solutions.

Figure 5.20: shows the superposition of the *SR-CD* spectra of ESV1 alone (pink) and when mixed with amylopectin (violet) or amylose (blue) solutions.

Figure 5.21: Thermal denaturation of LESV followed by SR-CD. Each plot represents consecutive scans on the protein collected at a set of temperature between 20 to 90°.

Figure 5.22: Thermal denaturation of ESV1 followed by SR-CD. Each plot represents consecutive scans on the protein collected at a set of temperature between 20 to 90°.

Figure 5.23: Starch granules seen under fluorescence microscopy.

Figure 5.24: The molecular model shows the complex formed between the C-terminal domain of LESV and the double helices of amylopectin.

Figure 6: ESV1 and LESV proposed mechanism

Figure S1: Sequence alignment produced by ClustalO of ESV1 and LESV protein sequences from different organisms.

Figure S2: Analysis of ESV1 and LESV binding to A600 by EMSA gel with a concentration of 0%, 0.1%, 0.2% and 0.3% of potential ligands.

Figure S3: Analysis of ESV1 and LESV binding to pullulan by EMSA gel with a concentration of 0%, 0.1%, 0.2% and 0.3% of potential ligands.

Figure S4: Structural conserved motifs on the Face A of LESV (top) and ESV1 (bottom).

List of Tables

List of Abbreviations

At: *Arabidopsis thaliana*

AA: Amino acid

ADP-glucose: Adenosine 5'-diphosphate-glucose

AGPase: ADPglucose pyrophosphorylase

CD: Circular Dichroism

CSA: Ammonium *d*-10-Camphorsulfonate

CAZy: Carbohydrate-Active enZYmes

CERMAV: Research Center for Plant Macromolecules

Col-0: Columbia genetic background in *Arabidopsis thaliana*

CBM: Carbohydrate-binding modules

cTP: chloroplastidial transit peptide

DPE: Disproportionating enzyme

DP: Degree of polymerization

DTT: Dithiothreitol

DBE: Starch debranching enzymes

DUF: Domain uncharacterized and functional

ESV1: EARLY STARVATION 1

EMSA: Electrophoretic Mobility Shift Assay

EOD: End of the day

EON: End of the night

GBSS: Granule bound starch synthase

GBSSI: Granule bound starch synthase I

GBSSII: Granule bound starch synthase II

Glc: Glucose

Glc-6-P: Glucose-6-phosphate

GWD: Glucan water dikinase

ISA1: Isoamylase I

ISA2: Isoamylase II

ISA3: Isoamylase II

kDa: kilo Dalton

LESV: LIKE EARLY STARVATION 1

LDA: Limit dextrinase

LC: Liquid chromatography

LSF2: Starch-excess Four2

MOS: Maltooligosaccharides

MRC: MYOSIN-RESEMBLING CHLOROPLAST PROTEIN

PU1: Pullulanase

PTST: Protein Targeting to Starch

PII1: PROTEIN INVOLVED IN STARCH INITIATION1

PWD: Phosphoglucan water dikinase

Pi: Inorganic phosphate

SAXS: Small angle X-ray scattering

SSI: Starch synthase I

SSII: Starch synthase II

SSIII: Starch synthase III

SSIV: Starch synthase IV

SSV: Starch synthase V

SS: Starch synthase

SDS: Sodium dodecyl sulfate

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

SEX: Starch EXcess

Tm: Melting temperature

Publications

Published in Science advances Journal

Chun Liu, Barbara Pfister, **Rayan Osman**, Maximilian Ritter, Arvid Heutinck, Mayank Sharma, Simona Eicke, Michaela Fischer-Stettler, David Seung, Coralie Bompard, Melanie R. Abt, Samuel C. Zeeman. "LIKE EARLY STARVATION 1 and EARLY STARVATION 1 Promote and Stabilize Amylopectin Phase Transition in Starch Biosynthesis". Science Advances Edition. 2023 May 26.

Submitted:

The work in this study about ESV1 and LESV protein is included in a journal paper which is submitted.

Rayan Osman, Mélanie Bossu, David Dauvillée, Corentin Spriet, Chun Liu, Samuel Zeeman, Christophe D'Hulst and Coralie Bompard. "Structural basis of the interaction between ESV1 and LESV from *Arabidopsis thaliana* with starch glucans".

Communications

Poster presentations:

Rayan OSMAN, Chun Liu, Nicolas Szydlowski, David Dauvillée, Sam Zeeman, Christophe d'Hulst & Coralie Bompard. "Structural and Functional study of ESV1 and LESV from *Arabidopsis thaliana* reveals a new fold designed for interaction with starch molecules". Poster presented during the International conference on the glycoproteins (the 8th Symposium on the Alpha-Amylase Family - ALAMY_8) from 9 till 13 of October 2022 in the Smolenice Castle, Slovakia.

R. OSMAN, C. Liu, N. Szydlowski, F. Wattebled, X. Roussel, D. Dauvillée, S. Zeeman, C. d'Hulst & C. Bompard. "Structural and Functional study of ESV1 and LESV from *Arabidopsis thaliana* & *Solanum tuberosum*, two new proteins involved in plant starch metabolism". Poster presented during the International workshop on Structural Biophysics from 5 till 10 of December 2021 in the city of Bordeaux, France.

Oral Presentation:

"Structural and Functional analysis of ESV1 and LESV, two new proteins involved in plant starch metabolism", Presented during the 20ème Journée André Verbert 2021 (Online seminar).

Foreword

When I began my academic journey, I initially enrolled in the Business Management program. However, I soon realized that my true passion lay elsewhere: in the complex interaction of the smallest elements that make up life. As a result, I switched my major to Biochemistry in the middle of my second year, a decision that led me down a fascinating and challenging path.

After obtaining my bachelor's degree in biochemistry in Lebanon, I pursued a master's degree where I had the chance to study the impact of environmental pollutants on the health of people living near a landfill. This experience opened my eyes to the real impact of biochemistry as it applies to human health and the environment.

In September 2017, I moved to France to complete two master's degrees at the University of Lorraine. My first master 2 was an international RNA/enzyme science degree, focusing my research on the neuropathological consequences of methionine synthase deficiency in aging. In the second, I specialized in microbiology, developing diagnostic tools for the identification of Mayaro virus and its differentiation from other viruses.

I joined the UGSF laboratory on December 2020 for my second year of Ph.D. on the project proposed by Mme Coralie Bompard entitled "Structural and Functional Study of ESV1 and LESV Proteins from *Arabidopsis thaliana* reveals a Novel Starch-Binding fold Module". This project felt like a perfect fit for me, and I was drawn to the complex mystery of it - sometimes it feels like we don't choose our destiny, but rather they choose us.

My research work focused on the characterization at the molecular level the structure and function of two new, highly-conserved non-enzymatic proteins namely EARLY STARVATION 1 (ESV1) and LIKE EARLY STARVATION 1 (LESV). These proteins have been identified to be involved in the control of starch metabolism in *Arabidopsis thaliana (At)* plant model, but their molecular functions in the plant remain unclear.

My research aimed to elucidate the mysteries of the structure and function of the two proteins ESV1 and LESV using integrative structural biology approaches. The first part consists in producing and purifying the proteins: Early Starvation 1 (ESV1) and Like Early Starvation 1 (LESV), and then validate their AlphaFold predicted patterns by checking for model's

compatibility with the SAXS and Circular Dichroism data, confirming the presence of the predicted β-sheet. The second part consists in examining protein-ligand interactions, by analyzing the different putative ligands using different techniques that allow us to study this interaction and to check for structural rearrangements such as electrophoretic mobility shift assay and circular dichroism, to find the appropriate ligands for crystallographic determination of 3D protein structures.

I am grateful to have had the opportunity to work on this project in collaboration with Prof. Sam Zeeman's team at ETH Zurich, who originally discovered these proteins. The guidance and support of my supervisor, Coralie Bompard were invaluable.

This journey has been both challenging and rewarding, and I hope that the insights provided in this dissertation will help us better understand these unique proteins and their role in plant metabolism."

1. Introduction

1.1. Carbohydrates

Biomolecules are essential components to all living organisms and are divided into four main classes: carbohydrates, proteins, lipids, and nucleic acids. Carbohydrates play a critical role in the proper functioning of all living organisms, serving as a primary source of energy and carbon. Chemically, carbohydrates are a diverse group of organic compounds primarily composed of carbon, hydrogen, and oxygen atoms, with a general molecular formula of Cx(H2O)y.

Carbohydrates can be classified based on their complexity, size, and functional groups. Monosaccharides, or simple carbohydrates, are the fundamental building blocks of more complex carbohydrates, including glucose, fructose, and galactose. Disaccharides, such as sucrose, lactose, and maltose, are formed when two monosaccharide units are linked together through a glycosidic bond. Oligosaccharides and polysaccharides are examples of complex carbohydrates, consisting of a few or long chains of monosaccharides, respectively. Complex carbohydrates can be further classified into four types: starch and glycogen, which serve as energy storage molecules in plants and animals, and cellulose and chitin, which function as structural components in plants and arthropods, respectively. However, the structure of glucans can vary depending on the specific type of glucan from starch, glycogen, cellulose or chitin. Also, the specific arrangements of the glucose monomers can result in different properties and functions of the different types of glucans listed above. Starch is the most abundant and widespread non-structural carbohydrate in plants.

Starch and glycogen, both branched homoglycans composed of α-D-glucosyl residues, share chemical similarities. They contain two types of glycosidic bonds: α-1,4-glycosidic bonds in the backbone and α-1,6-glycosidic bonds in the branches. Glycogen is characterized by short, extensively branched chains and a restricted solubility, much like amylopectin.

a. Glycogen: energy storage molecule in animals

Glycogen is a branched storage polysaccharide found in organisms ranging from bacteria and archaea to more complex organisms like yeast, animals, and humans (Adeva-Andany et al. 2016; S. G. Ball et Morell 2003; Wilson et al. 2010). It is composed of α-1,4-linked glucose

units and contains branching points with α-1,6-glycosidic linkages roughly every ten residues (Roach 2002) (Figure 1.1). The synthesis of glycogen polymers begins with the Glycogenin (GN) enzyme, working in tandem with glycogen synthase (GS) and glycogen branching enzyme (Zeqiraj et al. 2014; Hurley et al. 2006). The homodimeric GN enzyme possesses a tyrosine 194 (Tyr194) residue on each subunit and employs UDP-glucose as a donor for transferring glucosyl residues directly to Tyr194 or to glucose residues already connected to Tyr194 via an autoglycosylation process. This results in a polymer with a minimum of four glucose units linked by α-1,4-glycosidic bonds (Gibbons, Roach, et Hurley 2002; B. Lin et al. 1999).

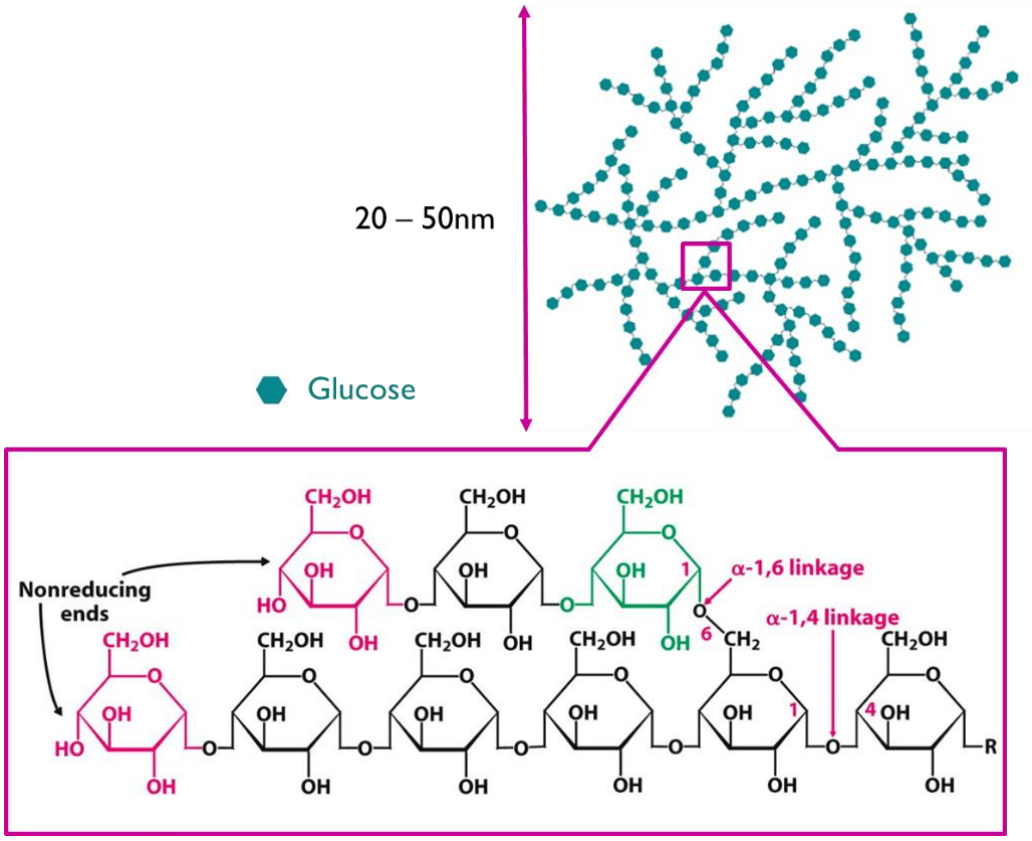


Figure 1.1: ***Structure of Glycogen.*** The linear configuration is a component of a glycogen molecule, where glucose units are interconnected by α-1,4-glycosidic bonds. At the branching points, they are connected by α-1,6-glycosidic bonds.

After an oligosaccharide chain of glucosyl monomers is established as a glycan primer, it is elongated by glycogen synthase, but only when complexed with glycogenin. During glycogen

molecule synthesis, glycogen synthase detaches from glycogenin, allowing glycogen molecules to grow and reach their full size (Smythe et Cohen 1991). The glycogen branching enzyme then creates branches by transferring the chain's end to form α-1,6-position branch points within the same or nearby glucan chain in glycogen (K. Wang, Henry, et Gilbert 2014).

Conversely, the breakdown of glycogen polymers is facilitated by the combined action of glycogen phosphorylase and glycogen debranching enzyme. Glycogen phosphorylase releases glucose 1-phosphate from linear glucan chains, but its activity is inhibited when it approaches the last four glucosyl residues from a glucan branch point. Thus, the glycogen debranching enzyme is needed to complete degradation and release branch points (Alonso-Casajús et al. 2006; Burwinkel et al. 1998; Adeva-Andany et al. 2016). In humans, glycogen is primarily stored in the liver, providing glucose to the bloodstream during fasting periods, and in skeletal muscle, supplying glucose to muscle fibers during muscle contractions. Additionally, glycogen is present in other human tissues, such as the brain, heart, kidney, adipose tissue, and erythrocytes (Roach et al. 2012).

b. Starch: energy storage molecule in plants

Starch is a type of carbohydrate that is primarily used by plants for storage purposes, and it serves as a key source of calories for both humans and animals. Additionally, this polymer is utilized in a wide range of industrial applications, including those related to food and non-food products (Samsudin et Hani 2017). Although its unique properties make it ideal for various uses, starch is typically processed after extraction from plants rather than being used in its native form.

Key factors, such as starch content, granule size, phosphate and protein contents, and glucan structure, significantly impact the extractability, yield, and end-use of starch. Additionally, these parameters influence the digestibility and nutritional properties that starch provides to humans. Consequently, understanding and optimizing these factors is essential to maximize the benefits of starch in both food and non-food applications (Y. Zhang, Rempel, et McLaren 2014).

Starch can be classified into two types - storage starch and transitory starch - depending on its biological functions. Typically referred as **transitory starch**, this type of starch is produced in the leaves (source organs) during the day through photosynthesis. It accumulates in the chloroplasts of photosynthetic cells during the day, and is broken down and used as carbon and

energy source during the night to sustain metabolism and biosynthesis in the absence of photosynthesis, as well as for energy production (Pfister et Zeeman 2016). This process is under a photoperiodic control guaranteeing progressive consumption of starch reserves until dawn (Scialdone et al. 2013). However, plants can experience severe starvation if their starch reserves are depleted before the end of the night. It happens in wild type plants when they are exposed to a night longer than the time it takes to use up their starch reserves (Gibon et al. 2006), or in mutants with defective starch synthesis (Pal et al. 2013). In addition, the same effect is also observed in a mutant with a shortened circadian clock, which depleted its starch reserves three hours before the end of the night (Graf et al. 2010). On the other hand, **Storage starch** is the type of starch found in plant organs that don't perform photosynthesis, such as seeds, roots, and tubers. It is typically stored for extended periods of time and serves as long-term carbohydrate reserve used only when photosynthesis is unable to provide all the energy and carbon needed for biosynthesis which is used during fuel germination, and seasonal re-regrowth (Pfister et Zeeman 2016). This type of starch, produced and stored in large amounts, is commonly consumed in our daily life and used by the industries (Lafiandra, Riccardi, et Shewry 2014). Starches obtained from different botanical sources have different characteristics that affect their properties and possible applications (K. Wang, Henry, et Gilbert 2014). The variations arise from differences in the structure of starch. For example, the size of starch granules can vary depending on the plant source, which affects how quickly they gelatinize. The composition of the starch granules, such as the relative amounts of amylose and amylopectin, also influences their functional properties. Finally, the molecular architecture of the constituent polymers, which make up the starch granules, can vary between different plant sources, resulting in different functional properties (K. Wang, Henry, et Gilbert 2014); (Pfister et Zeeman 2016).

1.2. Regulation of Transitory Starch Biosynthesis and Degradation in *Arabidopsis thaliana* Leaves: The Role of Circadian Clock and Carbon Allocation

Plants accumulate starch as a product of photosynthesis in their leaves during the day and use it as a carbon source for normal growth and to provide energy at night for the plant's metabolism (Stitt et Zeeman 2012). Approximately 50% of the carbon assimilated by photosynthesis in plants is stored as starch in leaf chloroplasts (Sulpice et al. 2009). Almost all of the starch is

consumed by dawn due to the nearly linear conversion of starch to sucrose that occurs at night (Graf et al. 2010).

The rate of starch biosynthesis and degradation is adjusted to optimize carbon allocation in response to changes in day length, for example, plants exposed to only 2 hours of light in a 12-hour light/12-hour dark cycle and are then placed in darkness show a slower rate of starch degradation compared to a normal night. This results in the conservation of starch reserves until dawn (Graf et al. 2010). This rate of starch degradation is adapted to match the length of the night and depends on both the circadian clock and the amount of starch in the leaf (Graf et al. 2010).

However, the mechanism by which starch levels are measured in plants is not fully understood, but the plant's circadian clock is responsible for tracking the length of the dark period. The *Arabidopsis* short-period circadian clock mutant cca1/lhy experiences a shortened clock period of about 17 hours instead of the normal 24 hours and adjusts its starch degradation rate accordingly (Alabadí et al. 2001). When grown under a 12-hour light/12-hour dark cycle, cca1/lhy mutants exhaust their starch reserves three hours before dawn, leading to carbon starvation as indicated by the induction of sugar-repressed transcripts (Graf et al. 2010).

1.3. The building blocks of starch: structure and composition

Starch is a storage polymer made of glucose residues (Figure 1.2) that accumulates as water insoluble, partly crystalline granules of 0.1 - 100 µm in diameter. In higher plant cells, starch granules are stored within plastids as discrete, semi-crystalline structures. These granules are composed of two types of polyglucans, namely amylose and amylopectin that adopt different 3D structures (Yu et al. 2017) and have different physicochemical properties. Their structure are not the same and their synthesis does not involve the same molecular mechanisms (Pfister et Zeeman 2016). Amylopectin the major component accounts for 70-80% of the starch content and the component responsible for determining the granule's precise structure, while the rest being amylose.

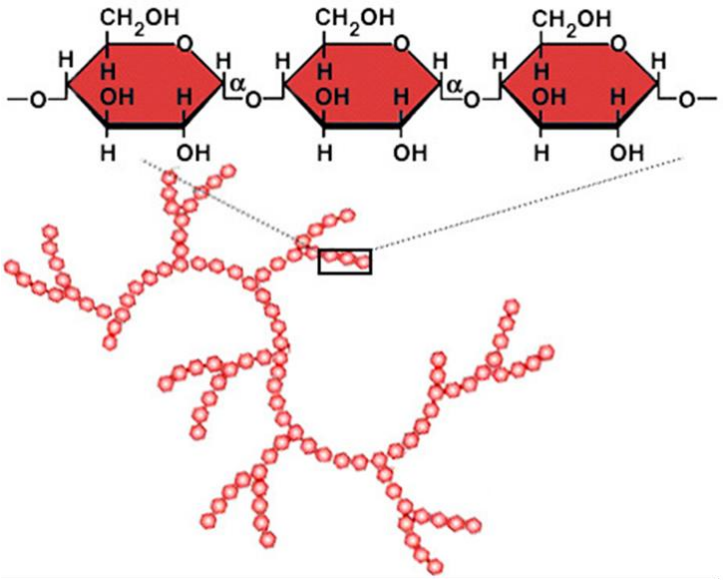


Figure 1.2: ***Structure of starch.*** The linear configuration is a component of a starch molecule in which the glucose units are connected by α α-1,4-glycosidic linkages, and the branching points are connected by α-1,6-glycosidic linkages.

Minor components less than 1% of the starch weight are also found in starch such as proteins, lipids and phosphate (Grimaud et al. 2008).

1.3.1. Structure of amylopectin

Amylopectin is a large, highly branched macromolecule with an enormous molecular weight ranging from 10^7–10^8 Dalton. It is composed of numerous relatively short α-glucan chains, each containing around 18 to 25 glucosyl units connected by α-1,4-D-glycosidic linkages. These chains are interconnected through α-1,6-linkages (Figure 1.3), which make up about 5% of all glucosyl bonds in most starch types (Manners 1989).

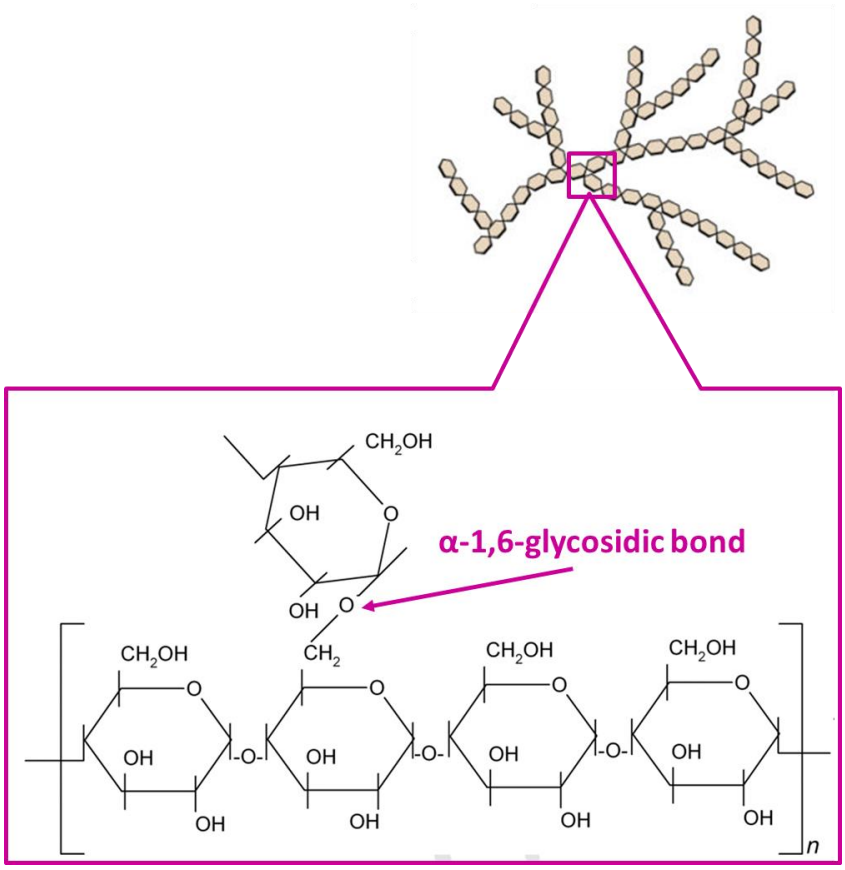


Figure 1.3: ***Chemical structure of amylopectin units.*** Starch granules are made up of α-1,4-glycosidic bonds with few α-1,6- glycosidic bonds, parallel to each other allowing the formation of double helix. The branches are placed in a specific manner, resulting regions of branching called amorphous lamella and regions of exclusively linear chains called crystalline lamella, which form defined growth rings. This highly ordered insoluble structure requires specific enzymes to synthesize and degrade (Coultate, 2001).

The frequency and clustering of α-1,6-linkages play a key role in starch's water-insoluble nature, in contrast to water-soluble glycogen, which has more evenly distributed α-1,6-linkages accounting for roughly 9% (S. Ball et al. 1996; Brust, Orzechowski, et Fettke 2020). Amylopectin has a crucial role in the organization of the starch granules and in determining the structure of the granules in which the distribution of its branches is asymmetrical (Figure 1.4). It gives the molecule a semi-crystalline structure in successive clusters, and this distribution is fundamental for the formation of the granules. This structure is at the origin of the physico-chemical properties of starch (Tetlow et Bertoft 2020).

a. Molecular structure of amylopectin

Amylopectin consists of two major types of unit chains: short and long chains. Short chains have a degree of polymerization (DP) ranging from 6 to 36, with the upper limit varying based on the amylopectin source (Eric Bertoft et al. 2008). Long chains possess a DP of 36 or more. Chains are classified into A-, B-, and C-chains (Hobson, Whelan, et Peat 1951). A-chains are, by definition, completely external and do not carry any branches themselves. In contrast, B-chains include one external segment, a total internal chain segment, and one or more branches. The C-chain is the single B-chain per molecule with a free reducing end (Figure1.4). To analyze the distribution of these chains, enzymatic debranching of the macromolecule is performed, followed by separation of the debranched products using size-exclusion chromatography (SEC) (J.-L. Jane et al. 1994), high-performance anion-exchange or chromatography (HPAEC) (Wong et Jane 1997). The molar ratio of short to long chains (S:L) typically falls between 6 and 19 in most starches, with B-crystalline starches having a generally lower ratio compared to A-crystalline types (Eric Bertoft et al. 2008). A-chains are not connected to other chains, while B-chains are connected to other chains (either A- or B-chains). Each macromolecule contains one C-chain, which has the only reducing-end group but is otherwise similar to B-chains (Hobson, Whelan, et Peat 1951). Additionally, the outermost chain segments in the macromolecule, between the outermost branch points and the non-reducing ends, are known as external chains. On the other hand, the segments located between the branches are known as internal chains, which includes the segment located between the internal branch point and the reducing terminal (Manners 1991). A total internal chain segment includes the entire chain except for the external segment, containing all internal segments and branch points (E. Bertoft 1991).

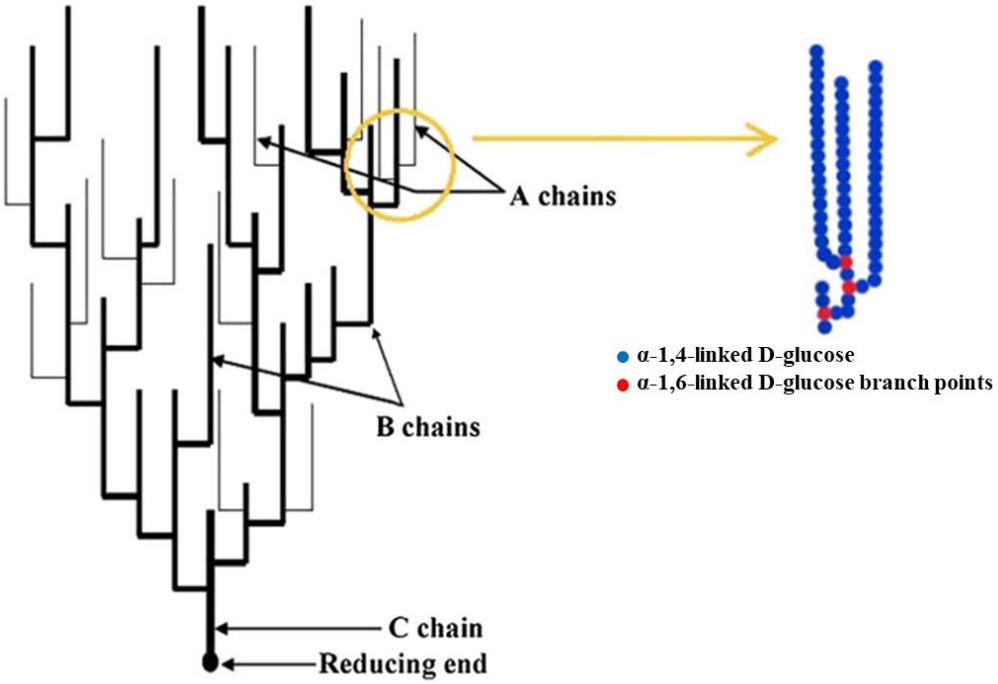


Figure 1.4: ***The clustered structure of amylopectin based on the model proposed by Manners 1989.*** A tree-like structure for the amylopectin. Double helices formation stabilizes the structure of linear α-1,4-chains and this latter form clusters of various arrangements, giving rise to the granule's semi-crystalline characteristics in amylopectin (Fasahat, Rahman, et Ratnam 2014).

Hizukuri discovered a polymodal size-distribution of unit chains in the amylopectin and a periodicity in chain length of about 27 to 28 glucosyl units through size-exclusion chromatography (Hizukuri 1986). He proposed that B2-chains span across two clusters, B3-chains across three, and B4-chains across four, with each cluster being a component of the repeating crystalline and amorphous lamellae. Each cluster corresponds to a 9 nm repeat distance, which aligns with the observed periodicity in chain length.

In most amylopectins, B2-chains make up the majority of long chains, while B3-chains only form a small portion of this group. This indicates that there is a minor part of long chains with decreasing numbers as their length increases, and without any periodicity in length (Eric Bertoft et al. 2008). As a result, the backbone primarily consists of B2-chains, but longer chains can also be found occasionally.

X-ray scattering and electron microscopic analyses indicate that amylopectin molecule clusters stack with a periodicity of about 9 to 10 nm (Figure 1.5). This 9 nm repeat distance of the lamellae corresponds to around 27-28 glucose units in a double-helical conformation (Gernat

et al. 1990). X-ray diffraction patterns further show that neighboring linear chain segments within the clusters form parallel double helices, with each full turn containing 6 glucose units per chain and having a period of 2.1 nm and a diameter of 1.05 nm (Anne Imberty et al. 1991).

Double-helices in starch are arranged in either an A-allomorph or a B-allomorph crystalline pattern, distinguishable through wide-angle X-ray diffraction (WAXS or XRD) (Buleon et al. 1997). A-allomorphs have more tightly packed double-helices and are commonly found in cereal starches. B-allomorphs, on the other hand, are less densely packed, with water molecules essentially replacing one out of four double-helices in the crystalline lattice. They are often found in tuber and root starches, which are more susceptible to amylolysis (Pfannemüller 1987). Some starch granules contain both allomorph types, referred to as "C-types," which can be found in banana (*Musa sp*.) fruits, cassava (*Manihot esculenta L*.), and many legume cotyledons (Bogracheva et al. 1998). The relative crystallinity in most starch granules ranges from 20 to 45% (Gérard et al. 2000). *Waxy* starches (a type of starch containing a high percentage of amylopectin) generally have slightly higher values than non-*waxy* counterparts, implying that the amylose component does not contribute significantly to the crystallites. It is important to note that external chain segments in amylopectin usually make up around 60% of the macromolecule.

b. Higher order structure of starch

Although starch granules display a wide variety of appearances, their internal structure is remarkably consistent across plant species (Paul J. Jenkins, Cameron, et Donald 1993). Most granules, perhaps excluding the smallest ones, exhibit a regular pattern of darker and lighter ring-like structures when viewed under a microscope (Ambigaipalan et al. 2011). These "growth rings" surround the starch granule's core, known as the hilum (Figure 1.5), which is generally considered the starting point of starch granule biosynthesis (Baker, Miles, et Helbert 2001).

The exact differences in structure between the light and the dark rings remain unclear. It is thought that the rings are composed of alternating regions amorphous (less compact), and the crystalline (more compact) regions, which give rise to different refractive phenomena viewed under the microscope. The semi-crystalline rings likely contain more ordered structures, with amylopectin being the main contributor. The organized arrangement of the macromolecules within the granules is demonstrated by the birefringent of starch granules, which allows them

to exhibit a "Maltese cross" pattern when viewed under cross-polarized light (Figure 1.5), and that amorphous regions are lost more quickly than ordered crystalline rings under controlled digestion conditions (Eric Bertoft 2013).

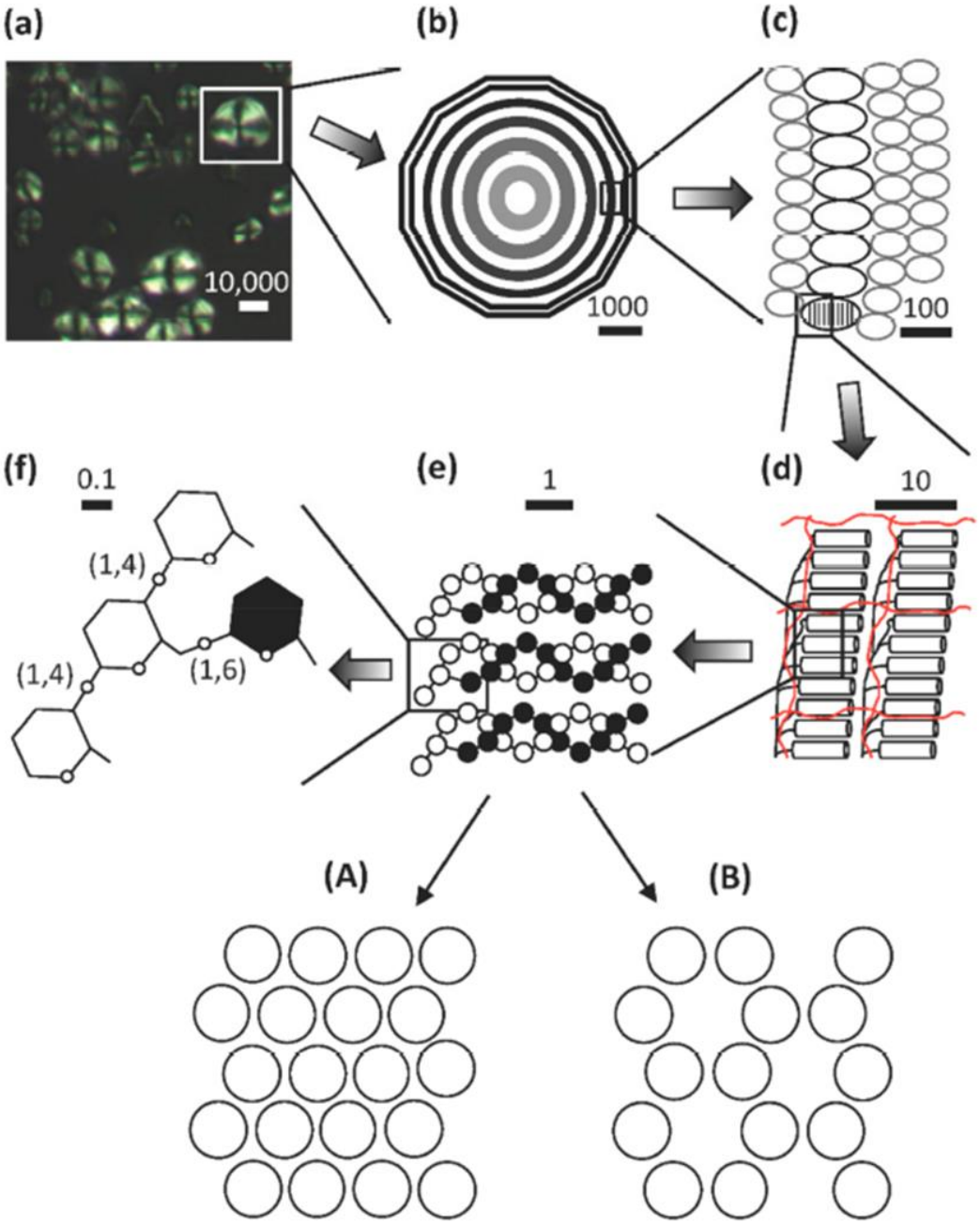


Figure 1.5: ***From granules to individual glucosyl units***: (a) Maize starch granules show a characteristic "Maltese cross" pattern when viewed under polarized light. (b) A theoretical granule is shown with growth rings radiating outward from the hilum. (c) Blocklets are shown within semi crystalline (black) and amorphous (gray) rings. (d) Crystalline and amorphous

lamellae are formed by double helices (shown as cylinders) and branched segments of amylopectin (shown as black lines). Amylose molecules, indicated by the red lines, are interspersed between the amylopectin molecules. (e) Three double helices of amylopectin are shown, with each double helix consisting of two polyglucosyl chains. Glucosyl units are represented by black circles, which are the A-chains that are unsubstituted, and white circles, which are the B-chains that are substituted with other chains. Double helices can form either A- or B-type allomorphic crystals (shown as A and B, with circles symbolizing double helices viewed from the edge). (f) Glucosyl units are shown with α-(1,4) and α-1,6-linkages at the base of the double helix. A bar scale (measured in nanometers) gives a rough indication of the size dimensions (Tetlow et Bertoft 2020; Pérez, Baldwin, et Gallant 2009).

Atomic force microscopic studies have revealed bulb-like structures, known as blocklets, with diameters ranging from 20 to 100 nm on the surface and within starch granules. These blocklets were isolated as discrete units from *waxy* maize starch (Perez Herrera, Vasanthan, et Hoover 2016). While their exact nature is uncertain, blocklets are hypothesized to be a result of amylopectin super-helix interactions (Gallant, Bouchet, et Baldwin 1997). Based on their size, the blocklets are thought to represent either individual amylopectin molecules or smaller cluster of those molecules. In this structure, amylopectin is organized in layers of alternating amorphous and crystalline sheets, with blocklets containing layered stacks of these lamellae contributing to the formation of the semi-crystalline rings. Tang et al. suggested that the amorphous rings also contain blocklets, but with a more "imperfect" architecture (Tang, Mitsunaga, et Kawamura 2006).

Tang et al. (Tang, Mitsunaga, et Kawamura 2006) and Perez-Herrera et al. (Perez Herrera, Vasanthan, et Hoover 2016) further proposed that amylose is distributed both between and within the blocklets, extending across multiple blocklets while interacting with amylopectin molecules. It is generally believed that amylose molecules are primarily present in the amorphous state within granules, either in amorphous rings or lamellae. However, they may also penetrate the stacks of lamellae, introducing "imperfections" into the crystalline organization (P. J. Jenkins et Donald 1995; Kozlov et al. 2007).

c. Understanding the macromolecular structure of amylopectin

The fundamental structure of the amylopectin is mainly determined by the orientation of its constituent chains. Therefore, insight into the configuration of the long and short chains, including the A- chains, B-chains, C chain, and the outer and inner segments, is critical to understanding the complex structure of the macromolecule. The simplest approach to determine this structure is to separate the branched components of the macromolecule. For this purpose, the endo-acting α-amylase of *Bacillus amyloliquefaciens* was used (E. Bertoft 1991). This specific α-amylase from *B. amyloliquefaciens* primarily attacks extended chain segments between branching points, whereas internal chain segments with three or fewer residues prove to be resistant to its action, and probably those with four residues as well (Robyt et French 1963).

Nevertheless, this enzyme also cleaves external chain segments, which reduce their lengths to an average of about two residues (Umeki et Yamamoto 1972). The remaining branched limit dextrins consist of closely-related branching points with inner chain lengths (ICL) of three residues or less. Thus, these limit dextrins constitute the branched components of amylopectin and are known as building blocks (Eric Bertoft et al. 1999) (see Figure 1.6). The size distribution of building blocks has been analyzed using size-exclusion chromatography (SEC) and high-performance anion-exchange chromatography (HPAEC), revealing a remarkably similar distribution regardless of the botanical source of amylopectin (Eric Bertoft 2013). Building blocks can be isolated, structurally characterized, and classified into different groups based on the number of chains they contain (Figure 1.6a) (Eric Bertoft, Koch, et Åman 2012).

Group 2 building blocks consist of only two chains (and a single branch point), meaning they have an A-chain connected to a B-chain (more accurately defined as a C-chain in this case; Figure 1.6c) with a DP range between 5 and 9. Group 3 building blocks comprise three chains (and two branches) with a DP range of 10 to 14, while Group 4 building blocks have four chains (DP 15 to 19) (Figure 1.6). Group 5 building blocks are isolated as mixtures of dextrins with a DP range of 20 to 35 and consist of an average of six chains, while Group 6 (DP > 35) contains 9 to 12 chains (Eric Bertoft, Koch, et Åman 2012). The most abundant category of building blocks in all amylopectins is Group 2, which represents about 50% of all blocks by number (Figure 1.6b), and about 25% of all amylopectin branch points are located within this group. The second most common group is Group 3, which accounts for about 30% of the blocks, and represents another 25% of amylopectin's branch points. Group 3 building blocks display one of

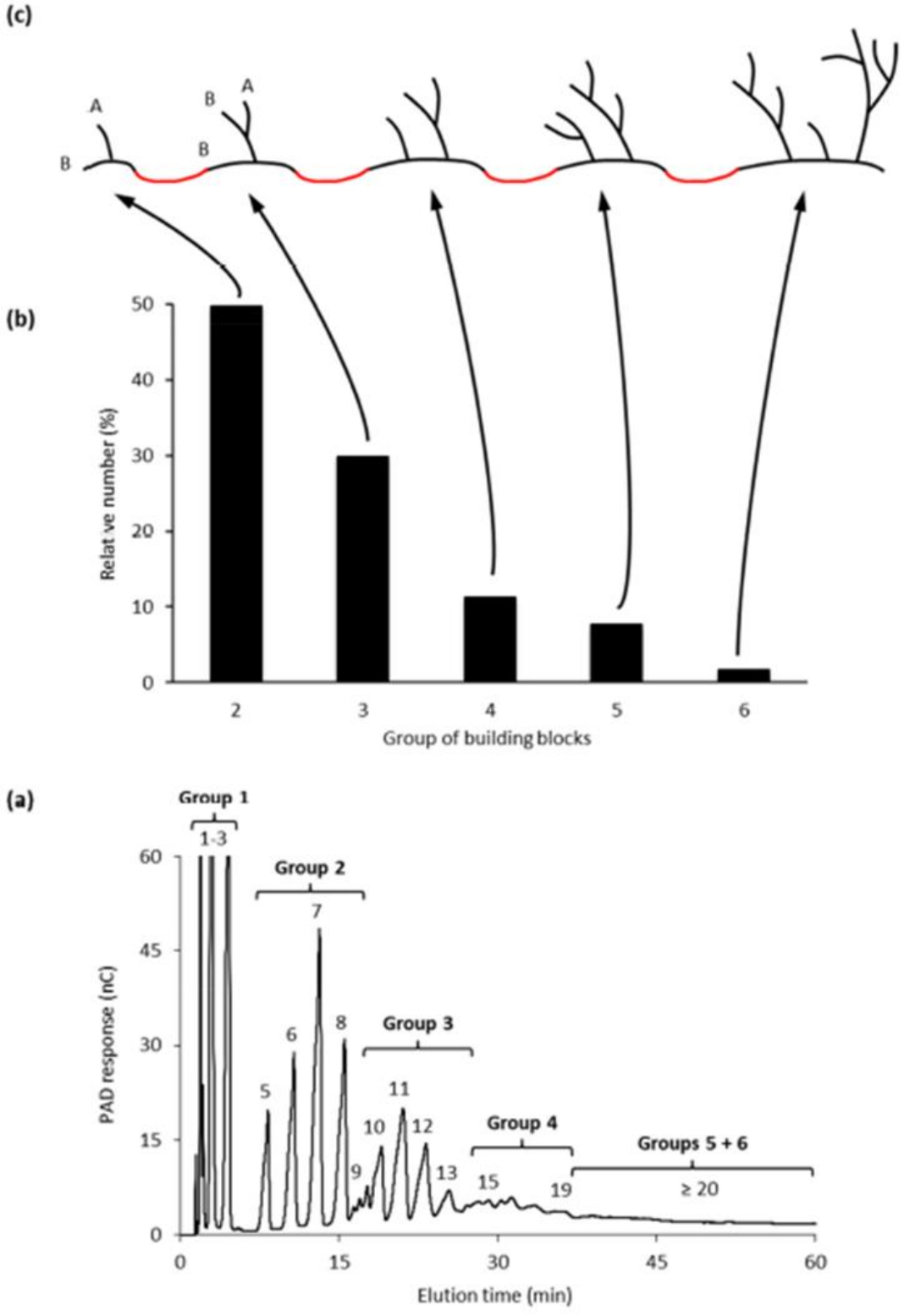


Figure 1.6: ***Building block unit composition in amylopectin***: **(a)** The size distribution of amylopectin building blocks from mung bean cotyledon was determined by HPAEC-PAD analysis after β-amylolysis and successive α-amylolysis. The building blocks are categorized into groups 1-6, with their corresponding DP values. Group 1 is composed of glucose, maltose,

and maltotriose, whereas groups 2-6 are composed of branched dextrins. However, HPAEC does not resolve individual peaks for groups 5 and 6. **(b)** The relative amount of each branched building block for each group, with the PAD response adjusted to measure the amounts of each DP. **(c)** The basic structures of the branched building blocks, showing A- chains or B-chains for both groups 2 and 3. Group 5 consists of building blocks with 5 to 7 chains, whereas group 6 contains 8 or more chains. The diversity of the possible building block structures greatly increases with the chain number. The Red lines between building blocks indicate the interblock segments (Tetlow et Bertoft 2020).

two primary structural arrangements, referred to as the Haworth (one A- chain and two B-chains, (Haworth, Hirst, et Isherwood 1937)) and Staudinger configurations (two A chains and one B-chain). The first is shown in Figure 1.6c. It's likely that a combination of these configurations exists in amylopectin, but the ratio may vary among different types of starches. Group 4 represents about 10% of the building blocks in terms of quantity, and the potential chain combinations in the dextrins increase exponentially with the number of chains. Group 5 constitutes between 5 and 8% of the building blocks, and the number of different dextrins is so vast that single peaks are not distinguishable by HPAEC, but can be quantitatively evaluated by SEC. Finally, Group 6 represents only a few percent of all blocks, with about 16% of all branch points included in this group.

The individual α-glucan chains within the building blocks are remarkably short. HPAEC shows a peak at DP 2 or 3 (mainly A-chains) and another peak at DP 5, 6, or 7 (B-chains) for most starches. The internal chain length (ICL) in building blocks is minimal, averaging from 1.2 to 2.2 glucosyl units. The total ICL varies from 4.1 to 5.8 and tends to increase with the size of the block (Eric Bertoft, Koch, et Åman 2012).

Building blocks are connected by interblock segments (IB-S, shown in red in Figure 1.6c). The sizes of these segments are predicted by studying the number and structure of the building blocks within the larger α-dextrins composed of two or more blocks. When *B. amyloliquefaciens* α-amylase acts between the building blocks, small, linear saccharides such as Glucose, maltose, and maltotriose are obtained as "Group 1" (Figure 1.6a), and the amount of these saccharides measures the inter-block chain length (IB-CL). Typically, the IB-CL varies between 5 and 8 Glucose residues (Eric Bertoft, Koch, et Åman 2012).

The resulting question concerns the location of the IB-Ss within the amylopectin macromolecule. Extremely small clusters of building blocks (α-dextrins) that have been purified do not have longer chains. However, somewhat larger clusters consisting of three or more building blocks tended to have longer chains (Eric Bertoft, Koch, et Åman 2012).

This strongly suggests that the majority of the blocks are linked to each other by the long chains of amylopectin. Therefore, the most likely structural model of amylopectin is that the majority of the building blocks are distributed along these extended chains of amylopectin (Figure 1.7a). These elongated B-chains are likely connected to each other by α-1,6-linkages, creating a longer backbone within the macromolecule. The distribution of building blocks from diverse groups within the backbone appears to be totally random (Källman et al. 2015).

It has been proposed that shorter B-chains (DP < 36) form short branches that connect "external" building blocks to the backbone (Figure 1.7a). Some types of amylopectins, principally A-crystalline storage starches from cereal endosperms (with high S:L ratios), may contain more of such branches and also involve some shorter B-chains in the backbone, compared to the amylopectin found in B-crystalline starches (with low S:L ratios) (Eric Bertoft, Koch, et Åman 2012).

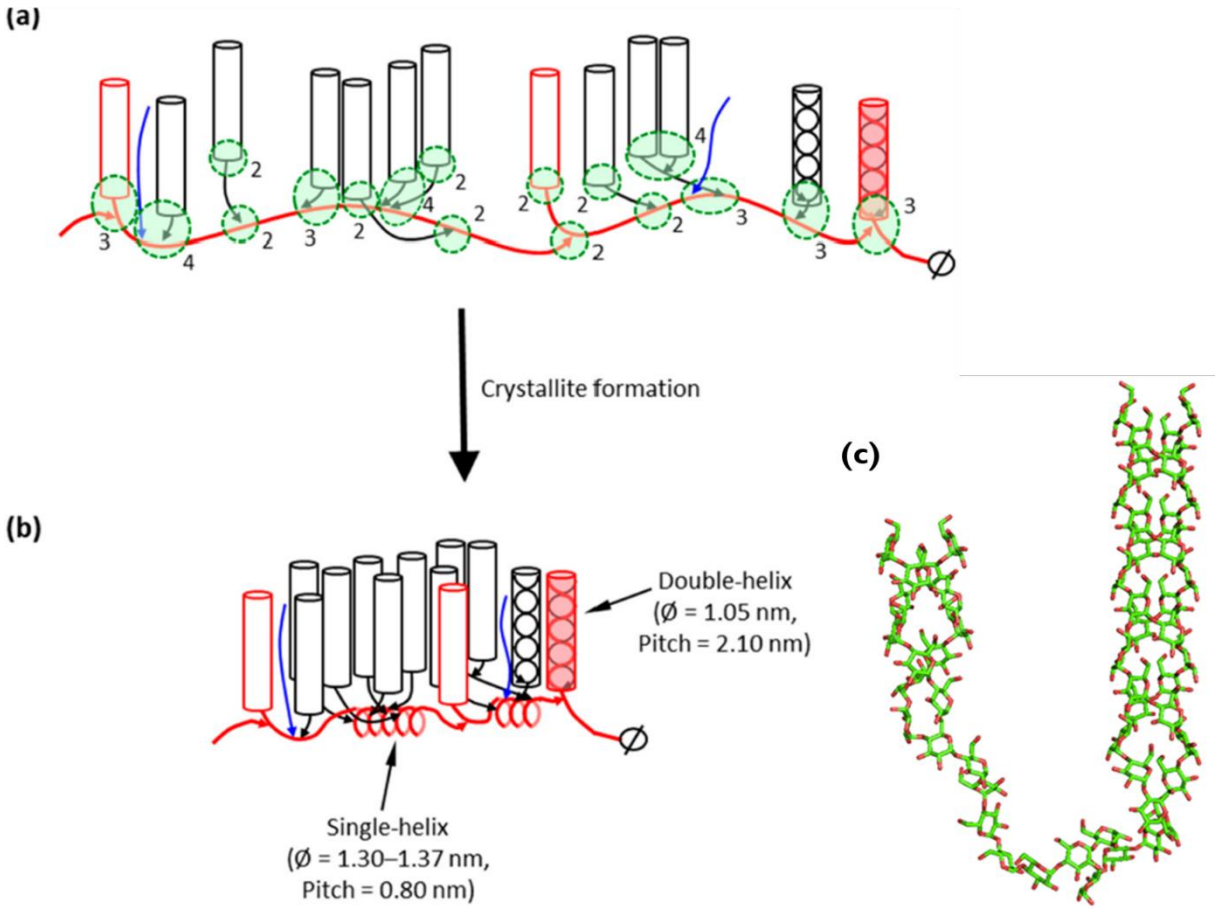


Figure 1.7: ***Amylopectin Backbone Structure***: **(a)** The primary structure shows the backbone (in red) with unique interblock segments between the branched building blocks (encircled in green). The numbers represent the building block groups, with group 2 having two chains, group 3 having three chains, and group 4 having four chains (minor groups 5 and 6 are not shown in the figure). The short chains create double-helices (cylinders), and the outer segments of the long chains that form the backbone also contribute to double-helices (red cylinders). Some of the chains do not make double-helices (blue chains) and are speculated to induce distortions among the double-helices, affecting crystallite formation. **(b)** The flexible backbone can generate single-helical segments, that constrict the macromolecule and bring individual double-helices closer together, allowing for crystallization. To construct bigger crystallites with one of the characteristic A- or B-allomorphs, double-helices of several amylopectin molecules must be in close proximity. The sizes of the single- and double-helices are constructed on data provided by (Zobel, French, et Hinkle 1967) and (Anne Imberty et al. 1991), respectively. **(c)** Structure of amylopectin. The image displays the double helix structure of amylopectin. This structure was obtained from https://glyco3d.cermav.cnrs.fr.

The backbone of amylopectin is believed to be a highly flexible structure. The larger the interblock segments, the more flexibility they give to the backbone, which significantly affects the properties and functionality of starch (Vamadevan, Bertoft, et Seetharaman 2013). The internal chains of amylopectin have been shown to bind iodine (I_3^-) (Shen et al. 2013), which suggests that the backbone segments can create helical inclusion complexes with iodine and that helical segments can exist in native starch even in the absence of a complexing agent. Those segments are probably the interblock segments, the length of which might theoretically create single helices with about six residues per turn (Figure 1.7b). This would facilitate the approach of double helices to each other along the backbone and enhance the interaction between the double helices, like the formation of crystallites.

In addition to the two primary groups of short and long chains, amylopectin from certain plants, such as *indica* rice varieties, wheat, potato, and cassava, also contains "superlong" (or extra-long) chains (Isao Hanashiro et al. 2005). These superlong chains are composed of hundreds or even thousands of glucosyl units, making them similar in length to the amylose moiety. The super-long chains seem to have some elongated branches, reflecting branched amylose (Isao Hanashiro et al. 2005). The location of the super-long chains inside the amylopectin macromolecule remains unknown, but it is reasonable to assume that they are attached to the backbone without bearing double helices. The quantity of super-long chains varies considerably, ranging from a few percent by weight in wheat, potato, and cassava (Laohaphatanaleart et al. 2009) to over 10% in *indica* rice varieties (Takeda et al. 1987). Super-long chains do not occur in *waxy* starches, and the enzyme that is responsible for their synthesis is granule-bound starch synthase I (GBSS I), which is the same enzyme that synthesizes amylose and is either inactive or absent in *waxy* plants (Wikman et al. 2014).

d. Chain Length Correlation with Crystalline Allomorphs in Amylopectin

It has long been accepted that the average chain length (CL) of amylopectin is associated with the crystal allomorphs of starch granules. Shorter CLs result in the A-allomorph, while longer CLs lead to the B-allomorph, and intermediate CLs tend to produce the C-allomorph pattern (Hizukuri 1986). However, it is important to note that double-helices are comprised of only the external segments of the chains (Eric Bertoft et al. 2008). The chain length conditions for the two inner chain segments (m and n, in which the m is the inner segment of the main chain and

n is the inner segment of the side chain joining to another double helix) connecting two double helices in the crystal lattice have been studied. It was discovered that only a limited combination of chain lengths for these two segments would result in either the A- or B-allomorph (Palmer, Macaskie, et Grewel 1983). Only the combinations of ICL with m = 1 and n = 3, or m = 4 and n = 6, could produce the A-allomorph structure, while the latter combination, along with m = 6 and n = 4, results in the B-allomorph. Additionally, m = n = 7 could produce two parallel double helices. All other combinations of m and n did not yield parallel double helices.

1.3.2. Structure of amylose

Amylose, the second component and the minor polyglucan of starch accounts for approximately 15 to 35% of the carbohydrate content in most starches (Seung et al. 2020). It is mostly linear, with very few branching points about 1% of α-1,6-linkages (Tetlow et Bertoft 2020) (Figure 1.8). This loosely branched molecule (Buleon et al. 1998) is present in all-natural starch granules, suggesting an essential structural role. In some plant species, amylose content can be as low as 5 to 8%, like in *Arabidopsis thaliana* while in *waxy* starches, it is significantly reduced or even absent (Eric Bertoft 2017). The exact biological function of amylose remains unclear, but it is thought to enhance the storage capacity and packing of starch granules (Donald 2001).

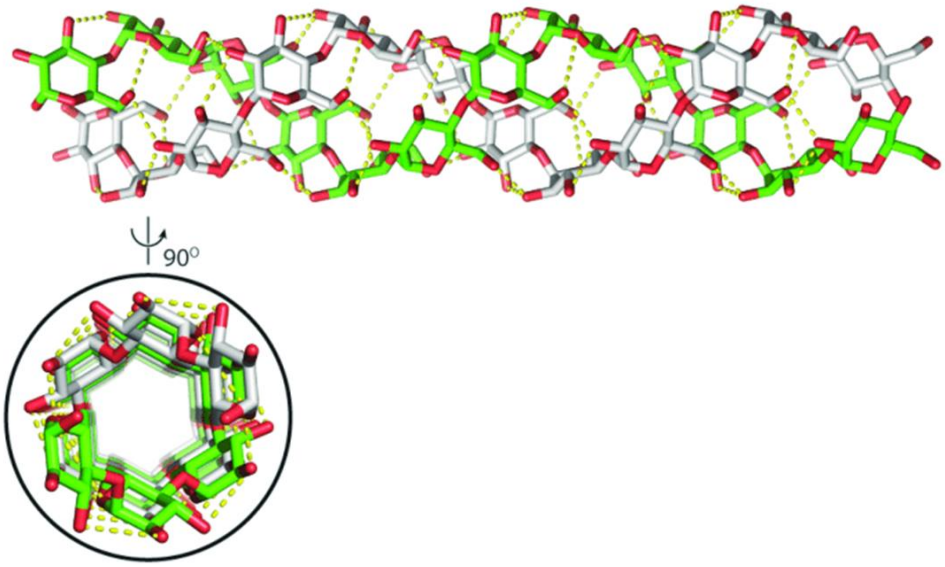


***Figure 1.8:* Structure of Amylose.** Image displays the double helix structure of A-type amylose. This structure was sourced from http://polysac3db.cermav.cnrs.fr.

Amylose α-glucan chains are significantly longer than those of amylopectin, generally consisting of several hundred to thousands of α-1,4-linked D-glucosyl units with an estimated mass of 10^6 Da (Doutch et Gilbert 2013; Eric Bertoft 2017) (Figure 1.9). Amylose molecules can be linear with a single long chain or slightly branched via α-1,6-linkages, with the number of chains in branched amylose varying between 5 and 20 (Hizukuri et al. 1981), which is depend on the starch source and these chains are shorter than those in the linear fraction. It is synthesized in the semi crystalline matrix: created by amylopectin. Research on amylose branching patterns has revealed a bimodal chain length distribution: AM1 consists of relatively short chains with a degree of polymerization (DP) of 100-700, while AM2 comprises longer chains with a DP of 700-40,000 (K. Wang, Henry, et Gilbert 2014).

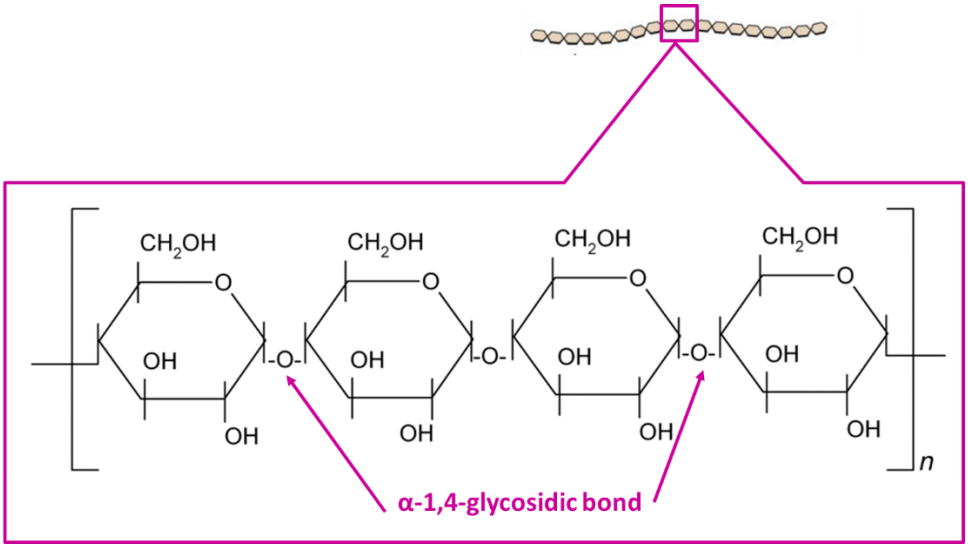


Figure 1.9: ***Chemical structure of amylose units.***

X-ray scattering data indicate that amylose is likely situated within the amorphous regions of amylopectin blocklets (P. J. Jenkins et Donald 1995), with a higher abundance near the granule surface (Pan et Jane 2000; J. Jane et Shen 1993). The amylose makes the starch granule denser by filling the spaces in the semi-crystalline matrix that is formed by the amylopectin (Pfister et Zeeman 2016). Early research suggest that amylose, but not amylopectin, leaks from starches into the water, indicating a weak association with the amylopectin matrix (R. F. Tester et Morrison 1992).

1.3.3. Other components

Starch granules also include minor components like lipids, minerals, and proteins, which together make up a small percentage of the granule's total weight. Among these components, only phosphate is covalently attached to the starch, specifically to the amylopectin part. Starch contains a low protein concentration (around 0.1–0.7%), primarily consisting of granule-bound starch synthase (GBSS), responsible for amylose synthesis, and other enzymes involved in amylopectin synthesis, such as starch synthases (SSs) and starch-branching enzymes (BEs) (Grimaud et al. 2008).

Root and tuber starches typically have very low lipid content but are relatively rich in phosphate groups covalently linked to amylopectin's α-glucan chains (Nitschke et al. 2013; Hizukuri et al. 1970). B-crystalline starches often contain higher levels of covalently linked phosphate compared to the A-crystalline. Potato (*Solanum tuberosum L.*) starch, in particular, is known for its relatively high phosphate content. Approximately two-thirds of the phosphate is bound to the C6 position of glucosyl units, and 20 to 30% at the C3 position (Hizukuri et al. 1970). A tiny percentage of phosphate in starch is located at the C2 position (Nitschke et al. 2013; Hizukuri et al. 1970). Other elements, such as potassium, calcium, magnesium, and sodium, are comparatively rare (Blennow et al. 2005; Dhital et al. 2011).

Additionally, starch granules comprise several proteins attach within the water-insoluble amylopectin matrix, recognized as granule-associated proteins (Mu-Forster et al., s. d. 1996; Denyer et al. 1993). Most of these proteins are present in the granules and involved in granule biosynthesis at some point. They remain attached even after thorough washing of starch granules with detergents and solvents.

1.4. Starch granules morphology

Starch granules occur in a wide range of sizes and shapes, depending on their botanical origin. The size of a starch granule can vary from about 100 nm for the smallest to 100 µm for the largest (Figure 1.10) (Buléon et al. 1998). Like in potato and orchid (*Phajus grandifolius*) who have large size distributions of about 100 µm (Chanzy et al. 2006), while in some cereals like rice the size distributions is about 0.1 - 7 µm (Waterschoot et al. 2015). It exists also an

intermediate size distributions of about 10 - 20 µm like in the case of maize (*Zea mays L.*) (Chen et al. 2006). There are two populations of granules which have the characteristic of bimodal size distribution in the endosperm of the starch granules: the A-granules which is the large granules with a diameter ranging from 15 to 35 µm with a flatten shape, and the B-granules which is the smaller granules with a diameter ranging from 2 to 10 µm with a spherical shape characteristic (Saccomanno et al. 2017). Important members of this group such as wheat (*Triticum aestivum L.*), barley (*Hordeum vulgare L.*), rye (*Secale cereale L.*) and oats (*Avena sativum L.*) (Saccomanno et al. 2017).

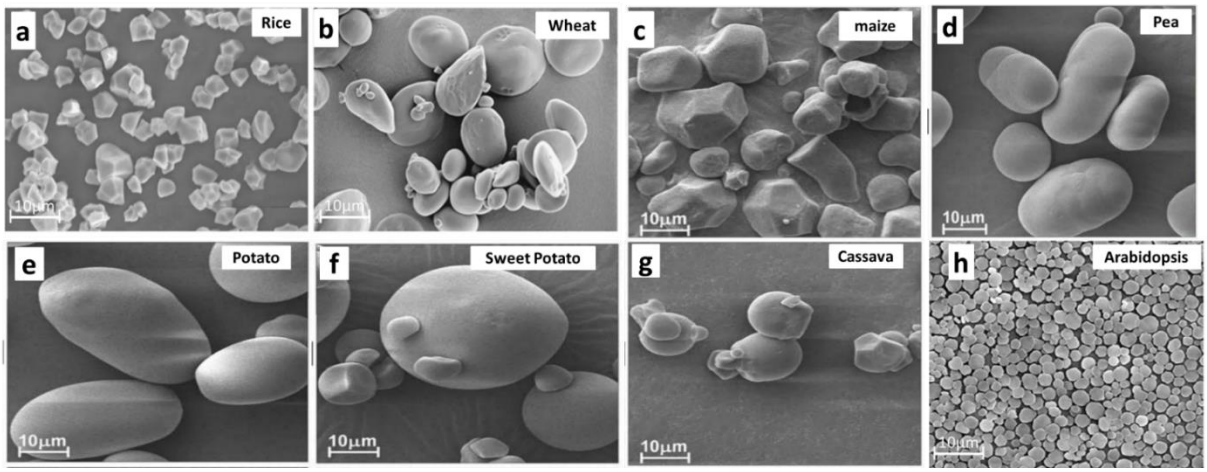


Figure 1.10: ***Scanning electron microscopy (SEM) of starch granules from different species***: (a, at 2000X) rice *(Oryza sativa)*, (b, at 850X) wheat (*Triticum aestivum*), (c) maize (*Zea mays L.*), d) pea (*Pisum sativum L.*), (e) potato (*Solanum tuberosum*), (f) sweet potato (*Ipomoea batatas*), (g) cassava, and (h) *Arabidopsis thaliana*. Adapted from (Fasahat, Rahman, et Ratnam 2014; Khalid et al. 2017) and (amidotheque.cermav.cnrs.fr).

The shape of a starch granule can vary also from round or polygonal (e.g., *Arabidopsis thaliana* transitory starch and maize (*Zea mays*)), pyramidal (lesser yam (*Dioscoreaceae species*)), disc shaped (A-granules in wheat and barley), spherical shape such as the reserve starch of sweet potato (*Ipomoea batatas*), to oval ((*potato (Solanum tuberosum*), peas (*Pisum sativum L.*)) (J.-L. Jane et al. 1994). Also, some plant groups, such as rice (*Oryza sativa L.*), produce compound granules (Tetlow et Emes 2017). Waxy varieties generally have granule shapes similar to their non-waxy counterparts, while high-amylose starches often exhibit irregular or elongated granule shapes (e.g., high-amylose (*amylose-extender*) maize, wrinkled (rr) peas) or amylose-only barley (Chen et al. 2006; Goldstein et al. 2016; Gallant, Bouchet, et Baldwin 1997). This

further supports the notion that amylopectin is the primary contributor to the fundamental architecture of starch granules in non-mutant (wild-type) plants.

1.5. *Arabidopsis thaliana*

Arabidopsis thaliana (Figure 1.11), a small diploid plant, is part of the *Brassicaceae* family, which also includes plants like mustard, cabbage, and radish. This plant offers significant advantages for fundamental research in genetics and molecular biology. It has become a crucial model organism in the field of plant biology and genetics due to its small genome, short generation time, and ease of cultivation under laboratory conditions. In 1943, Friedrich Laibach demonstrated its potential as a model organism by correctly identifying the number of chromosomes in the plant and isolating the first mutants using X-rays.

The study of *Arabidopsis thaliana* gained momentum in the 1980s, thanks to the development of efficient transformation methods that enabled the production of insertion mutants (Koornneef et Meinke 2010). *Arabidopsis thaliana* has become the most widely used model organism for studying starch metabolism. The most commonly used varieties for research are *Columbia (Col)*, *Wassilewskija (WS)* and *Landsberg erecta (Le)*.

***Figure 1.11:* Arabidopsis thaliana plant**

Arabidopsis thaliana possesses several characteristics that make it an attractive model for study, its small size and rapid life cycle of *Arabidopsis thaliana* make it an ideal model organism for laboratory research. It can grow to maturity in just six weeks, and its small stature allows for the cultivation of a large number of plants in limited space. Additionally, *Arabidopsis thaliana* can self-fertilize, which simplifies the process of generating stable genetic lines for study (Meinke et al. 1998).

2. Starch biosynthesis and degradation

2.1. Starch metabolism

2.1.1. Starch biosynthesis

Despite that starch is a chemically simple molecule, its structure is complex, and therefore involves a complex synthesis mechanism that requires several dozens of enzymes/proteins and protein-complexes to regulate the entire process. In short, amylopectin is produced by the coordinated action of several enzymes, including soluble starch synthases (SSs), starch branching enzymes (BEs), and starch debranching enzymes (DBEs), while, amylose is synthesized by the activity of granule-bound starch synthase (GBSS) (Grimaud et al. 2008). In this section, the different steps of starch synthesis and the exact role of each enzyme involved will be explained.

In higher plants, starch is synthesized within plastids, such as chloroplasts in leaves and amyloplasts in starch-storing tissues such as endosperm, roots and tubers (Tetlow et Bertoft 2020a; Pfister et Zeeman 2016). Research on crops and model species like *Arabidopsis* has identified the enzymes responsible for starch synthesis and degradation that are well-conserved across various plant species, providing valuable insights into starch formation (Tetlow et Bertoft 2020). However, starch structures can vary widely in terms of shape, size, and number of granules between different tissues and species, due to the duplication and specialization of starch enzymes with unique or partially overlapping functions. Starch degradation is more complex than synthesis, involving more steps and enzymes to convert starch to maltose and glucose, which are also fine-tuned at multiple levels (Tetlow et Bertoft 2020).

There are four major steps in starch synthesis: (1) Initiation (2) glucan chain elongation, (3) branching, and (4) debranching. The elongation process is catalyzed by starch synthase (SSs) enzymes, which use adenosine 5'-diphosphate-glucose (ADP-glucose) as a glucosyl donor to elongate glucose chains through α-1,4-linkages (Fujita et al. 2006). Starch branching enzymes (SBEs) are responsible for branching glucan chains, transferring chain segments to adjacent chains via α-1,6-linkages through glucanotransferase reactions (Tetlow et Emes 2014). Mispositioned glucosyl chains are removed by starch debranching enzymes (DBEs), which hydrolyze α-1,6 linkages to homogenize the structure (Yasunori Nakamura 1996). Several

isoforms of SSs, SBEs, and DBEs play specific roles in synthesis, contributing to the periodic branching pattern of amylopectin and the overall granule matrix structure. These enzymes exert their actions simultaneously and interdependently, and there are different classes among each type of enzyme (Pfister et Zeeman 2016).

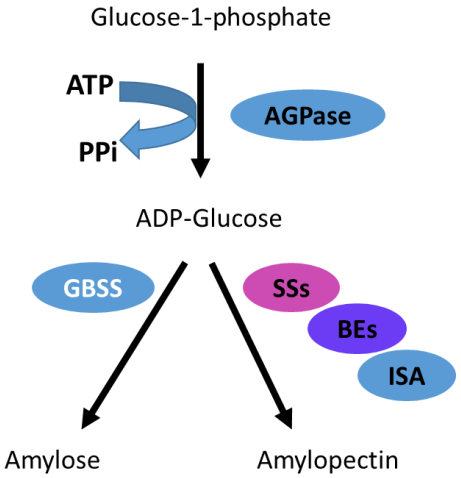


Figure 2.1: ***Overview of the starch biosynthesis pathway***. ADPglucose pyrophosphorylase (AGPase) produces ADPglucose, the substrate of starch synthases (SSs). Granule-bound starch synthase (GBSS) synthesizes amylose, while soluble SSs, branching enzymes (BEs) and isoamylase-type debranching enzyme (ISA) collectively synthesize amylopectin (Pfister et Zeeman 2016).

2.1.2. Synthesis of the precursor ADP-glucose for the synthesis of starch

The starting point for starch synthesis is ADP-glucose (ADP-Glc), with its precursors varying depending on the tissue. In photosynthetically active leaf chloroplasts, ADP-glucose production is directly connected to the Calvin-Benson cycle, in which fructose-6-phosphate is converted to glucose-6-phosphate (Glc-6-P) by phosphoglucose isomerase (PGI), and then to glucose-1-phosphate (Glc-1-P) via phosphoglucomutase (PGM) (Tiessen et al. 2002). This Glc-1-P is obtained by gluconeogenesis from Glyceraldehyde-3-phosphate (G3P) in the Calvin cycle (Tiessen et al. 2002). ADP-glucose pyrophosphorylase (AGPase) initiates starch synthesis in a committed rate-limiting reaction that converts Glc-1-P and adenosine triphosphate (ATP) to ADP-glucose and pyrophosphate (PPi) (Tiessen et al. 2002), however, the more abundant the Pi, the more inhibition to the activity of the AGPase. This process accounts for around 30%-

50% of *Arabidopsis* leaf photoassimilates being allocated to starch (Stitt et Zeeman 2012). In contrast, heterotrophic tissues acquire sucrose from source tissues, which is then converted into hexose-phosphates in the cytosol. Glucose-phosphates (mainly Glc-6-P, even though Glc-1-P transport has also been observed) and ATP are transported into the plastid for ADP-glucose synthesis (Geigenberger 2011). The cereal endosperm pathway is unique, as most AGPase activity (usually 80% or more) is present in the cytosol, and ADP-glucose is directly imported into the plastid through a cereal-specific adenine nucleotide transporter called BRITTLE1 (Bt1) (Martha G James, Denyer, et Myers 2003).

AGPase, crucial for carbon partitioning and starch synthesis, comprises two large regulatory subunits and two small catalytic subunits. These four subunits cooperatively interact to form a heterotetramer that responds to various post-translational regulations. In *Arabidopsis,* there are two genes encoding small subunits and four genes encoding large regulatory subunits, each with distinct spatial expression patterns. This may lead to the formation of AGPase with unique properties in different plant tissues.

2.1.3. Starch granule initiation

The initiation of starch granules is currently an active and developing area of research. It seems that the process of starch granule initiation is dependent on a network of proteins, both catalytic and non-catalytic. In order for starch granule initiation to occur, either MOS needs to be synthesized from soluble sugars in cells that produce storage starches, or semi-crystalline amylopectin needs to be built from existing structures in cells that produce transient starch, with a remnant of the granule often remaining after a period of darkness (Zhu et al. 2015).

a. Proteins Associated with Granule Initiation

The starting point of a starch granule is typically a central core called the hilum, which has an unclear structure but is believed to have a disordered α-glucan structure based on X-ray data (Gregory R. Ziegler, John A. Creek, et Runt 2005; Buleon et al. 1998; Vamadevan et al. 2014). The initiation of the hilum and the subsequent formation of starch granules require the SSIV isoform of a single starch synthase (Lundquist et al. 2017). Recent research suggests that another protein, SSIII, may also be involved in the initiation process, as there seems to be some overlap in its action with SSIV (Szydlowski et al. 2009; Leterrier et al. 2008). In *Arabidopsis,* SSIV interacts with a group of non-catalytic proteins called PTST proteins, which act as

regulatory scaffold proteins (Lohmeier-Vogel et al. 2008). Specifically, SSIV interacts with a protein called PTST2 (Seung et al. 2017), which recognizes a specific three-dimensional shape of MOS via its CBM48 and then forms a complex with SSIV, allowing SSIV to elongate the α-glucan (Janeček, Svensson, et MacGregor 2011). PTST2 may then be released and bind to other MOS structures for further interaction with SSIV (Seung et al. 2017). The ability of SSIV to dimerize is crucial for its catalytic activity and for forming interactions with other proteins.

According to research, SSIV requires pre-existing α-glucan chains in order to elongate, and it is particularly active with maltotriose (Szydlowski et al. 2009). On the other hand, SSIII has been found to be capable of forming unprimed α-glucan in the presence of ADP-Glc (Szydlowski et al. 2009), and starch phosphorylase (SP), also known as Pho1, has the ability to generate and elongate maltodextrins (MOS) even without the α-glucan primer, using Glc1P (Malinova et al. 2014). The priming actions performed by SSIII and SP have been recognized as essential elements in the granule initiation pathway (Malinova et al. 2014; 2017; Yasunori Nakamura et al. 2017). In rice endosperm, SP forms a protein assembly with the disproportionating enzyme (DPE) (Hwang et al. 2016), , and it's postulated that this assembly could provide MOS substrates for additional elements of the granule initiation system (Seung et Smith 2019). The DPE also plays a role in adjusting the length of the MOS chain, and could contribute to the granule initiation process and subsequent steps in starch synthesis (Hwang et al. 2016; Bresolin et al. 2006).

Non-catalytic proteins are also believed to play crucial roles in starch granule initiation by facilitating protein scaffolding and suborganellar positioning. One such protein is PROTEIN INVOLVED IN STARCH INITIATION1 (PII1), a newly discovered chloroplast protein in *Arabidopsis* that forms a complex with SSIV and may be essential for SSIV catalytic activity (Vandromme et al. 2019). PII1, also known as MYOSIN-RESEMBLING CHLOROPLAST PROTEIN (MRC), was discovered to be one of the two proteins that interact with PTST2 (Seung et al. 2018). PTST2, in turn, interacts with other coiled-coil containing proteins, including thylakoid-associated MAR-BINDING FILAMENT-LIKE PROTEIN (MFP1). Research on *Arabidopsis* has revealed many components of the granule initiation machinery. Recently, another component called SSV has been identified in *Arabidopsis.* SSV looks to modulate the number of starch granules produced in plastids through interacting with PII1/MRC via an α-glucan-binding domain. Although SSV is closely related to SSIV, it lacks glycosyltransferase activity. Mutants deficient in SSV show a reduction in the number of starch granules per chloroplast, but the size of these granules is larger, suggesting that other elements

of the granule initiation apparatus may partially compensate for the absence of SSV (Abt et al. 2020).

The hilum, which is believed to be structurally disorganized, serves as a glucan scaffold that supports further growth of the granule by the enzymes involved in amylopectin biosynthesis. The machinery involved in granule initiation produces linear, unbranched maltodextrins and initiates the formation of the hilum, which serves as a starting point for the growth of the emerging starch granule. The association of SSIV with PTST2, PII1/MRC, and SSV may create a micro-environment that promotes the formation of semi-crystalline α-glucans. Linear MOS can self-assemble into helical coils, which may be the first water-insoluble structures formed and facilitate starch granule initiation (Gidley et Bulpin 1987; Putaux et al. 2006). These structures may be less susceptible to degradation by α- and β-amylases, especially if physically protected by the protein complexes associated with the initiation machinery (Seung et Smith 2019). The branching of linear MOS structures by starch branching enzymes (SBEs) is necessary for granule formation, but it is not known when SBEs act during the hilum structure's formation or extension. Studies suggest that SBEIIa may play a role in granule initiation in barley endosperm (Malinova et al. 2017). Starch phosphorylase (SP) may work with starch synthases (SSs) and starch branching enzymes (SBEs) to provide a starting point for continued hilum expansion and development of the amylopectin backbone and its basic units (Y. Nakamura et al. 2012).

2.1.4. Amylose Synthesis

a. Granule Bound Starch Synthase

Granule bound starch synthase (GBSS) is a single enzyme responsible for amylose biosynthesis in plants and green algae. Its loss of catalytic activity results in *waxy* starch, indicating that no other synthase can replace it in this function (Shure, Wessler, et Fedoroff 1983; S. G. Ball, van de Wal, et Visser 1998). The process of amylose deposition takes place in a preexisting matrix of a starch granule and involves the water-insoluble amylopectin backbone to guide granule-bound starch synthase (GBSS) to the granule. There are two tissue-specific isoforms of GBSS, GBSSI and GBSSII, which are encoded by separate genes and act in different tissues (Tsai, Salamini, et Nelson 1970; Geddes, Greenwood, et Mackenzie 1965). Granule-bound starch synthase (GBSS) incorporates glucose (Glc) from ADP-Glc to the non-reducing end of an α-

1,4-linked glucan chain in a processive manner (K Denyer et al. 1999). Its product is largely linear and is protected from branching activity, explaining why it is linear. GBSS activity is stimulated by MOS and is regulated by protein phosphorylation and oligomerization. The synthesis of amylose necessitates the involvement of a member from the PTST family (Leterrier et al. 2008), PTST1, which targets GBSS to the starch granule via a CBM48 (Seung et al. 2017; 2015). Loss of PTST1 results in a *waxy* phenotype, indicating its critical role in amylose biosynthesis. A proportion of amylose contains α-1,6-branch linkages, but the enzymes responsible for branching amylose are not known.

Amylose is distributed unevenly in the starch granule, being more prominent at the periphery of the granule and in the equatorial fold or groove of developing granules (R. Tester et Morrison 1990). It is likely confined to regions that are more easily accessible to enzymes than the denser, less hydrated crystalline regions. The elongated chains of amylose and the extremely long chains of amylopectin produced by GBSS may contribute to the structural support of the amylopectin backbone within the non-crystalline regions.

2.1.5. Amylopectin Synthesis

Amylopectin is synthesized from the hilum, which results from the initiation process, by three major groups of enzymes: starch synthases (SSs), starch branching enzymes (SBEs), and debranching enzymes (DBEs). These three groups of enzymes functionally and physically interact with each other during the process. They have multiple isoforms within each enzyme class with distinct biochemical properties.

a. Starch Synthases

Three of the six known SS isoforms (SSI, SSII, and SSIII) play a direct role in amylopectin biosynthesis. They produce α-1,4-linked glucan chains of various lengths that can be further extended, branched (by starch branching enzymes, SBEs), or debranched (by debranching enzymes, DBEs) (Pfister et Zeeman 2016). In contrast to granule-bound starch synthase (GBSS), all SS isoforms involved in amylopectin biosynthesis operate in a distributive manner, in which the α-1,4-linked chain is elongated by a single glucose per enzyme-substrate interaction. The possible donor chains for the starch synthase (SS) isoforms could be different and originate from different sources such as pre-existing α-glucans (maltodextrins, MOS) produced by other SS isoforms, debranching enzymes (DBEs) that release branch chains while

trimming amylopectin, disproportionating enzyme (DPE), or starch phosphorylase (SP). SSs have a highly conserved C-terminal catalytic glycosyltransferase domain and a variable N-terminal extension (Figure 2.2) (Pfister et Zeeman 2016). The catalytic domain is conserved between SSs and bacterial glycogen synthases and contains both a GT5 and a GT1 domain (CAZy) (Leterrier et al. 2008).

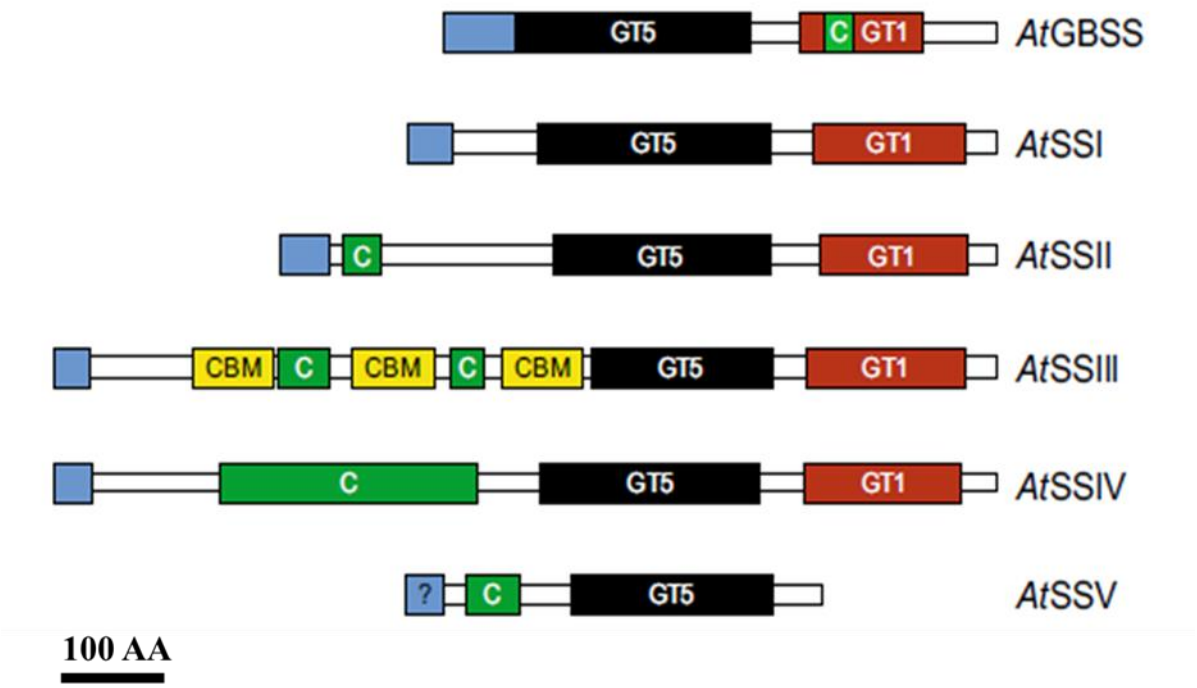


***Figure 2.2:* Domain structure of different Starch Synthases (SS) classes from Arabidopsis.** Different structural domains are highlighted: Plastidial transit peptides (N-terminal blue boxes), carbohydrate-binding modules of Family 25 (yellow boxes, CBM), coiled-coil domains (green boxes, C), glycosyltransferase-5 domains (black boxes, GT5), glycosyltransferase-1 domains (red boxes, GT1). The scale bar represents 100 amino acids (AA) (Pfister et Zeeman 2016).

ADPglucose binding to SSs may involve specific conserved motifs such as Lys-X-Gly-Gly (Furukawa et al. 1990; Edwards et al. 1999) and other charged/polar residues (Buschiazzo et al. 2004; Busi et al. 2008). The N-terminal extensions of different SS classes are diverse. In SSIII and SSIV, these extensions may facilitate protein-protein interactions, potentially via coiled-coil motifs (Hennen-Bierwagen et al. 2008). Additionally, the N-terminal part of SSIII contains three conserved carbohydrate-binding modules (CBMs) that are involved in substrate binding (Busi et al. 2008; Wayllace et al. 2010).

1. Starch synthase I (SSI)

Starch synthase I (SSI) extends the shortest MOS/α-glucan chains, with a degree of polymerization (DP) ranging from 6 to 7, generating intermediate-sized α-glucan chains with a DP between 8 and 12. The exact chain lengths used and produced by different SS isoforms depend on the plant source. These intermediate chains are likely further lengthened by SSII and potentially other SS enzymes. Nonetheless, recombinant *Arabidopsis* SSI can synthesize chains up to DP 15 when using maltoheptaose as a primer (Seung et Smith 2019) and, with the aid of a branching enzyme, can generate the full range of chain lengths typically found in one crystalline layer of amylopectin, such as A and B1 chains (Malinova, Qasim, et al. 2018).

A lack of SSI activity leads to noticeable changes in the chain length distribution (CLD) of amylopectin, especially regarding the A- and B1-chains that form clusters (Fujita et al. 2006). Barley-derived SSI studies reveal that the enzyme does not bind to maltotriose or maltotetraose, suggesting that it utilizes short MOS generated by other enzymes (Wilkens et al. 2014).

2. Starch synthase II (SSII)

Starch synthase II (SSII) deficiency has been investigated in various plants, such as potato tubers (Kossmann et al. 1999), pea seeds (Denver et al. 1995), rice (Umemoto et al. 2002), maize (X. Zhang et al. 2004), and *Arabidopsis* leaves (Pfister et al. 2014). The resulting phenotypes share remarkable similarities, characterized by noticeable changes in the fine of polymerization (DP) around 8 and a decrease in chains with a DP around 18, indicating a shift towards shorter chain lengths. Additionally, *ssII* mutant starches typically have higher amylose content, altered granule morphology, and reduced starch crystallinity. In *Arabidopsis,* small quantities of soluble glucan were found to accumulate alongside starch (Pfister et al. 2014). Based on these changes in chain length distribution (CLD), it seems that the intermediate chains with a DP between 8 and 12, generated by SSI, are elongated by SSII, which subsequently produces longer chains with a DP ranging from 12 to 30.

3. Starch synthase III (SSIII)

The role of Starch synthase III (SSIII) is less well-defined compared to SSI and SSII. It has been proposed to be involved in synthesizing long B chains, extending cluster-filling chains, and regulating other starch-biosynthetic enzymes. Additionally, SSIII is important for initiating

starch granules, especially in the absence of SSIV. Among all SS enzymes, SSIII possesses the longest N-terminal extension, which contains starch-binding domains and predicted coiled-coil domains (Klucinec et Thompson 2002). The most well-understood function of SSIII is the synthesis of long, cluster-spanning B chains such as B2 and B3. Alterations in the short chain profile of amylopectin in ssIII mutants suggest that SSIII also participates in the synthesis of short A and B chains (Q. Lin et al. 2012).

Apart from its function in granule initiation and hilum formation, SSIII also continues to elongate the long glucan chains, initially produced by SSII with a degree of polymerization (DP) between 12 and 30, to create the longest linear chains in amylopectin which can have a degree of polymerization (DP) more than 30.

SSI, SSII, and to a lesser extent SSIII show strong association with starch granules and are resistant to significant removal even after thorough washing with detergents. The mechanism by which granule-associated proteins bind to the granule is not fully understood (Commuri et Keeling 2001). However, for some SS isoforms such as SSI, it's been proposed that their binding affinity for the α-glucans improves with glucan chain length, causing them to become locked onto the substrate (F. Liu et al. 2012). The activities of other SS isoforms and starch branching enzymes (SBEs) may decrease the likelihood of individual enzymes becoming trapped in the granule during starch synthesis.

b. SS Isoforms influence on the Building Block-Backbone Model

The majority of the amylopectin structure is composed of building blocks, which consist of short α-glucan chains extended to form short outer segments. The configuration of these medium-length chains facilitates the formation of helices that contribute to the crystalline characteristic of the 9-nm repeating structure. Starch synthase isoforms SSI and SSII are responsible for creating the short and intermediate-length glucans, which make up the primary structure of amylopectin. This includes the building blocks and short outer segments that are considered part of the clusters in the models proposed by Hizukuri and French. Research in *Arabidopsis* suggests a critical role for functional interactions between starch synthase isoforms SSI and class II starch branching enzymes (SBEs) in establishing the characteristic multimodal chain length distribution found in plant starches (Brust et al. 2014). SS enzymes and SBEs have been shown to physically interact (Tyynelä et Schulman 1993; Tetlow et al. 2008). The long

chains that form the backbone are likely a result of SSIII activity (Martha G James, Denyer, et Myers 2003). In fact, one proposed function of SSIII in amylopectin synthesis, according to other structural models, is the provision of long, cluster-connecting chains that make up the amylopectin backbone (Martha G James, Denyer, et Myers 2003). α-glucan chain groups are abundant in plant amylopectins, and along with the analysis of various SS mutants, this suggests a vital role for SSI, SSII, and SSIII in constructing the building blocks and backbone. Additionally, the backbone may be reinforced by the super-long chains formed by granule-bound starch synthase (GBSS) (Isao Hanashiro et al. 2005; I. Hanashiro et al. 2008).

2.1.6. Starch Branching Enzymes

Branching enzymes (E.C.2.4.1.18) catalyze an irreversible reaction that forms an α-1,6-linked glucan linkage from α-1,4-linked glucan chain. This occurs by hydrolytic cleavage of an α-1,4-glycosidic bond inside the α-glucan chain. The reducing end of the released α-glucan chain is subsequently transferred either to a C6 hydroxyl of the original glucan chain leading to an intra-chain transfer (Drummond, Smith, et Whelan 1972; Borovsky et al. 1979) or to an adjacent α-glucan chain (inter-chain transfer). The factors that determine whether inter- or intra-chain transfer occurs are not fully understood; however, research on potato tuber starch SBE reactions suggests that the relative concentrations of the α-1,4- glucan chains (maltodextrins, MOS) may have a significant influence (Borovsky et al. 1979; Borovsky, Smith, et Whelan 1976).

In starch biosynthesis, branching frequency is limited to about 5%, much lower than in other polyglucans like glycogen (approximately 9%). Branching linkages are restricted to the non-crystalline regions (amorphous regions) of the 9-nm repeat structure in the amylopectin (Brust, Orzechowski, et Fettke 2020). In plants, there are typically two classes of starch branching enzymes (SBEs), categorized based on the relationships of their conserved amino acid sequences: SBEI and SBEII (Figure 2.3) (Larsson et al. 1996; Jobling et al. 1999). SBEI and SBEII are the products of separate genes and have distinct biochemical properties, suggesting specific roles in determining amylopectin structure (Fisher et al. 1996). Typically, the two classes - SBEI and SBEII differ in their substrate specificity and the length of the α-glucan chains they transfer. There are different minimum chain length requirements for SBEI and SBEII classes for branching: approximately DP 15 for SBEI and approximately DP 12 for SBEII (Guan et Preiss 1993; Fisher et al. 1996). SBEI shows a preference for amylose transferring long chains with a degree of polymerization (DP) of about 30, predominantly in

the DP 10 to 13 range. On the other hand, SBEII isoforms show increased catalytic activity when working with amylopectin and transfer relatively shorter α-glucan chains from about DP 6 to 14 (Fisher et al. 1996).

In grasses and cereals, starch branching enzyme class II (SBEII) is subdivided into tissue-specific isoforms: SBEIIa (mainly found in leaves) and SBEIIb (predominantly endosperm-specific), each the product of separate genes and exhibiting different glucan length transfer properties (Fisher et al. 1996; Regina et al. 2005). Typically, SBEI is expressed more predominantly in storage tissues than in leaves and other photosynthetic tissues, suggesting its critical role in determining the structural characteristics of storage starches as opposed to the transient starches (Blauth et al. 2001; Satoh et al. 2003). The absence of SBEI in cereals leads to minor changes in the amylopectin structure, implying some functional overlap between different SBE isoforms. Interestingly, the absence of SBEI in the green alga *Chlamydomonas reinhardtii* results in a decreased in the degradation of transient starch, suggesting that SBEI expression promotes the formation of amylopectin structures that are more susceptible to amylolysis (Tunçay et al. 2013). It's remarkable that not all plants express SBEI, while some plants, such as Canola (*Brassica napus L.*) and *Arabidopsis*, harbor only SBEII class enzymes (Dumez et al. 2006).

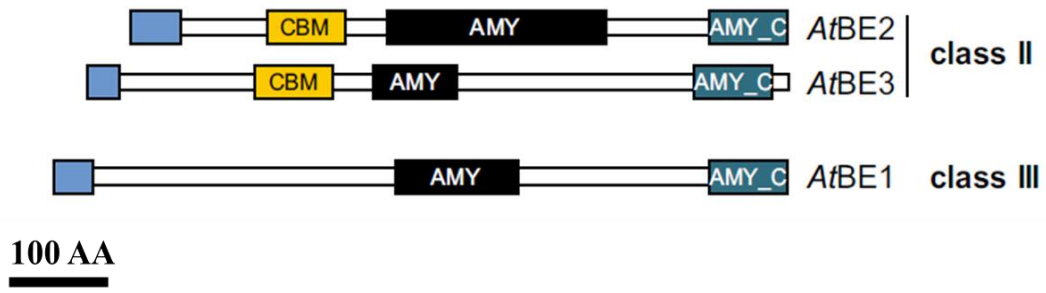


***Figure 2.3*: Domain structure of different classes of branching enzymes (BE).** While these classes can be distinguished by sequence, they share a common domain structure. Different structural domains of these enzymes are highlighted: Plastidial transit peptides (N-terminal blue boxes), carbohydrate-binding modules of type 48 (orange boxes, CBM), catalytic domains of the α-amylase family (black boxes, AMY), and the all-β domains typically found in the C-terminus of α-amylase family members (blue boxes, AMY_C). The scale bar represents 100 amino acids (AA) (Pfister et Zeeman 2016).

SBEI plays a role in forming long branch chains found in amylose and the "super-long" chains of amylopectin. Many studies highlight the significance of the SBEII class in determining the fine structure of amylopectin and affecting overall starch content in various plants (Pfister et Zeeman 2016; Tetlow et Emes 2014). This is not unexpected, considering that the majority of amylopectin structure consists of building blocks and interblock segments, with branching points attached to these structures that are likely formed through the action of SBEII. In many plants such as grasses and cereals, SBEII is responsible for the majority of measurable SBE activity (Regina et al. 2006; Boyer et Preiss 1981), while the absence of SBEII leads to significant changes in the architecture of amylopectin. Loss of SBEII results in an amylopectin with reduced branching frequency, an increase in long chain content (referred to amylose content), and a decrease in total starch content (F. Liu et al. 2009; J. Wang et al. 2018; Jobling et al. 1999; Schwall et al. 2000). Conversely, overexpression of SBEII results in starches with low crystallinity, highlighting the importance of adjusting the rates of α-glucan chain elongation (starch synthases, SS) and branching (Brummell et al. 2015; Tanaka et al. 2004).

2.1.7. Starch Debranching Enzymes

Starch debranching enzymes (DBEs) play a critical function in shaping the water-insoluble properties of amylopectin within starch granules in green algae and higher plants by truncating the amylopectin structure (Cenci et al. 2013; 2014; S. Ball et al. 1996). DBEs belong to the α-amylase superfamily and hydrolyze the α-1,6-branch linkages created by branching enzymes (Møller, Henriksen, et Svensson 2016). During plant and green algae evolution, some DBE isoforms acquired a catabolic function, becoming essential components of the starch biosynthetic pathway. Selective hydrolysis of α-1,6-branch points is thought to play a role in clustering the remaining branches within amylopectin. This clustering promotes local interactions between α-glucan chains and the formation of α-helices, which are likely crucial for the development of the semi-crystalline structure of amylopectin (Gidley et Bulpin 1987; Regina et al. 2004; Mouille et al. 1996; Nielsen, Baunsgaard, et Blennow 2002).

There are two classes of debranching enzymes (DBEs) in plants: the isoamylase (ISA) type and the pullulanase (limit dextrinase) (LDA) type (Figure 2.4). ISA-type DBEs are responsible for debranching amylopectin and other polyglucans, while LDA acts on both amylopectin and the fungal polymer pullulan's α-1,6-linkages (Møller, Henriksen, et Svensson 2016). LDA and an

ISA-type DBE isoform (ISA3) primarily function in removing α-1,6-branch linkages from chains during starch degradation (Wattebled et al. 2008; Yun, Umemoto, et Kawagoe 2011).

During starch biosynthesis, two isoforms of isoamylase (ISA1 and ISA2) trim exposed branches created by SBEs. This trimming results in the formation of clustered branches that are thought to be less accessible to the action of debranching enzymes (DBEs) or amylases such as α- and β-amylases (M. G. James, Robertson, et Myers 1995; Morris et Morris 1939). The function of DBEs in starch biosynthesis has been assumed primarily from mutant analyses, such as the *sugary1* mutant found in sweet corn varieties. This mutant produces a water-soluble polyglucan similar to glycogen there is a polyglucan in plants known as phytoglycogen (M. G. James, Robertson, et Myers 1995; Morris et Morris 1939). Biochemical analysis of *Chlamydomonas* ISA1 supports the role of isoamylase (ISA) in *glucan trimming* and the formation of water-insoluble polyglucan. The enzyme has a low affinity for closely spaced branches, preferentially removing more open and accessible α-1,6-branch points that prevent the clustering of branches necessary for the accumulation of insoluble polyglucan (Dauvillée et al. 2001; Sim et al. 2014). On the other hand, debranching enzymes (DBEs) act by releasing maltodextrin (MOS) chain fragments. These fragments can be used by granule-bound starch synthase (GBSS) for amylose synthesis or by starch phosphorylase (SP) and debranching enzyme (D-enzyme), which may control the availability of MOS for granule initiation (Bresolin et al. 2006). Unlike other key enzymes involved in amylopectin and amylose synthesis, such as starch synthase (SS) and starch branching enzyme (SBE) isoforms, DBEs remain soluble in the plastid stroma rather than becoming physically associated with the growing water-insoluble starch granule.

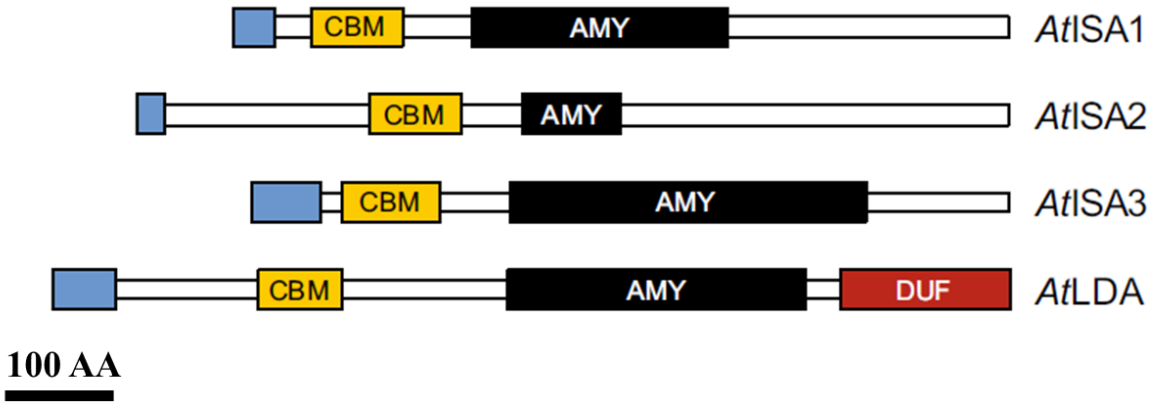


***Figure 2.4*: Domain structure of different classes of debranching enzymes (DBE).** The *Arabidopsis* debranching enzymes (ISA1, ISA2, ISA3, LDA). Different structural domains are highlighted: Plastidial transit peptides (N-terminal blue boxes), carbohydrate-binding modules of family 48 (orange boxes, CBM), catalytic α-amylase family domains (black boxes, AMY), and a Pfam domain of unknown function, DUF3372 (red box, DUF). The scale bar represents 100 amino acids (AA) (Pfister et Zeeman 2016).

DBEs have a critical, albeit indirect role in shaping amylopectin's fine structure. Within the context of the amylopectin cluster model proposed by Hizukuri and French (Hizukuri 1986; French 1972), amylopectin trimming is considered necessary for cluster formation, which optimally structures crystallites (S. Ball et al. 1996). The backbone model proposes that the debranching enzyme (DBE) isoforms involved in amylopectin synthesis, such as ISA1, ISA2, and LDA, act on the building blocks and interblock segments as they are synthesized at the periphery of the growing starch granule by starch synthases (SSs) and starch branching enzymes (SBEs). These DBEs are responsible for removing short outer chains, thereby modifying the structure of the amylopectin molecule. In the backbone model, trimming serves to prevent the formation of overly large, tightly branched entities (i.e., "clusters"), enabling the development of interblock segments, that contribute to the formation of a flexible backbone structure. In addition, the presence of helical segments along the backbone facilitates the alignment and crystallization of double helices, contributing to the overall crystalline structure of amylopectin (Inouchi, Glover, et Fuwa 1987; Boyer et Liu 1983). Initially, a part of the amylopectin molecule is formed with tightly positioned branches. Subsequent trimming by DBEs

facilitates further elongation of the emergent long chain in the growing backbone. This process occurs simultaneously with the formation of double helices from the remaining short chains within the trimmed section (Inouchi, Glover, et Fuwa 1987; Boyer et Liu 1983). As a result, backbone elongation and double-helix formation occur simultaneously, contributing to the overall structure of the starch molecule. This elongated region is continuously trimmed, and more double-helices are formed. In the absence of trimming (lack of ISA1 activity), extensively branched phytoglycogen is produced instead, which is known to lack long-chain fractions (Inouchi, Glover, et Fuwa 1987; Boyer et Liu 1983).

As mentioned earlier, Group 6 building blocks, which consist of approximately 10 chains, may indicate remnants of incomplete truncation of the amylopectin backbone. These building blocks are typically found in small amounts within the amylopectin structure (Eric Bertoft, Koch, et Åman 2012).

2.2. Starch degradation

For starch degradation in leaves, the breaking of glucan polymers is mainly catalyzed by two types of hydrolytic enzymes, *i.e.* exoamylase (β-amylase) that cleaves α-1,4 linkages, and the debranching enzyme isoamylase that cleaves α-1,6 linkages. Moreover, as starch is compacted into semi-crystalline structures, it is essential to decreases the degree of crystallinity to facilitate and speed up degradation allowing a quick turnover. Phosphorylation at the granule surface, probably by interfering with the packing of double helices, facilitates the opening of the granule surface, thus making starch more accessible for hydrolytic enzymes. However, for full degradation of starch, removing the inhibitory phosphates (dephosphorylation) is also required to allow β-amylase to access (Fulton et al. 2008).

Plants rely on stored metabolites to survive and maintain their metabolism at night. Given that starch is the major storage metabolite in leaves, its degradation into glucose and other sugars is essential at night to reach different organs and maintain plant growth (Smith, Zeeman, et Smith 2005). Different enzymes regulate the process of starch degradation through multiple non-linear reactions. First, glucans phosphorylation occurs to facilitate the initiation of starch degradation, whereby the enzymes GWDs and phosphoglucan water dikinases (PWDs) transfer the β-phosphate of ATP to phosphorylate amylopectin glucosyl residues at C6 and C3, respectively (Edner et al. 2007; Ritte et al. 2006). As a result, the semi-crystalline packed structure of glucans

at the surface of granules becomes disrupted, thereby initiating the process of starch degradation (Mahlow et al., 2016).

The breakdown of amylopectin glucans is dependent on the phosphorylation and dephosphorylation reactions (Smirnova et al., 2015). The enzyme Starch EXcess 4 (SEX4) catalyzes the dephosphorylation of amylopectin glucans at C6 and C3 positions (Kötting et al. 2009) whereas and Like Starch-excess Four2 (LSF2) catalyzes the dephosphorylation at C3 position only (Santelia et al. 2011). The dephosphorylation step is crucial for allowing β-amylase to access and break down the glucan chains because the presence of a phosphate group hinders its enzymatic activity (Mahlow et al., 2016). Thus, starch degradation process is achieved through destabilizing the granule surface via phosphorylation, then its degradation by glucan hydrolytic enzymes and dephosphorylation by SEX4 and LSF2 in a process known as reversible glucan phosphorylation (Mahlow et al., 2016; Smirnova et al., 2015; Silver et al., 2014).

Several enzymes have been found to be involved in the starch degradation process. A key enzyme in this process is β-amylase, which catalyzes the cleavage of α-1,4-glycosidic bonds from the non-reducing ends of starch to produce maltose (Lloyd et al., 2005). The enzymes isoamylases (ISAs) are also involved in starch metabolism, where they hydrolyze the α-1,6-branching points in amylopectin to release long linear glucan chains (Mahlow et al., 2016; Smirnova et al., 2015). Other enzymes such as α-amylases and debranching enzymes hydrolyze the α-1,4 and α-1,6-glycosidic bonds of glucans, respectively (Smith et al., 2005). These reactions produce a mixture of linear and branched malto-oligosaccharides. Further degradation of these products can be done by chloroplastic glucan phosphorylase, which releases G1P that can be converted to triose phosphate (Weise et al., 2004). The latter can be exported from the chloroplast in exchange for inorganic phosphate via the triose-phosphate transporter (Walters et al. 2004). In addition, the produced maltose by the action of β-amylase can be transported from the chloroplast into the cytosol through the maltose transporter MALTOSE EXCESS 1 (MEX1) (Lu et al. 2006). Lastly, evidence suggests that DPE1 is implicated in maltotriose metabolism during starch degradation (Critchley et al. 2001). It converts two maltotriose molecules into maltopentaose and glucose, which is exported from the chloroplast to the cytosol through a glucose transporter (O'Neill et Field 2015).

The process of starch degradation is regulated in response to a vast array of factors and conditions, including the circadian clock, CO_2 levels, nutrients, hormone signaling, and water supply (Lloyd et al., 2005). Environmental alterations that cause a decreased level of

accumulated starch storage during the day affect the process of starch degradation at night (Streb et Zeeman 2012). Similarly, changes in pH and redox potential affect the catalytic activity of enzymes implicated in this process (Monroe et al. 2014). These factors are also affected by the onset of light-night transition. Short photoperiods are associated with an increased rate of starch biosynthesis and decreased rate of degradation (Gibon et al. 2009). Likewise, changes in the levels of metabolites may affect the fluxes into and out of the starch degradation process (Keurentjes et al. 2008). In summary, starch degradation is an essential process in plants that allows for the efficient use of stored energy. It is regulated by various enzymes and signaling pathways and is particularly important in plants that experience periods of prolonged darkness.

3. Two new proteins involved in starch metabolism

In order to understand the mechanism by which starch break down is adjusted to the length of the night, an *Arabidopsis thaliana* mutant with a clock-regulating phenotype was studied, allowing the discovery of novel proteins involved in starch metabolism.

Like PTSTs, two new proteins, ESV1 (Early Starvation 1) and LESV (Like Early Starvation 1), with no predicted enzymatic activity but with an important role in starch metabolism in plants, have recently been identified (Feike et al. 2016). The presence of these proteins in all photosynthetic organisms highlights their importance in these organisms. My PhD thesis focuses on the study of these two proteins.

I am co-author of one of the papers describing the function of these proteins, which is currently in press (C. Liu et al. 2023). In that paper, I performed the structural part that I will describe in the results. However, as I was not involved in the functional studies, I have chosen to describe these data in the introductory part of my thesis.

3.1. Early Starvation 1 (ESV1)

3.1.1. ESV1 identification

ESV1 was discovered through a forward genetic screen in a study aimed at investigating the regulation of transitory starch between the day-night cycle, by understanding how *Arabidopsis thaliana* plants modulate leaf starch degradation rates to the length of the night by identifying mutants defective in the regulation of carbon availability that exhaust their starch reserves at night earlier than normal plants in which they identified ESV1 (Feike et al. 2016). In addition, ESV1 was identified for the second time in a study aimed at identifying the genes responsible for the absence of starch granules in the hypocotyl endodermis and root columella in mutants in which they found that the *esv1* mutant is the responsible for the lack of starch granules phenotype (Song et al. 2021).

3.1.2. Phenotype analysis of *esv1* mutants

a. Starch turnover in plants

esv1 mutant is a mutant discovered by Feike et al. 2016, and named early starvation1 (*esv1-1*) in which the plant depleted in this gene has a starvation phenotype (Feike et al. 2016). This mutant displayed strong bioluminescence at the end of the night compared to the wild type, indicating rapid starch degradation at night due to a mutation in an unannotated gene encoding a highly conserved tryptophan-rich protein encoded by At1g42430 found in chloroplasts and within starch granules (Feike et al. 2016; Song et al. 2021). However, this mutant maintained near-normal starch levels at the end of the day.

esv1 mutants displayed abnormal, nonlinear starch turnover during the day and experienced sucrose starvation, leading to the depletion of their starch reserves too quickly at night, approximately two hours before night's end. Consequently, this resulted in early carbon starvation under a normal 12-hour light/12-hour dark cycle despite having normal levels of starch at the end of the day (Figure 3.1.a) (Feike et al. 2016) .

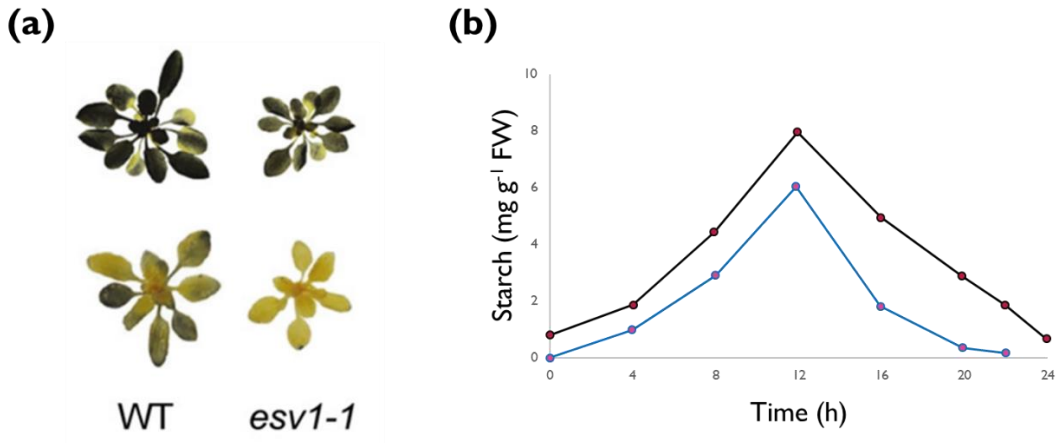


***Figure 3.1:* Rosettes stained with iodine two hours before the end of the night.** Control plant (Col-0) is to the left, while the *esv1-1* mutant obtained from the forward genetic screen is to the right. (a) Wild type rosettes stained pale (on the bottom), indicating that some starch was still present, while esv1 mutant did not stain two hours before the end of the night (at the bottom), indicating a significant decrease in starch content. *esv1-1* mutant had near-normal starch levels at the end of the day (at the top), as shown by the iodine staining. (b) Starch degradation in

leaves of esv1 mutants. The graph shows starch content throughout the day-night cycle in 3-week-old wild-type rosettes (black) and esv1 (blue) (Feike et al. 2016).

In addition, under 8-hour light/16-hour dark cycles, which represent an unexpectedly early night, *esv1* mutants depleted their starch reserves 4 hours before the end of the night indicating an inability to adjust the rate of starch degradation properly Figure 3.1 (Feike et al. 2016). On the other hand, wild-type plants can adapt starch degradation rates based on night length by degrading starch at an almost linear rate and didn't exhaust their reserves before night's end during the 12-hour night or during the early night (8-hour light) (Feike et al. 2016). The time it takes for esv1 mutants to exhaust starch reserves was demonstrated by Feike et al. 2016 to be dependent on their initial starch content when exposed to different light levels. They subjected both esv1 and wild-type plants to varying light levels throughout a single day, resulting in different starch levels at the beginning of the 12-hour night. *esv1* mutant also have lower net starch accumulation during the day meaning low starch content at the start of the night. *esv1* mutant plants have higher sucrose and maltose levels during the day but lower levels at night's end which means that they degraded starch during both day and night. This led to premature starch depletion in *esv1* plants also resulting in less starch accumulation than in wild-type plants. increased maltose levels are a sign of starch breakdown (Weise et al., 2004) while increased sucrose levels are a sign of defects in starch biosynthesis or degradation pathways (Mugford et al. 2014).

ESV1's mutation has an impact on starch granule formation and this impact varies depending on the tissue, and the reasons for this variation are unknown (Song et al. 2021). However, the severity of the ESV1 mutant phenotype may be linked to the balance between starch synthesis and degradation in a given tissue, which in turn determines the formation of starch granules. In *Arabidopsis* sink tissues, starch is synthesized from glucose 6-phosphate in the cytosol and then transported to the plastids via GPT (Kunz et al. 2010), while in photosynthetic source tissues, it is synthesized from fructose 6-phosphate in the Calvin-Benson cycle (Streb et Zeeman 2012). The absence of starch granules in tissues lacking ESV1 coincides with those that use transported glucose 6-phosphate for starch biosynthesis, suggesting a possible function for ESV1 in glucose 6-phosphate transport (Song et al. 2021). However, it is important to note that this information does not necessarily indicate ESV1's involvement in glucose 6-phosphate transport into the plastids (Song et al. 2021). ESV1 may have a role in regulating metabolic processes that are common to all tissues, but the severity of the mutant phenotype may be linked

to the ratio of starch synthesis to degradation in a particular tissue (Song et al. 2021). This ratio determines the balance between starch synthesis and degradation, which in turn affects the formation of starch granules (Song et al. 2021).

b. Phenotype of starch granules

Alteration levels of *esv1* will lead to change in the starch granule morphology, number, or composition, in *esv1* mutants (Feike et al. 2016). The shape of starch granules in *esv1* mutant was less regular compared to those of the wild-type as discussed in chapter 1 and as we can see in Figure 3.2 (Feike et al. 2016). in addition, in esv1 mutant starch had a change in the starch composition of approximately 60% more amylose than wild-type plants (Figure 3.3), but this change doesn't have any effect on starch degradation since the esv1 mutant had a faster degradation of starch even if the amylose is present in the starch or no (Feike et al. 2016).

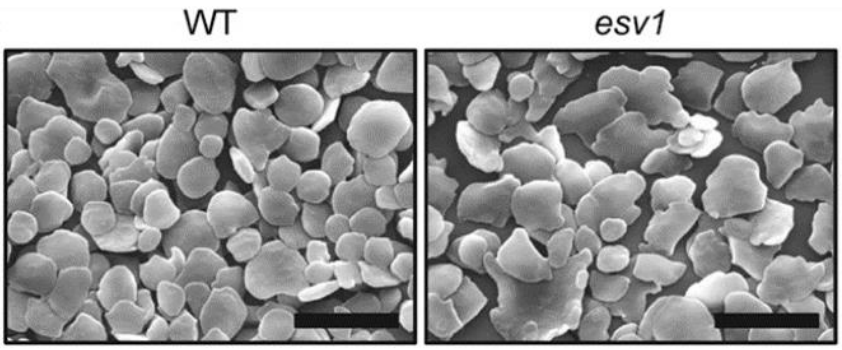


***Figure 3.2:* esv1 mutant starch granule morphology.** Scanning electron micrographs (SEM) of isolated starch granules of wild-type (Col) and esv1; Bars = 2 µm (Feike et al. 2016).

Rice ESV1 is important for accumulating starch in leaves and forming tightly packed starch granules. Mutants lacking ESV1 were found to have small, loosely packed starch granules with crevices between them, whereas wild-type starch granules are large, tightly packed with narrow crevices between them. During the day, rice ESV1 mutants store less starch in leaves compared to the wild type. However, they deplete starch at a slower rate at night, resulting in similar levels of starch depletion at dawn (Song et al. 2021).

3.1.3. Overexpression of ESV1

With the aim of understanding the function of ESV1 protein in regulating leaf starch content, Feike et al. 2016 created plants with increased levels of esv1 protein using constitutive YFP fusion expression. Elevated starch contents were observed in plants expressing ESV1-YFP, regardless of whether they were wild-type or esv1 mutants. Increased starch content was seen only at the end of the night for some lines, whereas in line 3-2 as we can see in Figure 3.3 is the most noticeable effect where starch content was significantly higher throughout the entire day-night cycle (Figure 3.3). It has the highest ESV1-YFP expression (Feike et al. 2016).

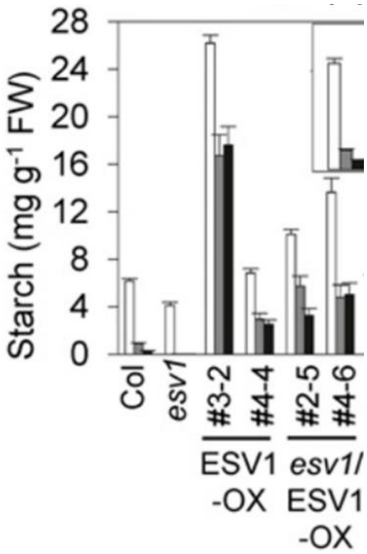


Figure 3.3: ***Starch contents at the end of the day*** represented in white, then 10 hours into the night represented in gray, and the end of the night represented in black were measured for wild-type plants, *esv1*, *lesv* mutants, and lines overexpressing ESV1-YFP in a wild-type or an esv1 background (Feike et al. 2016).

3.2. Like Early Starvation 1 (LESV)

3.2.1. LESV identification

LESV was identified by sequence homology to ESV1 protein encoded by At3g55760 in *Arabidopsis* genome and referred to as **LIKE EARLY STARVATION 1** (LESV). It shares 38% sequence identity with ESV1 in *Arabidopsis* (Feike et al. 2016).

3.2.2. *lesv* mutant phenotype

Feike et al. 2016 studied the effect of *lesv* mutants on starch degradation, in which they found that starch accumulation and degradation were similar to wild-type plants (Figure 3.4). In addition to the levels of sucrose of maltose that are also similar to wild-type plants during the day and the night cycle.

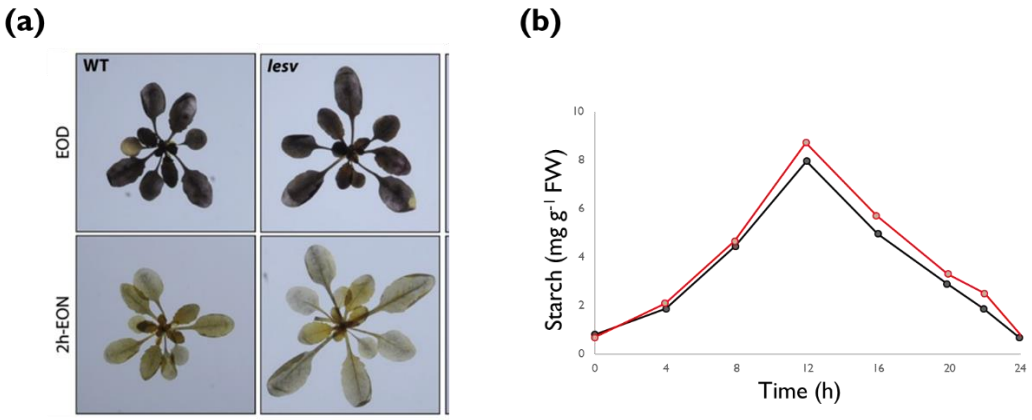


***Figure 3.4:* Starch degradation in leaves of lesv mutant.** Control plant (Col-0) is to the left, while the l*esv* mutant to the right. (a) Wild type rosettes stained pale and *lesv* mutant two hours before the end of the night (on the bottom), indicating that some starch was still present. Wild type rosettes and l*esv* mutant had near-normal starch levels at the end of the day (at the top). (b)The graph shows starch content throughout the day-night cycle in 3-week-old wild-type rosettes (black) and lesv-1 (red) (Feike et al. 2016).

a. Phenotype of starch granules

Feike et al. 2016 studied the alteration levels of LESV protein in terms of starch granule morphology; The shape of starch granules in *lesv* mutant was less regular compared to those of the wild-type (Feike et al. 2016).

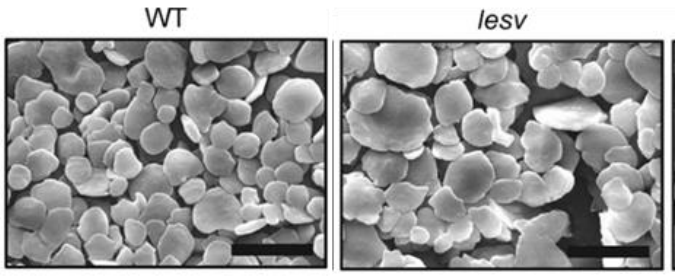


***Figure 3.5:* lesv mutant starch granule morphology.** Scanning electron micrographs (SEM) of isolated starch granules of wild-type (Col), and lesv mutant leaves (Feike et al. 2016).

Changes in the amount of LESV had a greater impact on the levels of starch-bound phosphate. When LESV was absent, starch phosphate content decreased by around 25%, and when it was overexpressed, it increased by about 75% (Feike et al. 2016).

Subsequent to this publication and in view of the results obtained for *lesv* and described in the rest of the manuscript, the phenotype of the *lesv* mutant, which is very close to that of the wild type, did not fully correspond to the expected results because of the important function of lesv. So TEM images have been look more closely and a small fraction of the chloroplasts (5%) that, in fact, contained aberrant granules has been found (C. Liu et al. 2023). In these plastids small granules with irregular surfaces were observed surrounded by very small particles, or just the small particles (Figure 3.6A).

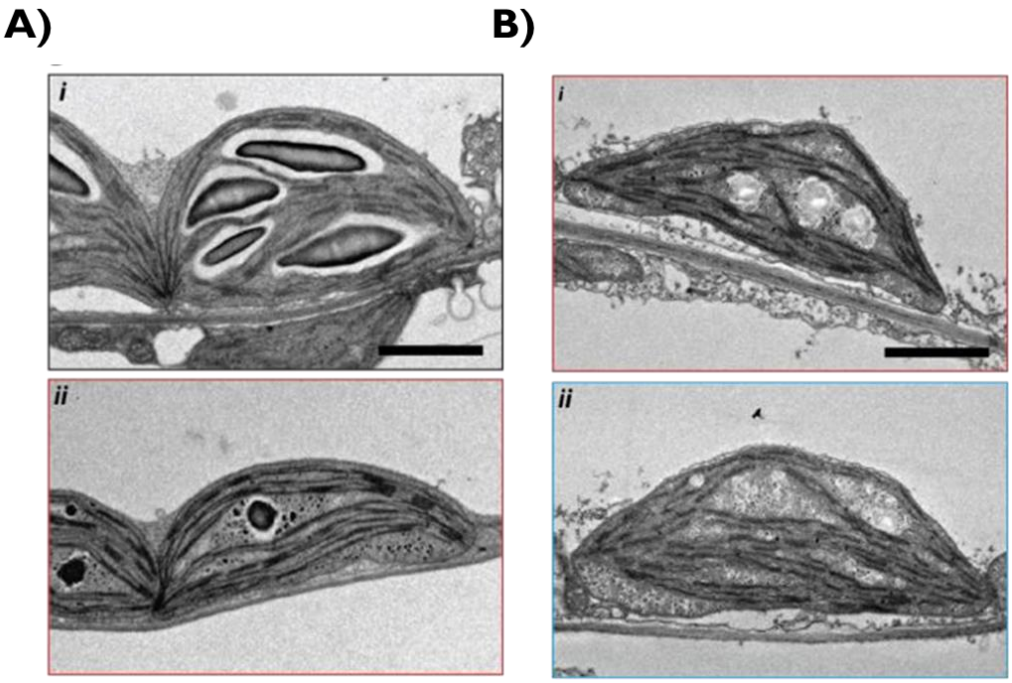


Figure 3.6: ***A) TEMs of selected lesv chloroplast sections, obtained from plants grown.*** Most *lesv* plastids appear wild-type like (*i*), but a few contain unusual glucans (*ii)*. B) TEMs of *lesv* plants grown and harvested. Most chloroplast sections contain either a mixture of starch and phytoglycogen (*i*) or phytoglycogen-like inclusions only (*ii*) Scale bar, 2 μm (C. Liu et al. 2023).

So, it has been decided to de-starch the plants *lesv* with an extended dark period and let them resynthesize the starch from scratch. This yields a spectacular phenotype, with the vast majority of chloroplasts containing a mixture of aberrant granules and substantial amounts of water-soluble polysaccharide. Visualization of these glucans by TEM showed that most *lesv* chloroplast sections contained a mix of aberrant granules and small, presumably soluble particles (58%) or exclusively small particles (37%) (Figure 3.6B). Normal starch was observed in only 5% of chloroplast which shows an important function for LESV in starch granules formation planta that will be described hereafter.

3.2.3. Overexpression of LESV

Overexpression of LESV resulted in a higher number of granules per chloroplast, many of which were much smaller than wild-type granules, and a reduction in the amylose content in

plants. with the aim of understanding the function of ESV1 and LESV proteins in regulating leaf starch content, Feike et al. 2016 created plants with increased levels of LESV protein using constitutive YFP fusion expression. Secondly, they studied the impact of LESV overexpression on starch content. In a wild-type background, LESV-YFP expression had minimal effect on starch content at day's end, but caused a 3-fold decrease in starch content 10 hours into the night and at night's end. This suggested that LESV-overexpressing plants might experience starvation by the end of the night (Feike et al. 2016).

3.3. Localization of ESV1 / LESV

ESV1 and LESV proteins were predicted to be localized in the plastids for the first time using a proteomic approach by (Bayer et al. 2011). However, analysis using TargetP and ChloroP software algorithms did not indicate the presence of a significant chloroplastidial transit peptide (cTP) for the ESV1 protein**,** in addition, similar proteins from other plant species have putative transit peptides. however, this absence was unexpected since most nuclear-encoded proteins require a targeting sequence for localization in plastids (Emanuelsson et al., 1999) . This absence was confirmed by mass spectrometry analysis, which showed that the first 58 amino acids were missing and probably were eliminated during the transportation of the protein into the chloroplast (Malinova et al. 2018).

On the other hand, analysis using the ChloroP program indicates the presence of a chloroplastidial transit peptide of about 56 amino-acid on the N-terminal for LESV protein (Feike et al. 2016).

The chloroplastic location of ESV1 and LESV proteins was later confirmed by Feike et al. 2016 through transient expression of both proteins in the leaves of a plant called woodland tobacco plant (*N. Sylvestris*) by C-terminal fusions to YFP (Feike et al. 2016). In both cases, YFP fluorescence was fully observed in the chloroplasts and also associated with structures expected to be starch granules called discrete bodies as we can see in Figure 3.7.

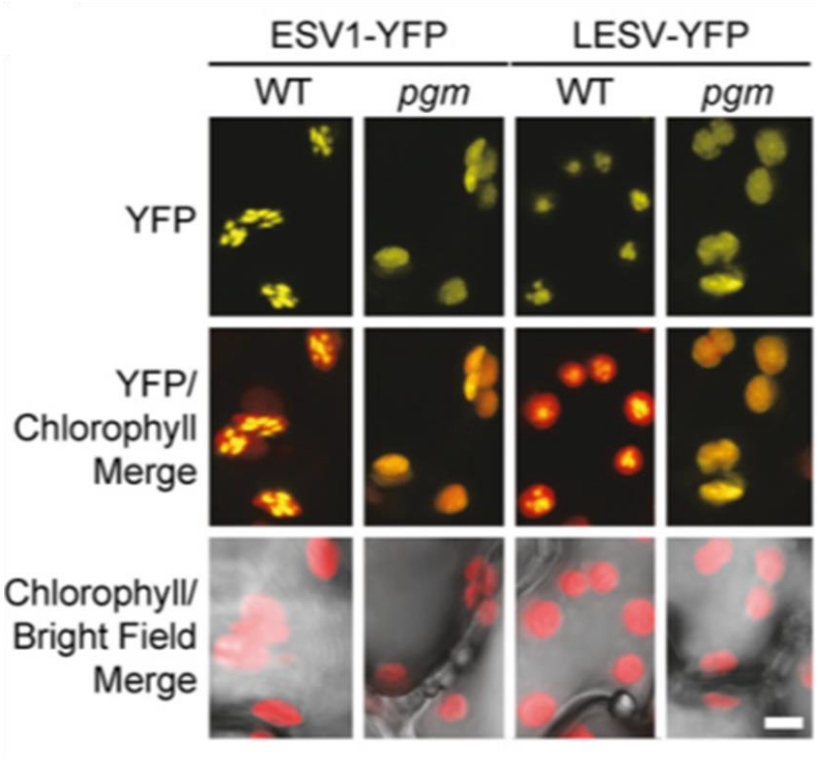


Figure 3.7: ***Transient expression of ESV1-YFP and LESV-YFP in wild-type and pgm mutant N. sylvestris leaves.*** The top panel displays YFP fluorescence, the middle panel shows the combined YFP fluorescence and chlorophyll, and the bottom panel shows only the chlorophyll fluorescence superimposed on a bright field image of the leaf area. All panels share the same magnification. The scale bar at the bottom right represents 5 mm (Feike et al. 2016).

Then, in order to have more information about proteins location within the chloroplast, they expressed the fusions protein in a plant that cannot synthesize starch due to the absence of the chloroplastic phosphoglucomutase enzyme which is the *pgm* mutant of *N. Sylvestris* (Hanson et McHale 1988). *pgm* mutant showed a diffuse YFP signal in the chloroplast stroma, suggesting that the fluorescence in wild-type plants was indeed associated with starch granules. In wild-type *Arabidopsis* leaves, ESV1 and LESV were found in both soluble and insoluble fractions. The proteins were mainly soluble when starch levels were low and insoluble when starch levels were high (Feike et al. 2016). The proteins were predominantly soluble when starch levels were low at night and insoluble when starch levels were high during the day. Additionally, these proteins were present in purified starch from economically significant plants, including *cassava*, potato, rice, and maize (Feike et al. 2016).

3.4. Conservation of ESV1 / LESV

The two non-enzymatic proteins, ESV1 and LESV, are highly conserved across the plant kingdom, including land plants and green algae, while being absent from red algae, prokaryotes and other eukaryotic organisms (Feike et al. 2016). The considerable evolutionary conservation of ESV1 and LESV protein sequences implies that these proteins possess crucial functions exclusive to the Viridiplantae (green plants).

ESV1 protein are conserved in different starch-producing plants such as rice (*Oryza sativa*), liverwort (*Marchantia polymorph*), maize (*Zea mays*) and *Arabidopsis thaliana*, as well as starch-producing unicellular green algae such as *Chlamydomonas* and multicellular green algae like *Chara* (Song et al. 2021). LESV, a lesser-known protein, but equally important protein that shares similarity with the well-studied ESV1 protein. Similar to ESV1, LESV is conserved among diverse plant species, including both land plants and algae (Figure 3.8). The conservation of these two proteins across different plants indicates their fundamental role in starch metabolism (Feike et al. 2016).

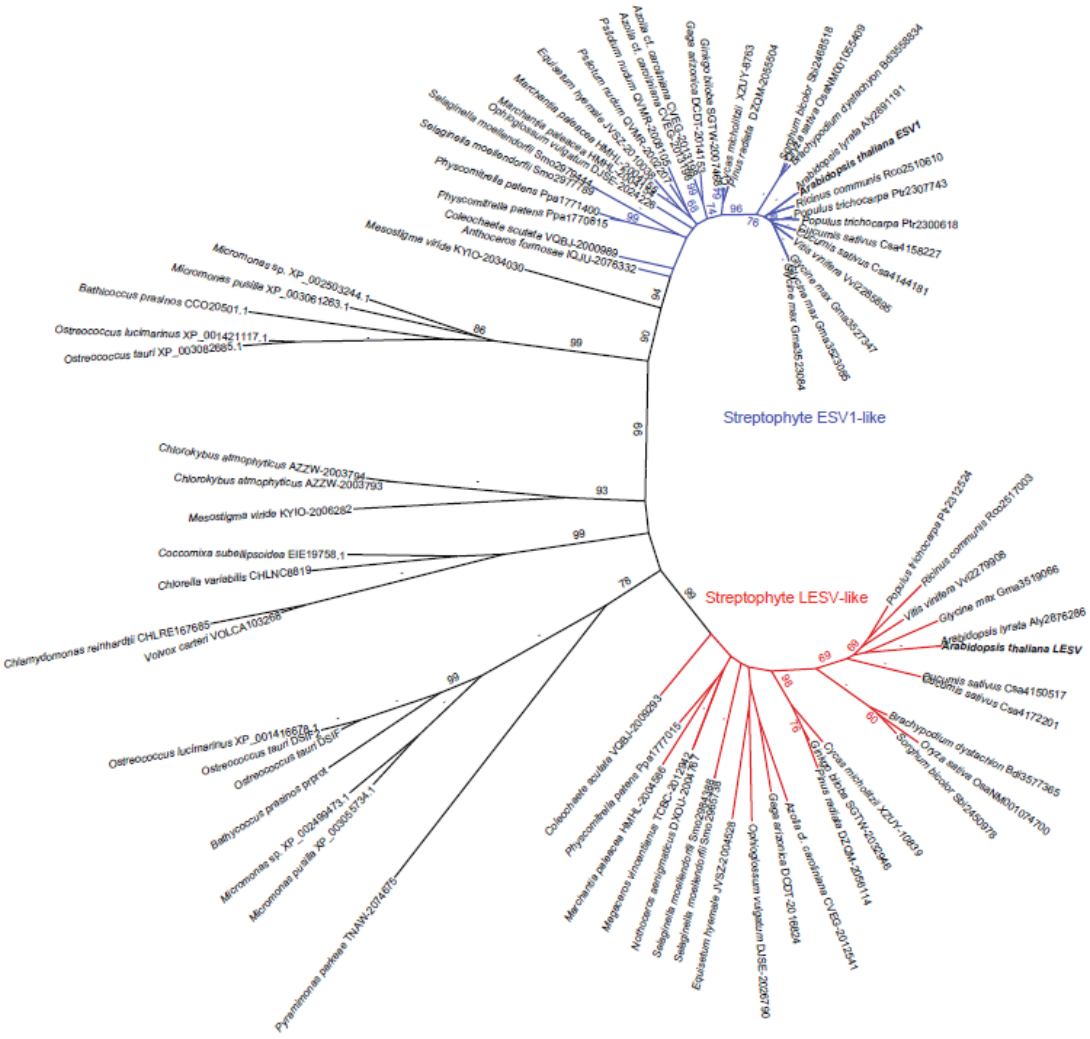


Figure 3.8: ***The conservation of LESV and ESV1 proteins in green algae and land plants*** is examined using a phylogenetic tree. ESV1 protein sequences are shown in blue, while LESV protein sequences are shown in red (Feike et al. 2016).

3.5. Double mutants

Feike et al., 2016 explored if ESV1 can interact with and inhibits a specific enzyme involved in starch degradation by creating double mutant plants lacking both ESV1 and a particular enzyme associated with starch degradation. They looked at how the phenotype of the GWD-deficient *sex1* mutant might change if ESV1 was lost. Glucan water dikinase (GWD) is an enzyme that catalyzes the phosphorylation of the C6-position of the amylopectin's glucose residues, which is the initial step in the breakdown process. In *sex1* mutants, starch is not properly degraded and accumulates in excessive quantities, in addition to having differences in

starch contents of leaves compared to both parents. These starch contents were significantly higher in the esv1 and wild-type plants while significantly lower in the *sex1* mutant (Figure 3.9).

However, Hejazi et al., 2008 suggested that the architecture of the starch polymers is assumed to be disrupted and made easier to degrade by enzymes that hydrolyze starch as a result of GWD's phosphorylation at the granule surface.

Also, starch turnover phenotypes in double mutants *esv1 sex1* were different from their parent plants. Starch contents were substantially higher than those of *esv1* and wild-type plants but much lower than those of the *sex1* mutant.

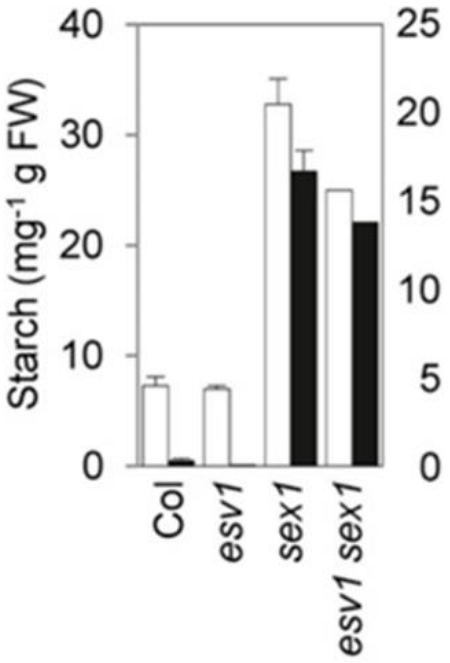


***Figure 3.9:* Starch contents in esv1 double mutants in 4-week-old plants**, specifically in wild-type (Col), esv1-1, sex1, and esv1-1 sex1 double mutants at the end of the day (white bars color) and the end of the night (black bars color) (Feike et al. 2016).

The results showed that the loss of ESV1 modified but did not eliminate the effects of mutations reducing starch degradation. Although the loss of ESV1 diminished the severity of the starch-excess phenotype in all mutant backgrounds, double mutants still retained more starch than esv1 and wild-type plants. This implies that ESV1's role in starch degradation may not be

primarily a negative regulator of any single starch degradation protein. However, additional research is necessary to confirm this.

3.6. ESV1 and LESV Proteins expression in *isa1isa2* mutants in yeast *S. cerevisiae*

The role of ISA1 and ISA2 in the crystallization of amylopectin molecules during starch granules formation has been described as crucial (Delatte et al. 2005; Streb et Zeeman 2014). Given the degraded starch granule phenotype observed in lesv mutants, the latter also appears to play an important role in this process.

To examine the role of ESV1 and LESV in starch biosynthesis within a simplified biological context, these proteins were expressed in *S. cerevisiae* cells that were previously modified to produce starch-like glucans. This yeast lacks the glycogen metabolic machinery and synthesizes starch like glucans (Pfister et al. 2016). Two yeast strains, referred to as lines 28 and 29, were utilized for this purpose. Line 28 contained *Arabidopsis* enzymes SS1 to SS4, BE2, and BE3, resulting in the accumulation of substantial amounts of soluble glucan. In contrast, line 29 additionally harbored the heteromultimeric isoamylase composed of ISA1 and ISA2, leading to a considerable fraction of insoluble and semi crystalline glucans (Streb et al. 2008). It was proposed that ISA activity are required to hydrolyze improperly positioned α-1,6-glucosidic linkages that disrupt the formation and packing of double helices in amylopectin, consequently preventing its crystallization (S. Ball et al. 1996). Line 29 produced high amounts of insoluble glucans in which the regular repeats of crystalline and amorphous layers were less well-ordered and happened with a space of 13.6 nm which is larger than that of starch which is 9 nm. This spacing may possibly occur due to the relative abundance of longer chains (DP>25) in this line (Pfister et Zeeman 2016). Moreover, a significant proportion of the glucans produced by line 29 remained soluble.

Through homologous recombination, sequences encoding ESV1 and LESV (excluding transit peptides) were inserted into the yeast nuclear genome, either individually or in combination. Subsequently, the expression of these proteins in different strains was confirmed using immunoblotting. The levels of accumulated glucans in the strains were assessed by separately measuring the soluble and insoluble pools. Intriguingly, introducing ESV1 into line 28 (referred to as line 559) caused a minor accumulation of insoluble glucan in addition to the soluble glucans typically observed in the parental line (Figure 3.10). This effect became more

pronounced when LESV (line 562) or both proteins (line 563) were expressed simultaneously. Similar results were observed when ESV1 and LESV were expressed in line 29, which already produced insoluble glucans. In this scenario, the supplementary expression of ESV1, LESV, or both (lines 569, 572, and 573, respectively) significantly elevated the levels of insoluble glucans, surpassing those of soluble glucans. Notably, the impact was most prominent when LESV was expressed or when both proteins were simultaneously expressed. In both genetic backgrounds, the increase in insoluble glucans resulting from ESV1/LESV expression coincided with a corresponding decrease in soluble glucans, implying that these two proteins might influence the distribution of glucans between the soluble and insoluble fractions, rather than modifying their overall accumulation.

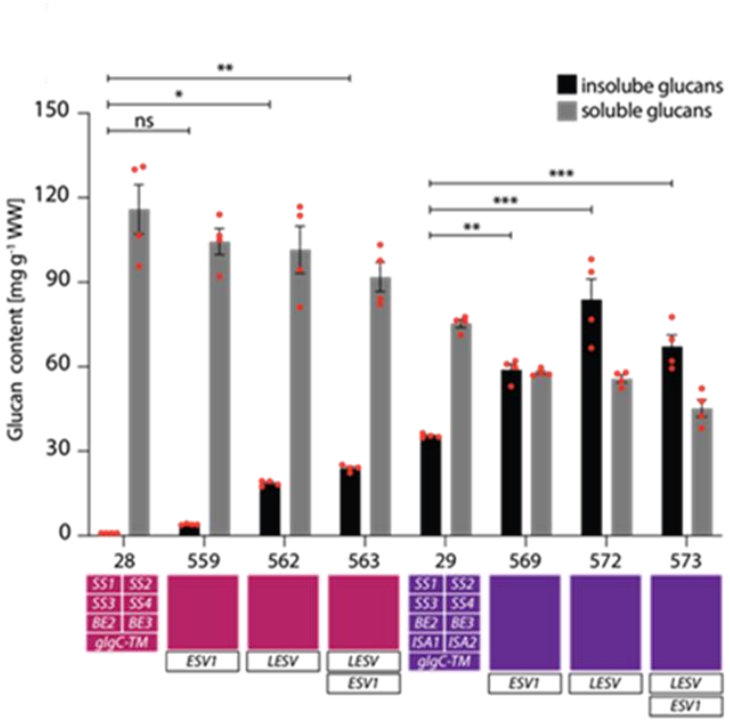


Figure 3.10: ***Insoluble and soluble glucans are accumulated upon the expression of ESV1 and LESV in different yeast strains.*** **(A)** Quantification of soluble glucans (grey bars) and insoluble glucans (black bars) in yeast strains grown in liquid culture with galactose as a C-source in order to induce protein expression (C. Liu et al. 2023).

Glucan accumulation patterns were observed in iodine-stained cells of various strains using light microscopy (LM). Line 28 and Line 559, which also expressed ESV1, displayed a uniform brownish stain. In contrast, lines 562 and 563, which expressed LESV or both ESV1 and LESV

respectively, showed less consistent staining, possibly due to the presence of both insoluble and soluble glucans. Insoluble glucans in cells from line 29 were visible as discrete, darkly stained patches and irregularly shaped particles (C. Liu et al. 2023). Interestingly, in line 569, which expressed ESV1, the patchy deposits appeared larger and more uniform in shape but appeared less darkly stained. This effect was more pronounced in line 572 (expressing LESV), and even more so in line 573. (expressing both ESV1 and LESV) (C. Liu et al. 2023).

To gain further insight into the appearance of the insoluble particles generated in various yeast strains, transmission electron microscopy (TEM) of embedded and fixed yeast cells as well as scanning electron microscopy (SEM) of purified insoluble glucans (Figure 3.11) were used (C. Liu et al. 2023). As previously reported (Pfister et al., 2016), TEMs revealed an abundance of small, glycogen-like particles but no larger particles of insoluble glucan-like particles in line 28's cytoplasm. However, lines 559, 562, and 563 (expressing ESV1, LESV, and both proteins, in the line 28 background, respectively) showed larger, irregularly-shaped particulate matter deposits were frequently visible in TEMs (Figure 3.11). These irregular deposits were surrounded by glycogen-like material like that in the parental line. Using SEM, the insoluble material of these three lines looked as small, aggregated particles (C. Liu et al. 2023).

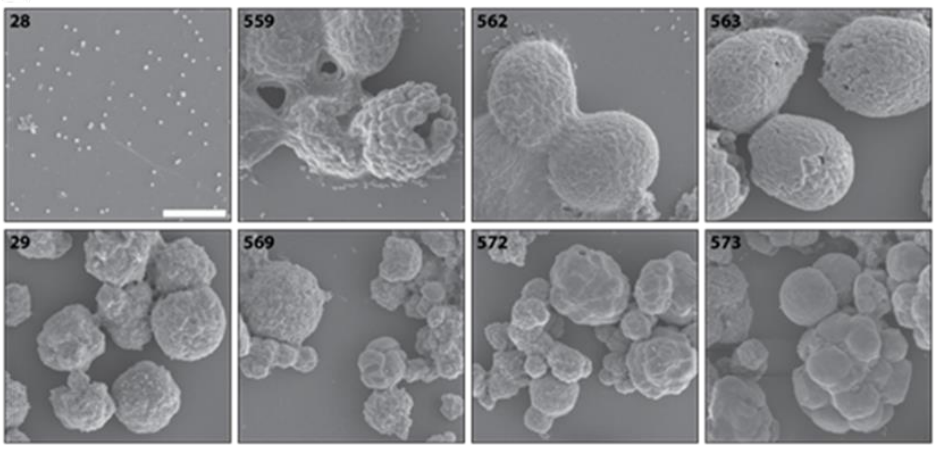


Figure 3.11: Visualization of insoluble and soluble glucans accumulated upon the expression of ESV1 and LESV in different yeast strains. Scanning electron micrographs (SEMs) of purified insoluble particles purified from the specified yeast strains lines (C. Liu et al. 2023).

In line 29, the insoluble glucans produced were visible as large irregular particles in the yeast cell cytoplasm, and upon extraction, appeared as granular structures with rough surfaces under the SEMs (Figure 3.11), consistent with previous findings (Pfister et al. 2016). The additional expression of ESV1, LESV, or both proteins simultaneously which are lines 569, 572, and 573, respectively changed the particles morphology in this background significantly; many of them appeared as aggregates of discrete, larger particles with much smoother surfaces compared to the particles in line 29 (C. Liu et al. 2023).

The effect on particle morphology, regarding size and surface uniformity, was more pronounced with LESV expression compared to ESV1 expression, but the most noticeable change occurred when both proteins were expressed together (C. Liu et al. 2023).

3.7. Starch crystallinity in the *Arabidopsis* isa*1isa2* background

isa1isa2 plant mutants was used with the aim to study the effect of *esv1* and *lesv* mutants or their overexpression on the starch structure and its crystallinity in plants.

The *isa1isa2* mutants produce glucans with a modified structure due to the lack of the debranching step, which is known to facilitate amylopectin crystallization. As a result, a significant portion of the glucan produced remains soluble and does not crystallize (Zeeman et al. 1998; Wattebled et al. 2005; Delatte et al. 2005). This aberrant soluble glucan, which bears a superficial resemblance to mammalian and bacterial glycogen, is referred to as "phytoglycogen" (Dvonch et Whistler 1949).

To investigate the specific function of ESV1 and LESV in glucan crystallinity in *Arabidopsis*, *esv1* and *lesv* mutants were crossed into the *isa1isa2* mutant background, which accumulates significant amounts of phytoglycogen as well as small quantities of starch in its chloroplasts. *Isa1isa2esv1* and *isa1isa2lesv* were chosen as triple mutants. They used Lugol's staining and quantitative glucan measurements to examine glucan metabolism in higher-order mutants and parental lines. By the end of the day (EOD), *isa1isa2* had accumulated a mixture of phytoglycogen and starch, with phytoglycogen being predominant (Figure 3.12A and B). This dominance was reflected in the stained rosettes' reddish-brown appearance, while the wild type, *esv1*, and *lesv* mutants, which all accumulated exclusively starch, displayed a characteristic dark brownish-blue color.

When compared to their respective parental lines, both *isa1isa2esv1* and *isa1isa2lesv* triple mutants accumulated significantly less total glucan at the EOD, with reductions in both starch and phytoglycogen. The relative proportion of phytoglycogen to starch in *isa1isa2lesv* was higher than in *isa1isa2*, while in *isa1isa2esv1*, the proportion shifted slightly towards more starch than phytoglycogen, despite the fact that glucan levels were extremely low (Figure 3.12B).

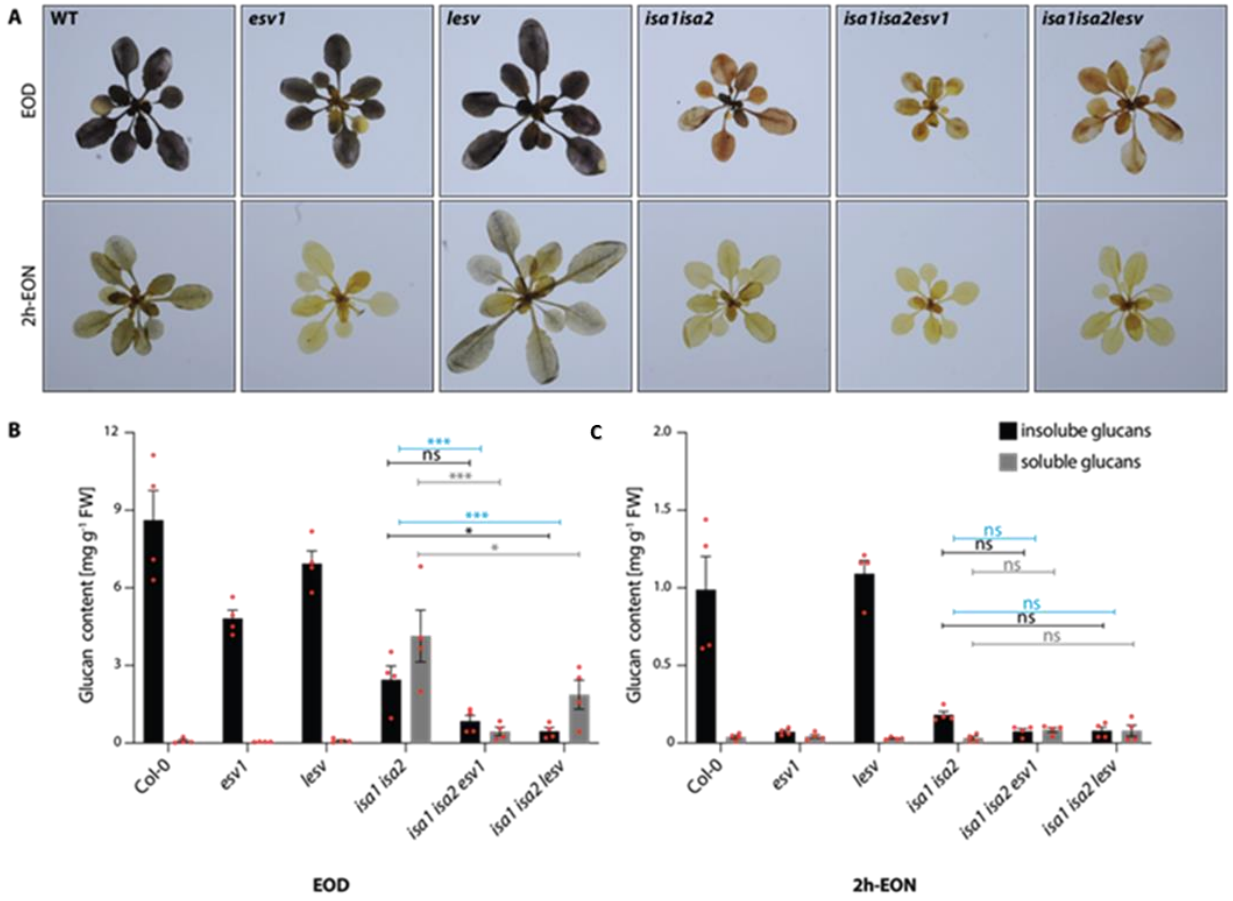


Figure 3.12: ***Starch and phytoglycogen content of isa1isa, esv1, lesv, and isa1isa2esv1, isa1isa2lesv mutants.*** **(A)** Entire rosettes were harvested at the EOD (upper images) and 2h-EON (lower images), and stained for starch with Lugol's solution. **(B)** and **(C)** Quantification of starch and phytoglycogen in plants at the EOD (B) and 2h-EON (C) (C. Liu et al. 2023).

At 2 hours before the end of the night (2h-EON), rosettes of *esv1*, *isa1isa2*, *isa1isa2esv1*, and *isa1isa2lesv* were hardly stained (Figure 3.12A). This suggests that these lines had prematurely

depleted their glucan reserves, possibly displaying an early-starvation phenotype similar to *esv1*.

They examined the presence of glucans in leaf mesophyll, epidermal, and bundle sheath cells (C. Liu et al. 2023). They found that phytoglycogen-like particles predominated in the mesophyll chloroplasts of *isa1isa2*, with few larger starch-like particles present. *Isa1isa2* epidermal and bundle sheath cell plastids contained only starch granules. There was less glucan visible in the mesophyll cells of isa1isa2esv1, but some solid starch-like particles remained. There are no starch granules or phytoglycogen in the epidermal cells. However, with *isa1isa2lesv* a different result found that the plastids in the mesophyll, epidermal, and bundle sheath cells all appeared to contain only phytoglycogen.

3.8. ESV1 or LESV Overexpression enhances starch accumulation in *Arabidopsis* isa mutants

Liu et al. 2023 investigated the function of ESV1 and LESV by altering their expression in the *isa1isa2* double mutant background, in order to study their effects on glucan crystallinity. Transgenic *Arabidopsis* lines overexpressing YFP-tagged ESV1 or LESV were crossed with *isa1isa2*, and the resulting lines were characterized with respect to their starch content and phytoglycogen content at the end of the day (EOD). When stained with Lugol's solution, ESV1-OX rosettes appeared darker blue (sign of high amylose) than wild-type rosettes, whereas LESV-OX stained equally to the wild type (Figure 3.13A). TEM microscopy showed that ESV1-OX had bigger granules than the wild type, whereas LESV-OX accumulated high number of small starch granules with variable appearance. Starch quantification revealed a significant starch excess in ESV1-OX and wild-type levels in LESV-OX.

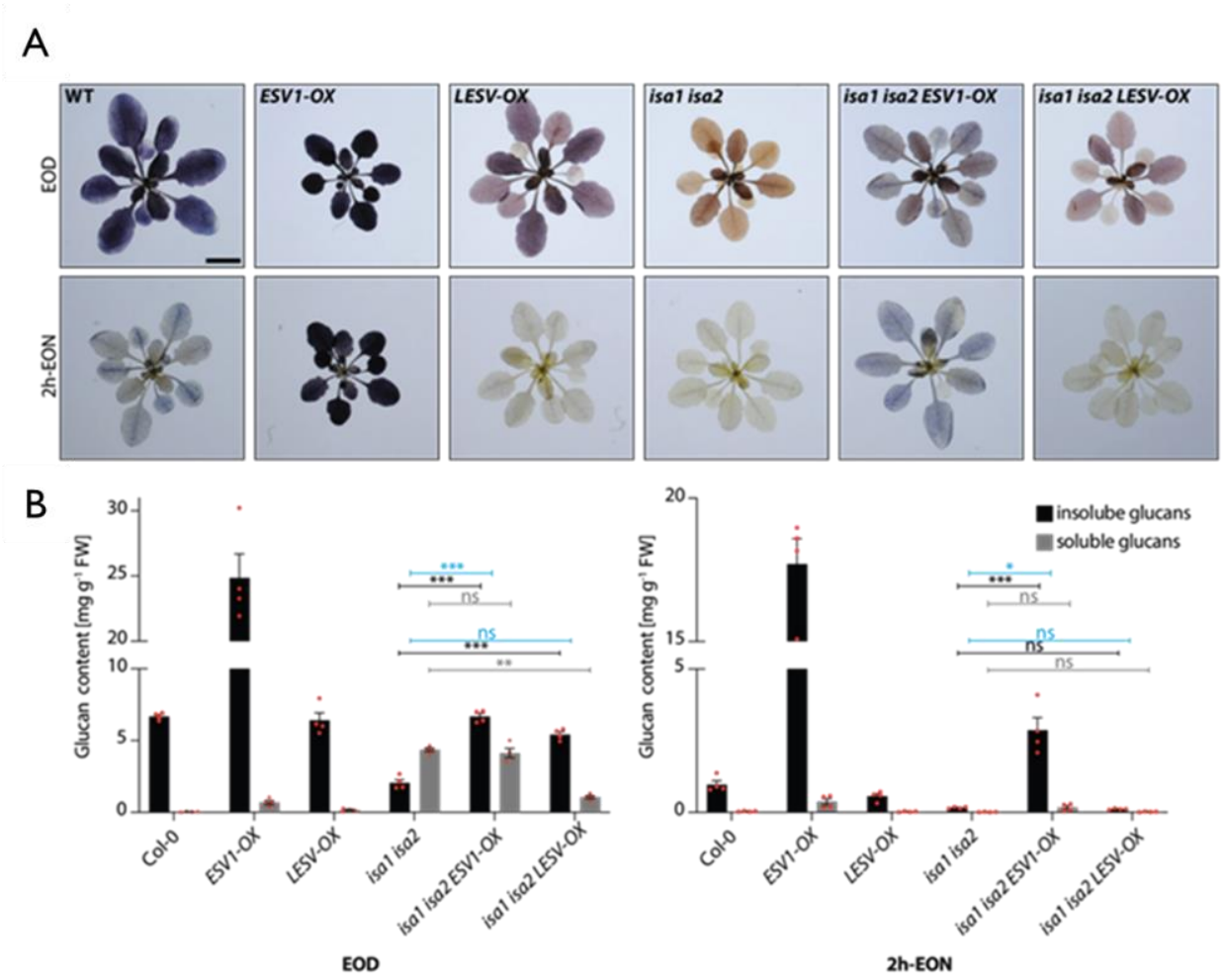


Figure 3.13: ***Glucan accumulation and turnover in plants lines overexpressing LESV and ESV1 in the isa1isa2 background.*** **(A)** Entire *Arabidopsis* rosettes were harvested at the EOD and 2h-EON colorized with Lugol's solution. **(B)** Quantification of starch and phytoglycogen in plants harvested at the EOD and 2h-EON (C. Liu et al. 2023).

Both the *isa1isa2*ESV1-OX and *isa1isa2*LESV-OX lines stained brown-red, indicating a mixture of starch comparable to the wild type and phytoglycogen like the *isa1isa2* line. Quantitative glucan measurements revealed the presence of similar amounts of soluble glucan in *isa1isa2ESV1-OX* as in *isa1isa2*, but significantly less in *isa1isa2LESV-OX* indicating that both *isa1isa2*ESV1-OX and *isa1isa2*LESV-OX lines displayed different phenotypes compared to isa1isa2 (Figure 3.13C). Nevertheless, in both lines, the starch fraction was significantly higher than in *isa1isa2*, predominating over the phytoglycogen (Figure 3.13B). In contrast, *isa1isa2*LESV-OX exhibited wild-type starch levels and very little phytoglycogen.

These findings indicate that ESV1 appears to primarily limit starch degradation, whereas LESV appears to promote starch formation instead of phytoglycogen, without hindering degradation. Under confocal microscopy, the distribution of YFP-tagged proteins showed patterns consistent with starch granule localization.

3.9. Biochemical characteristics of ESV1 and LESV proteins

In *Arabidopsis* leaves, ESV1 protein has a molecular weight of approximately 49 kDa (the product of the At1g42430 gene) and is made up of 426 amino acids, with no previously annotated regions (catalytic domain or carbohydrate-binding domain, or even protein-protein interaction domain). It possesses a proline-rich region at the C-terminal extremity consisting of 11 proline residues between amino acids 397 and 425. It is also characterized by a high abundance of tryptophan residues (W) residues and other aromatic amino acids in the C-terminal two-third end of the sequence (Figure 3.14) (Feike et al. 2016). Around 11% of the amino acids between amino acids 130 and 380 are tryptophan residues. This region contains conserved motifs of 5 to 9 residues in which aromatic residues are separated mostly by acidic residues (E or D). However, these conserved residues may form several Sugar Binding Sites (SBS) allowing the binding of the protein to the long glucanic chains of starch (Feike et al. 2016).

LESV protein has a molecular weight of approximately 66 kDa and is made up of 578 amino acids (Feike et al. 2016). It shares the tryptophan-rich region containing acidic and aromatic motifs with ESV1 (Figure S1, Liu et al 2023) but lacks the proline-rich C- terminal regions. Although its N-terminal region is different from ESV1 (Feike et al. 2016). The highest level of identity between ESV1 and LESV proteins is within the tryptophan region in which they have similar numbers of tryptophan residues (Figure 3.14) (Feike et al. 2016).

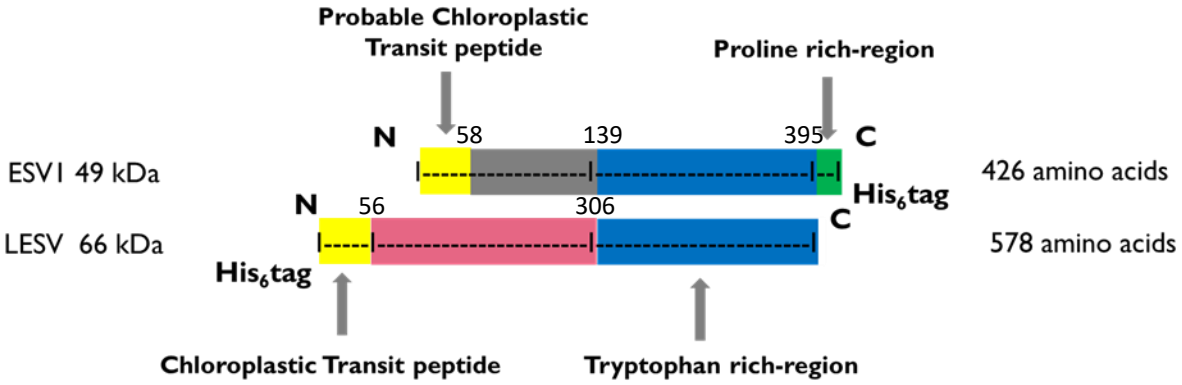


Figure 3.14: ***Schematic representation of the conserved structures of ESV1 and LESV.*** The proline-rich region, W-rich region, and transit peptide are colored green, blue, and yellow respectively.

3.10. Distinct roles of ESV1 & LESV in supporting starch granule formation

The function of ESV1 was investigated in plant systems (Feike et al. 2016), where mutations that affect its expression or overexpression have distinct phenotypic effects (Feike et al. 2016). ESV1 may function in the formation of starch granules, in which the conservation of this function was demonstrated by the ability of ESV1 orthologs to restore the phenotype of an *esv1* mutant by expressing the orthologs of ESV1 protein from other species that are capable of rescuing the production of starch granules, indicating that the ESV1 is a functionally conserved protein among many organisms and required for starch granules formation (Song et al. 2021). The primary function of ESV1 is not believed to be involved in the regulation of starch degradation, it has been suggested that ESV1 is likely involved in starch degradation at an earlier stage (before the beginning of starch degradation) (Malinova, Mahto, et al. 2018). Furthermore, ESV1 prefers binding to highly ordered starch-like glucan structures. This protein affects the activity of GWD on the surface of the starch granule, which, in turn, impacts the function of PWD as well (Malinova, Mahto, et al. 2018) . ESV1 plays a critical role in starch degradation by affecting the phosphorylation of starch via GWD and PWD in an antagonist manner. It reduces the activity of GWD, leading to a decrease in phosphorylation and an increase in PWD-mediated phosphorylation, likely due to prephosphorylation by GWD (Malinova, Mahto, et al. 2018). The absence of ESV1 in knock-out mutants results in an increase in the degradation of starches during the dark phase (Feike et al. 2016). The expected

higher GWD-mediated phosphorylation in the absence of ESV1 would lead to more rapid starch degradation. Furthermore, the influence of ESV1 is rather specific for the dikinases as it alters the structure of the starch granule surface specifically and differently used by the two starch-related dikinases downstream of GWD and PWD (Malinova, Mahto, et al. 2018). However, ESV1 does not directly influence the phosphorylation activity of GWD and PWD, nor does it significantly alter their binding to *Arabidopsis* wild-type starch (Malinova, Mahto, et al. 2018). However, mutation or overexpression of ESV1 in *Arabidopsis* resulted in only minor alterations in the amount of phosphate bound to starch (Feike et al. 2016). it was proposed that ESV1 may play a role in organizing the starch matrix, which affects the resistance of glucan chains to degrading enzymes (Feike et al. 2016).

The function of LESV was examined in plant systems (Feike et al. 2016). LESV is expressed across all plant organs. In leaves, its transcript levels are elevated at the end of the night and reduced during a significant portion of the day (Feike et al. 2016). It exhibits strong co-expression with numerous genes responsible for encoding starch metabolism enzymes, including Glucan water dikinase (GWD), starch branching enzyme 3 (BE3), and alpha-amylase 3 (AMY3) (Feike et al. 2016). In addition, LESV expression in *Arabidopsis esv1* mutant doesn't rescue the production of the starch granules like in ESV1 case (Song et al. 2021), and does not show the same phenotype as the *esv1* mutant, indicating that LESV doesn't required for starch granules formation and that LESV has different function than ESV1 protein (Song et al. 2021).

Prior to the study by Feike et al. in 2016, it was generally believed that starch chains synthesized on the surface of granules *in vivo* assemble into double helices and self-assemble by physical processes to form the semi crystalline lamellar structure of the starch matrix (Waigh et al. 2000). However, the discovery of the ESV1 and LESV proteins introduced a new hypothesis, suggesting that the proper assembly of the starch matrix involves not only physical processes, but also the involvement of proteins that bind to starch polymers, as shown by (Feike et al. 2016).

a. LESV may be a potential organizer of starch granule matrix.

In yeast cells previously engineered to produce starch-like glucans, LESV was found to enhance the accumulation and the transition of glucans into insoluble glucans in addition to affecting its morphology in yeast cells by resulting in larger and more regularly shaped starch granules.

LESV can reduce the accumulation of short maltooligosaccharides (Pfister et al. 2016), suggesting that LESV may influence the distribution of glucans (C. Liu et al. 2023), indicating a greater proportion of the glucan transitioned into an insoluble form with a lower degree of debranching (C. Liu et al. 2023). Further experiments demonstrated that the process of insoluble starch granules formation may be impaired in the lesv mutant, as evidenced by the accumulation of soluble glucans and short chains after a prolonged night treatment, which indicates that LESV plays an essential role in facilitating the formation of insoluble starch granules (C. Liu et al. 2023).

b. ESV1 may act as a protective shield of amylopectin structure against degrading enzymes.

ESV1's function is to stabilize amylopectin within the starch granule, rather than to promote its crystallization. ESV1 expression in yeast cells results in a smaller effect on promoting glucan phase transition compared to LESV in yeast cells. However, ESV1 overexpression in plants increases the amount of starch with larger granules compared to the wild type (Feike et al., 2016), and this increase come from blocking its degradation at night (C. Liu et al. 2023) rather than by decreasing the amount of soluble phytoglycogen which remains similar to the wild type (Feike et al. 2016). As previously mentioned, the esv1 mutation causes the starch to be broken down too quickly at night (Feike et al. 2016), indicating that ESV1 may be involved in stabilizing some weak points in the semi-crystalline structure. The surface of starch granules is not uniform, and certain areas may be more susceptible to degradation, such as where a crystalline lamella ends and a new one begins. These areas may have less stable double helices. ESV1 may recognize and stabilize these weak points, and without it, uncontrolled degradation can occur. On the other hand, overexpression of ESV1 could make the granule too stable and interfere with phosphorylation. (Malinova et al. 2018). As a conclusion, the amylopectin biosynthetic enzymes generate a glucan product with a tendency to crystallize, facilitated by LESV and subsequently stabilized by ESV1.

3.11. Mechanism of action

Liu et al., 2023 studies found that these two proteins greatly enhance the transition of branched amylopectin from a soluble to a pseudocrystalline state during starch granule synthesis,

indicating that they significantly improve the process. LESV is particularly important for promoting the nucleation of granules during starch biosynthesis, while ESV1 functions later on by stabilizing them (C. Liu et al. 2023). This new understanding challenges the previous assumption that the formation of starch granules was a spontaneous biophysical process (C. Liu et al. 2023).

All of these results suggested that starch granules can clearly form in the absence of LESV, also it promotes this process by seeding the glucan phase transition (C. Liu et al. 2023). They proposed that the tryptophan-rich domain acts as a template for aligning double helices that spontaneously happen between neighboring amylopectin chains (Figure 3.15). Each strip of aromatic amino acids on each side of the β-sheet domain could bind one double helix, and that once properly aligned on LESV, these double helices could then align others, resulting in a crystalline lamella.

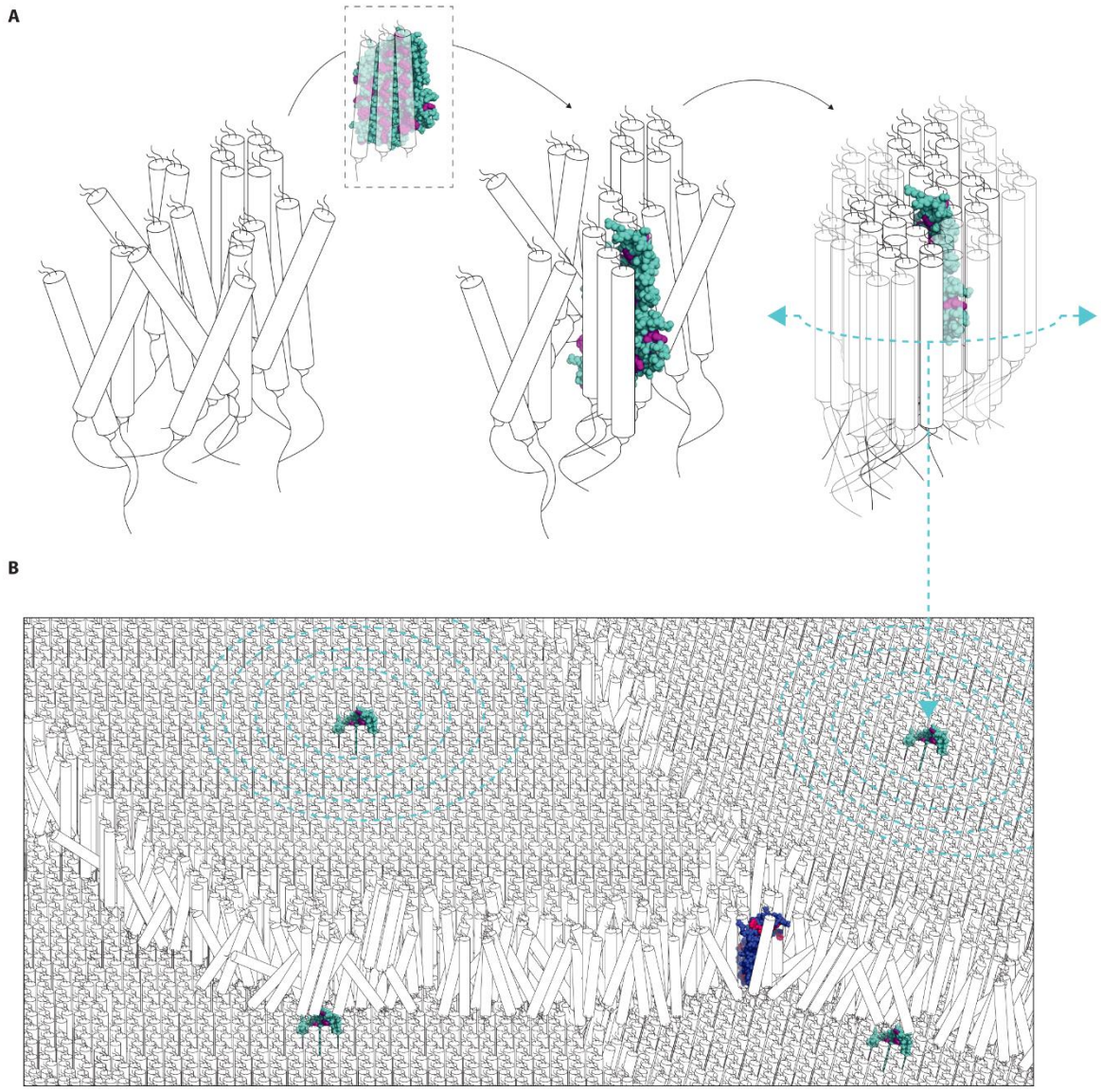


Figure 3.15: ***Proposed functions of ESV1 and LESV are illustrated in this model.*** **(A)** LESV interacts with glucans with helical secondary structure via its Trp-rich domain and, possibly assisted by its N-terminal domain (not shown), facilitating their arrangement into compact, ordered tertiary structures. Regular glucan arrangements can self-produce and spread once seeded. **(B)** ESV1 binds to and stabilizes exposed helices, preventing hydrolytic activities. It should be noted that the displayed areas of ordered glucan are most likely underrepresented, and the proteins to glucans ratio in starch would be lower. ESV1 and LESV proteins are represented as simplified versions of their predicted AlphaFold structures in both A and B, (only Trp-rich region shown; aromatic residues are highlighted by color; pLDDT values are disregarded) (C. Liu et al. 2023).

Other parts of the LESV protein, such as the conserved -helical domains or regions with additional conserved aromatic residues, may also be required for this seeding function. Many observations could be explained by this model. In cases, where the glucan branching pattern is suboptimal (e.g., in the absence of the trimming isoamylase), the likelihood of spontaneous self-alignment of double helices to seed a crystalline structure is reduced, but the presence of LESV template may allow it to happen and thus LESV facilitate crystallization. When branching is optimal, however, crystallization can occur spontaneously even in the absence of LESV template (C. Liu et al. 2023).

Objectives

The objective of my PhD project is to elucidate the molecular functions of the ESV1 and LESV proteins, which have recently been identified as being involved in starch metabolism in the model plant *Arabidopsis thaliana* by structural approaches. The widespread conservation of both proteins across the plant kingdom, along with their presence in starch particles from a variety of sources, suggests a critical role in plant physiology.

Since the precise assembly of the matrix may involve proteins binding to starch polymers in addition to physical processes, my goal is to investigate these phenomena. I intend to achieve this through a strategy that combines structural and functional studies as follows:

The aim of this Ph.D. project is threefold:

1. To validate the predicted models of ESV1 and LESV proteins involved in starch metabolism in the model plant *Arabidopsis thaliana* using biophysical techniques such as small-angle X-ray scattering (SAXS) and circular dichroism (CD) to understand their overall architecture and conformation.
2. To study the functional interactions between these proteins and glucans, such as amylopectin, using an interdisciplinary approach combining biochemical and biophysical methods to uncover their functions, reveal any structural rearrangements, and identify the molecular interactions involved.
3. To decipher the 3D structures of ESV1 and LESV proteins using X-ray crystallography to obtain a detailed structural understanding that may provide further insights into their role in starch metabolism throughout the plant kingdom.

4. Materials and Methods

4.1. ESV1 protein

4.1.1. ESV1 protein construct

a. DNA construction and bacterial strain

The ESV1 protein gene has been synthesized with codons optimized for expression in E. coli and inserted into the pET24a+ plasmid (Novagen). This plasmid allows the insertion of a 6His-tag at the C-terminus of the ESV1 protein, and is selected by its resistance to kanamycin and chloramphenicol. In the ESV1 construct, the coding sequence for the first 95 residues at the N-terminus has been deleted, which corresponds to a region with low conservation in orthologous sequence alignments. This modified construct is referred to as pET24a+ AtESV1 -Δ95. This construct was provided by our collaborator, Zeeman's lab (C. Liu et al. 2023).

4.1.2. ESV1 expression and purification protocol

a. Cell culture and Induction

The pET24a+ AtESV1 -Δ95 construct was used to transform the bacterial strain BL21 codon+ (Agilent) for overexpression, which is one of the most widely used hosts for protein expression and has the advantage of being deficient in both *lon* and *ompT* proteases. This strain was selected using kanamycin and chloramphenicol. Transformed bacteria were grown on sterile LB agar plates containing Kanamycin and chloramphenicol at concentrations of 50mg/mL and 35 mg/mL respectively.

For an overnight pre-culture, bacteria were grown in 100ml of Lysogeny Broth (LB) medium supplemented with kanamycin (50 mg/ml) and chloramphenicol (35mg/ml), with shaking at 180rpm overnight at 37°C.

A 1L volume of LB liquid broth, supplemented with kanamycin (50 mg/ml) and chloramphenicol (35mg/ml), was then inoculated with 50ml of the overnight pre-culture. The culture was incubated with agitation at 37°C until an optical density of 0.5 - 0.6 (A_{600}) was reached using a SPECTROstarNano. Protein expression was induced with 1mM isopropyl β-D-1-thiogalactopyranoside (IPTG) for 2- 3 hours at 37 °C, followed by overnight incubation at 20°C.

b. Protein purification

Cultures were centrifuged at (8,000 × *g*) for 20 minutes at 4°C, cells were collected and re-suspended in 50mM Tris buffer pH 7.5 with 300mM Nacl, 40mM Imidazole, 1/200 DNase, 1/100 $MgSO_4$, complete protease inhibitor cocktail EDTA-free tablets [one tablet per liter of culture] (Roche) at 4°C, and were disrupted through an emulsiflex high-pressure homogenizer at 1500 bars. Soluble bacterial extract was then collected by centrifugation (10,000 × *g*) for 1 hour at 4 °C.

The extract was subjected to two purification steps using the ÄKTA™ go purification system (Cytiva) (Figure 4.1). The first purification step using an Immobilized metal affinity chromatography (IMAC) using His-Trap FPLC 5 ml column (Cytiva) equilibrated in 50mM Tris buffer pH 8, 300mM Nacl, 40mM Imidazole. Recombinant protein was then eluted using 250mM Imidazole in the equilibrium buffer and 1.5 mL fractions are collected in 96-well plate.

Figure 4.1: ÄKTA™ go purification system (Cytiva)

This step was followed by a second purification step by size exclusion chromatography (SEC) step using a Superose™ 6 Increase 10/300 GL (Cytiva) pre-equilibrated with 50mM Tris pH 7.5, 100mM Nacl, 10% (v/v) glycerol, 2mM DTT (w/v) at 4°C. Recombinant protein was then eluted using the same equilibrium buffer 50mM Tris pH 7.5, 100mM Nacl, 10% (v/v) glycerol, 2mM DTT (w/v) and 2 mL fractions are collected in 96-well plate. Different fractions collected throughout the purification process are subsequently analyzed using sodium dodecyl sulfate-polyacrylamide gel electrophoresis (SDS-PAGE) to identify the fractions that contain the ESV1

protein. For structural study, protein samples were concentrated using Vivaspin centrifugal falcon with a 10 kDa cut-off (sartorius). Protein concentrations were determined using a Nanodrop Spectrophotometer (ND1000) from Thermo Scientific. The monodispersity of the obtained protein solution has been assessed by dynamic light scattering (DLS) by our collaborator in Zurich.

4.2. LESV Protein

4.2.1. LESV protein construct

a. DNA construction and bacterial strain

The gene encoding the LESV protein has been synthesized with codons optimized for production by *E. coli* and inserted into the pET28a+ plasmid (Novagen), resulting in the construct pET28a-AtLESV. The pET28a+ plasmid allows the insertion of a 6xHis-tag at the N-terminus of the protein, followed by a thrombin cleavage site and is selected for its resistance to kanamycin. In the LESV construct, the sequence encoding the 56 residues representing the N-terminal transit peptide was truncated. giving rise to the construct called pET28a+ AtLESV. This construct was used to transform the bacterial strain Rosetta (DE3) pLysS (Novagen).

4.2.2. LESV expression and purification protocol

a. Cell culture and Induction

The pET28a+ AtLESV construct was utilized to transform the bacterial strain Rosetta (DE3) pLysS (Novagen) for overexpression, which was previously selected among several species because it allows the best production of the protein. This strain contains the pRARE plasmid, which is selected by Chloramphenicol allowing the expression of tRNAs specific for rare codons in bacteria. The strain was selected using kanamycin. Transformed bacteria were grown on sterile LB agar plates containing Kanamycin at a concentration of 50mg/mL.

For an overnight pre-culture, bacteria were incubated in 20ml of LB media (lysogeny broth) supplemented with kanamycin (50 mg/ml) and agitated at 37°C overnight.

A 1L volume of ZY liquid broth, supplemented with kanamycin (50 mg/ml), was inoculated with 10 ml of the overnight pre-culture. The culture was incubated with agitation at 37°C until

an optical density of 0.5 - 0.6 (A_{600}) was reached using a SPECTROstarNano. Protein expression was induced by self-induction for 3 hours at 37 °C, followed by overnight incubation at 20°C.

b. Protein purification

Cultures were centrifuged at (8,000 × *g*) for 20 minutes at 4°C, cells were collected and re-suspended in 50mM Tris buffer pH 8 with 300mM Nacl, 10mM Imidazole, 1/200 DNase, 1/100 $MgSO_4$, complete protease inhibitor cocktail EDTA-free tablets [one tablet per liter of culture] (Roche) at 4°C, and were disrupted through an emulsiflex high-pressure homogenizer at 1500 bars. Soluble bacterial extract was then collected by centrifugation (10,000 × *g*) for 1 hour at 4 °C.

The extract was subjected to two purification steps using the ÄKTA™ go purification system (Cytiva). The first purification step using an Immobilized Metal Affinity Chromatography (IMAC) using His-Trap FPLC 5 ml column (Cytiva) equilibrated in 50mM Tris buffer pH 8, 300mM Nacl, 10mM Imidazole. Recombinant protein was then eluted using one step of 250mM Imidazole in the equilibrium buffer and 1.5 mL fractions are collected in 96-well plate. This step was followed by a second purification step by size exclusion chromatography (SEC) step using a HiLoad 16/600 Superdex 200 (Cytiva) pre-equilibrated with 50mM Tris pH 8, 150mM Nacl, 10% glycerol, 2mM DTT (w/v). Recombinant protein was then eluted using the same equilibrium buffer 50mM Tris pH 8, 150mM Nacl, 10% glycerol, 2mM DTT (w/v) and 2 mL fractions are collected in 96-well plate. Different fractions collected throughout the purification process are subsequently analyzed using sodium dodecyl sulfate-polyacrylamide gel electrophoresis (SDS-PAGE) to identify the fractions that contain the LESV protein. For structural study, protein samples were concentrated using Vivaspin centrifugal falcon with a 10 kDa cut-off (sartorius). Protein concentrations were determined using a Nanodrop Spectrophotometer (ND1000) from Thermo Scientific. The monodispersity of the obtained protein solution has been assessed by dynamic light scattering (DLS) using a Zetasizer pro (Malvern Panalytical).

The biophysical characteristics of ESV1 and LESV constructs calculated using Expasy Protparam Tool are shown in Table 4.1.

Table 4.1: ***Biophysical characteristics of ESV1 and LESV constructs based on Expasy Protparam Tool***

	ESV1	LESV
Amino acids (aa)	343	523
Molecular weight (Da)	39880.15	59595.26
Theoratical pI	5.27	5,04
Extinction coefficient, ε (M-1 cm-1)	166505	184035

4.3. Sodium dodecyl sulfate–polyacrylamide gel electrophoresis (SDS-PAGE)

This is a one-dimensional electrophoresis on polyacrylamide gel under denaturing condition, it allows a separation of proteins denatured by SDS and β-mercaptoethanol according to their molecular weight under the influence of an electric field.

Proteins are denatured in loading buffer (1X Tris-Glycine; 0.1X SDS; 25% glycerol; 10mM EDTA) in the presence of β-mercaptoethanol and bromophenol blue by incubation for 5 minutes at 90°C. Samples are loaded onto a 4% concentration gel and a 12% separation gel according to the compositions. Migration is done in a PowerPac Basic™ system (Biorad) at room temperature and 130V immersed in Tris-Glycine-SDS buffer at pH ≈ 9. Gel reading is done against a pre-stained molecular weight marker PageRuler™ 10 to 170 (ThermoSientific) after staining under agitation with Coomassie Brilliant Blue (7-8mL) InstantBlue™ G-25I (Expedeon) of the gel previously rinsed with distilled water.

4.4. Glucan Solutions Preparation

In this study, we prepared stock solutions of various glucans, including amylose, A600, pullulan, amylopectin, β-limit dextrin, and glycogen, for EMSA and CD experiments. The procedures for each glucan solution preparation are as follows:

For EMSA experiments, stock solutions of amylose 1% (w/v), A600 1% (w/v), pullulan 1% (w/v), amylopectin 1% (w/v), β-limit dextrin 1% (w/v), and glycogen 1% (w/v) were prepared. 0.1g of amylose powder (from potato, Sigma A0512) was dissolved in 1mL of 2M sodium hydroxide (NaOH) alkaline solution and then vortexed to ensure complete solubilization of the amylose powder. After adding 2 mL of distilled water, the solution was neutralized by adding 1 mL of 2M hydrochloric acid (HCl). The solution was adjusted to a final volume of 10 mL with distilled water, then heated at 50°C for 5 minutes and vortexed to ensure complete dissolution. The same procedure was performed for both A600 and pullulan from Sigma. 0.1g of amylopectin powder (from potato, Sigma A8515) was dissolved in 10 mL of distilled water under a hotplate (70°C) and then subjected to autoclaving to completely dissolve the amylopectin and produce a homogeneous solution. 0.1g of glycogen powder (from Oyster, Sigma 50573) was dissolved in 10 mL of distilled water and then stirred under heat (40°C) to obtain a homogeneous solution.

For CD experiments, glucan stock solutions were prepared using the same protocol by replacing water with protein buffer containing with 15mM Tris pH 7.5, 10% glycerol, 2mM DTT and 100mM NaCl for ESV1 protein, and 15mM Tris pH 8, 10% glycerol, 2mM DTT and 150mM NaCl for LESV protein.

4.5. AlphaFold2 modelling

The 3D structures of ESV1 and LESV proteins were modeled using AlphaFold2 (Jumper et al. 2021). For each protein, five different models were calculated and ranked according to their global pLDDT scores. These five molecular models were then superimposed to assess the potential positioning of dynamic regions. The models with the highest pLDDT scores were used for illustrations and additional molecular docking.

4.6. Synchrotron radiation Circular dichroïsm

Synchrotron radiation circular dichroïsm (SR-CD) measurements were conducted at the DISCO beamline of the SOLEIL Synchrotron, located in Gif-sur-Yvette, France. Samples containing 5 µL of ESV1 at 6.1 mg/mL and 2 µL of LESV at 13.4 mg/mL were placed between two CaF2 coverslips with wavelengths of 20 µm and 10 µm, respectively (Refregiers et al. 2012).

Utilizing a beam size of 4 × 4 mm and a photon flux per nm step of 2×10^{10} photons s^{-1} in the 270–170 nm spectral range, radiation-induced damage was avoided (A. J. Miles et al. 2008). CD spectra were obtained using IGOR software from WaveMetrics. Protein and buffer spectra were obtained consecutively, each being the average of three accumulations. Buffer baselines were recorded sequentially and subtracted from the spectra, with concentration in residues considered. Prior to measurements, the molar elliptical extinction coefficient of Ammonium *d*-10-Camphorsulfonate (CSA) was determined on the beamline and used as the calibration standard for all data measurements (Andrew J. Miles, Wien, et Wallace 2004). Data processing was conducted using CDToolX software (Andrew J. Miles et Wallace 2018). The impact of various glucans on the structure of LESV and ESV1 was examined by incubating protein/glucan mixtures for 2 hours and measuring spectra under conditions identical to those for native proteins. Mixtures consisted of 80% protein solution and 20% amylose or amylopectin solutions at 1% w/v. Spectra containing a mixture of 80% protein buffer and 20% glucan solutions were subtracted from the protein/glucan spectra prior to CSA calibration.

Temperature scans were performed to assess protein stabilization by glucan interaction. CD spectra were gathered from 20-30 to 90°C and processed as previously described with 3-5 ◦C temperature increases and 3 min of equilibration time.

The content of secondary structure elements in each protein, either alone or in the presence of glucans was estimated using BestSel (Micsonai et al. 2015).

4.7. Small-angle X-ray scattering (SAXS)

Each protein sample solutions were centrifuged for 10 minutes at 10,000 rpm before X-ray analysis to remove all aggregates. SAXS data were collected at the SWING beamline of Synchrotron SOLEIL (λ = 1.033 Å). All solutions were combined within a quartz capillary held at a constant temperature of 15°C. Monodisperse protein sample solutions were loaded onto a size exclusion column (SEC-3, 150 Å; Agilent) using an Agilent HPLC system and eluted into the capillary cell at a flow rate of 0.3 ml/min^{-1} (David et Pérez 2009). Subsequently, 50 µl of protein samples were injected for SAXS measurements. Numerous frames were collected during the first minutes of elution and averaged to account for buffer scattering. This was then subtracted from selected frames corresponding to the primary elution peak. Data reduction to absolute units, frame averaging, and subtraction were performed using FOXTROT (David et Pérez 2009). All subsequent data processing, analysis, and modelling stages were executed

using programs from the ATSAS suite (D. Franke et al. 2017). The radius of gyration (*Rg*) was determined through the Guinier approximation using PRIMUS (D. Franke et al. 2017). GNOM (D. I. Svergun 1992) was employed to calculate the pair-distance distribution functions [P(r)] and establish the maximum dimension of the macromolecule (Dmax).

DAMMIF/DAMAVER/DAMFILT (Volkov et Svergun 2003) were utilized to model protein envelopes and structures. BUNCH (Petoukhov et Svergun 2005) was used to model the missing parts of proteins not identified by Alphafold (Jumper et al. 2021).

The sample to detector distance is two meters and data have been collected to q_{max}=0,5Å^{-1}.

Kratky representation

SAXS analysis provides information on the level of structuring of the protein. One way to evaluate this level of structuring is Kratky representation (q^2. I(q) as a function of q). For a globular protein, the Kratky representation has the shape of a bell curve. On the other hand, for a protein with disordered regions, the values of q^2.I(q) increase for large values of q and can reach a plateau in the case of unstructured case of unstructured protein (Figure 4.2).

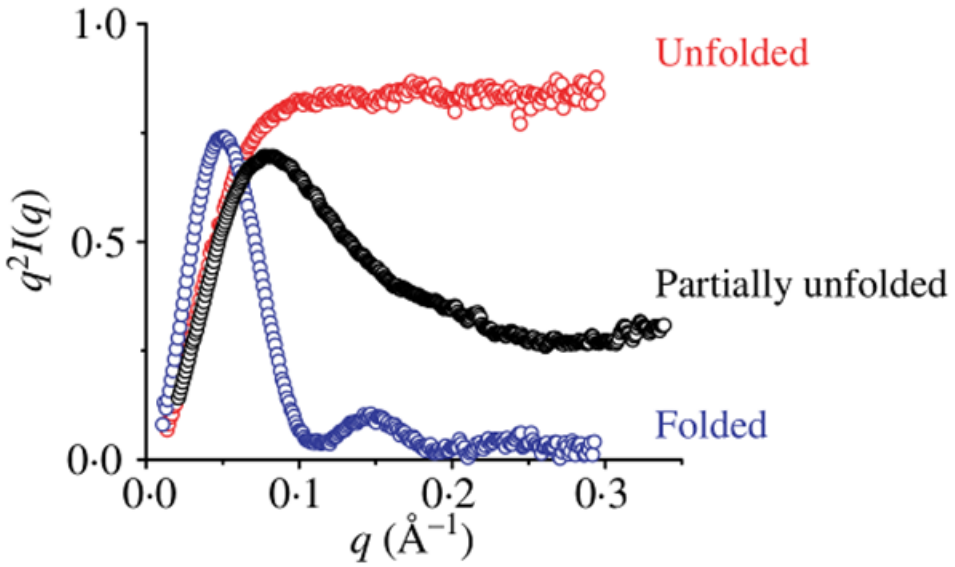


Figure 4.2: Kratky representation of proteins of different structuring levels. Kratky representation of a globular protein in purple, with disordered regions in black, a disordered protein in red.

Distance distribution function

The SAXS data also allows access to the dimensions of the protein, more particularly to its maximum diameter (Dmax). The distance distribution function P(r) is the Fourier transformation of I(q). It corresponds to a histogram where all distances interatomic distances (r) within the protein are represented. Nevertheless, its direct calculation from the scattering curve I(q) is limited by the cuts in resolution qmin, qmax and by the noise of the of the scattering curve.

The GNOM program (Svergun et al., 2001) is therefore used to compute the indirect Fourier transform of I(q). transform of I(q). The maximum diameter Dmax of the protein is then defined by the user thanks to the graph P(r)=f(r) calculated by GNOM.

The shape of the distance distribution curve obtained allows to deduce information on the shape of the protein.

Ab initio modelling

Using the distance distribution function obtained, the protein envelope can be modeled. For this purpose, the program GASBOR (Svergun et al., 2001) was used.

The GASBOR program uses the entire scattering curve for shape calculations. The amino acids of the protein are represented by N balls of 0.38 nm in diameter (in order to respect the constraints of interatomic distances). These N pseudo residues, correspond to the numbers of amino acids in the protein, GASBOR then proceeds by iteration: the position of a pseudo residue is changed randomly, then the curve of the obtained model is calculated and compared to the experimental curve, where the divergence between the two is defined by $\chi 2$. At each cycle, the program refines the position of the pseudo-residuals allowing to minimize this divergence. In order to achieve a reliable model, 10 envelopes were generated by GASBOR and we selected the most representative model.

Those envelopes generated by GASBOR are then averaged by the DAMAVER program (Volkov et Svergun 2003). The resulting envelope thus reflects an average of the conformations corresponding to the experimental SAXS data.

4.8. Docking

The alphafold model structure of the conserved C-terminal domain of ESV1 and LESV was used as a target to model the binding position of the double helix of the amylopectin model derived from Polysac3DB (CERMAV https://polysac3db.cermav.cnrs.fr). A genetic algorithm was used for the search phase within a 10 Å sphere centered on the tryptophan stripes of the conserved β-sheet. The scoring function was based on the ChemPLP force field as used by GOLD (Jones et al. 1997). All settings were left as defaults. A subsequent energy minimization was performed on the best model using the Amber force field. All images depicting the molecular structures of proteins and ligands were generated using Pymol (The Pymol Molecular Graphics System, version 1.8.0.0 Schrödinger, LLC).

4.9. 3D imaging by fluorescence microscopy

In this study, we used waxy corn starch granules. The waxy variety of starch is obtained from plants with mutations for GBSS, an enzyme that facilitates the production of amylose and is responsible for the autofluorescence of the granules. We cleaned the starch granules with water, acetone, and ethanol to remove any phenolic contaminants from plastic storage, following the procedure described in (Tawil et al. 2011).

We then monitored the binding of ESV1 and LESV to the starch granules using UV microscopy. This process was further investigated using both visible and UV microscopy, following the same protocol described in (Jamme et al. 2014). We regulated the excitation at λ=310 nm and used an emission filter at 329-351 nm to observe the tryptophan emission (at 345 nm) of the bound proteins.

Data were collected over a period of 10 seconds for each emission fluorescence image and 0.2 seconds for visible images. For 3D visualization, Z-slices (along the optical axis) were acquired with a step size of 300 nm over a 40 µm Z-range under the control of µManager according to the method described in (Edelstein et al. 2010).

4.10. Electrophoretic Mobility Shift Assay (EMSA)

Electrophoretic Mobility Shift Assay (EMSA) or Gel Shift Assay is a simple, efficient, rapid, and versatile technique used to study protein-glucan interactions. It is based on the decrease in the electrophoretic mobility of the protein in acrylamide gels caused by an interacting protein to an increasing concentration of glucans in a non-denaturing condition. The EMSA assay provides valuable insight into the binding specificity, affinity, and stoichiometry of the protein-glucan interaction. This interaction of the protein with the glucan will result in a reduced migration rate compared to the free glucan, resulting in a "shift" in the position of the band in the gel, a phenomenon that can be seen on the gel with the naked eye as we can see in Figure 4.3.

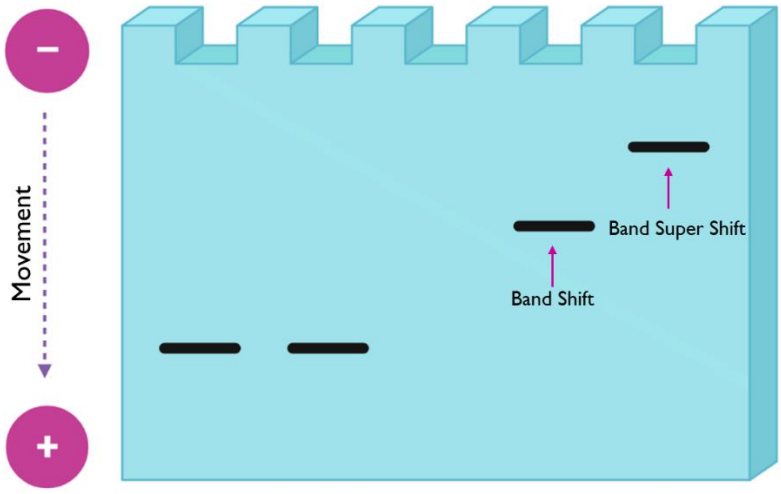


Figure 4.3: ***Electrophoretic Mobility Shift Assay gel***

The presence of a specific glucan in the gel causes the protein to interact and migrate differently with the different types and concentrations of the glucan used, resulting in a shift in the band visible upon detection. The technique is suitable for qualitative, quantitative, and kinetic analyses. In our experiment, we investigated whether the proteins can interact with the different ligands used and induce conformational changes upon interaction. To test this, we have prepared polyacrylamide electrophoresis gels containing 8% of 40% Acryl/Bisacryl solution (29:1) with an increasing concentrations of glucan solutions ranging from 0% to 0.5%. Once the gels are prepared, 1µg of the purified proteins ESV1 and LESV were loaded on the different

gels containing the glucans. We used a reference protein which has no affinity for sugars. Electrophoresis was carried out under native conditions at 4°C at a constant voltage of 130V for 2 hours in a migration buffer containing 25mM Tris and 192 mM Glycine. When the desired separation was achieved, the gels were rinsed with distilled water and then stained with InstantBlue™ Coomassie Brilliant Blue G-25I (Expedeon).

5. Results

5.1. Production and Purification of ESV1 and LESV proteins

Structural studies by SAXS and X-ray crystallography require pure and monodisperse protein solutions. In order to obtain accurate and reliable structural analysis, it is essential to have pure protein samples with consistent behavior in solution. Impurities and aggregates can lead to inaccurate data and hinder the ability to decipher the true structure of the protein. Therefore, obtaining high quality protein samples is critical for the success of SAXS and X-ray crystallography experiments. This task was the first part of my research work.

5.1.1. Purification of ESV1 protein

Before starting my PhD, preliminary experiments were conducted using the ESV1 and the LESV protein. pET24a+ AtESV1 -Δ95 and pET28a+ AtLESV were chosen in order to express and purify ESV1 and LESV proteins.

Our collaborators tried to work with the full-length of ESV1 protein, but they never succeeded in expressing the full length of the ESV1 protein. So, they decided to do a deletion of the first 95 residues in which they succeeded in expressing and produce the protein. They send us the construct and preliminary purification protocol. ESV1 is expressed at lower levels than LESV and is less soluble during the concentration process.

ESV1 protein was eluted using an imidazole-containing buffer and then dialyzed. The main issue is that the protein precipitate rapidly after few days. So, we tried a second purification protocol by using a Superose 6 column or by doing directly a desalting step after using the HisTrap chromatography column.

Later on, we tried to design a new construct for the ESV1 protein that contains only the β-sheet a construction from the residues E139 till E371 but we never succeeded in expressing this construct.

In the Materials and Methods segment, the protocol for expressing and purifying the ESV1 protein are described in details and the steps of the final optimum purification process for ESV1 protein are shown in Figure 5.1. To achieve stable ESV1 protein samples, many steps required optimization.

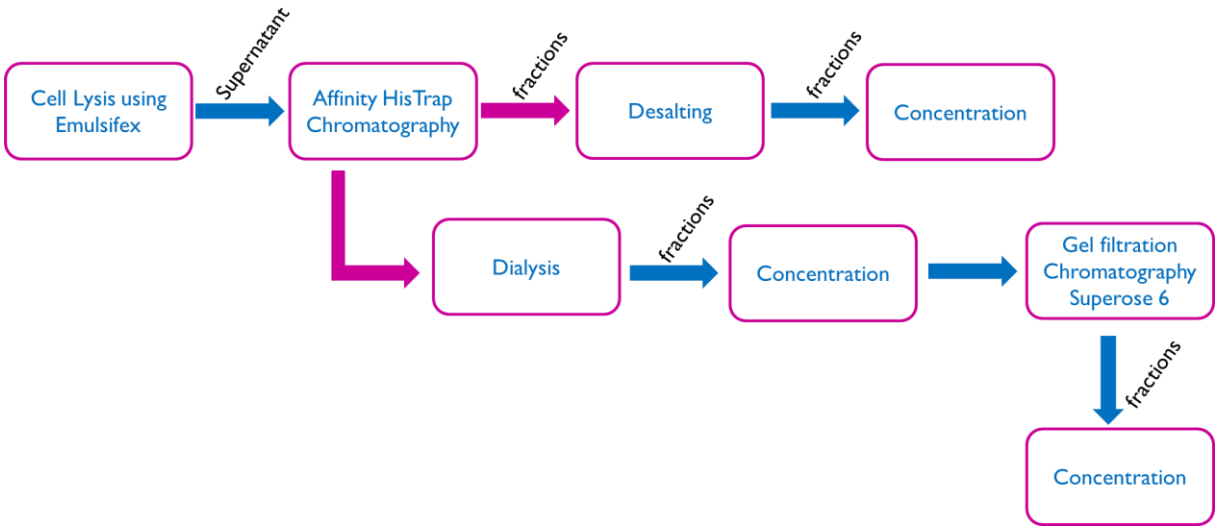


Figure 5.1: ***Diagram of the purification steps of the ESV1 protein***

The chromatogram of the ESV1 HisTrap purification with the UV absorbance curves at 280 nm, and the 12% SDS-PAGE with different fractions are shown in Figure 5.2.

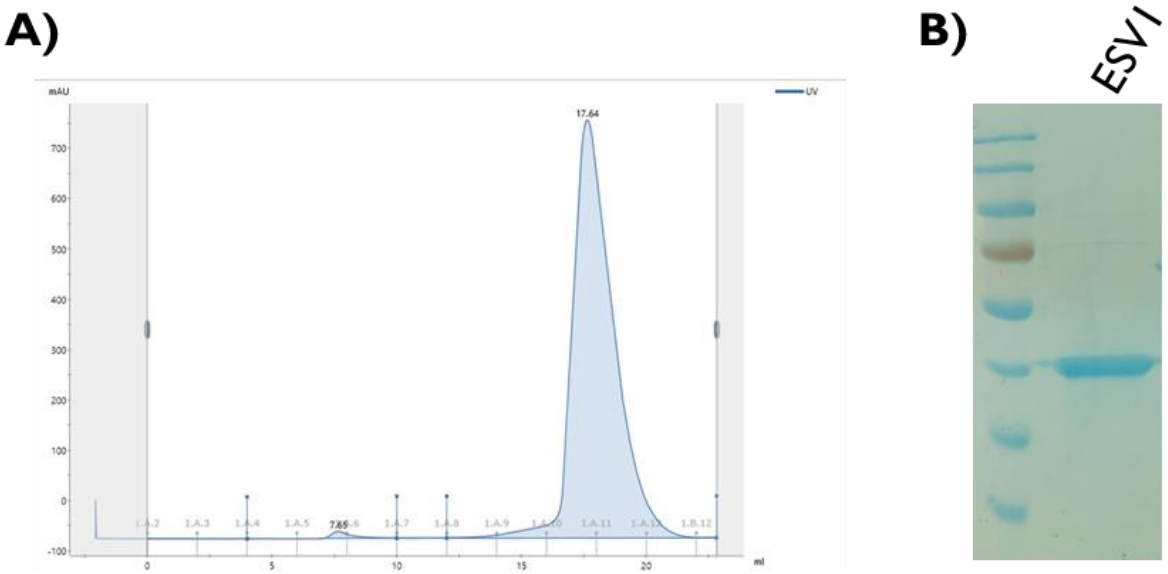


Figure 5.2: ***A) Purification profile of ESV1 by size exclusion chromatography Column*** volume: 120mL; Flow rate: 1mL/min. The blue curve represents the absorbance at 280 nm of the protein as a function of the volume. B) Analysis of the purified fractions of ESV1 using 12% SDS-PAGE in which Coomassie blue stained showing 2 µg of purified protein samples of ESV1.

The elution profile shows a single peak with elution volumes of 86ml (Figure 5.2) containing pure ESV1 (Figure 5.2). The protein was concentrated for crystallization and SAXS experiments to a maximum concentration between 4 and 6 mg/ml.

5.1.2. Purification of LESV protein

We succeeded in expressing and producing the full-length of LESV protein with a good expression and soluble. LESV protein was eluted using an imidazole-containing buffer for the first step of the purification, and with a second purification protocol by using a Superdex 200 column.

In the Materials and Methods segment, the protocol for expressing and purifying the LESV protein are described in details and the steps of the final optimum purification process for LESV protein are shown in Figure 5.3.

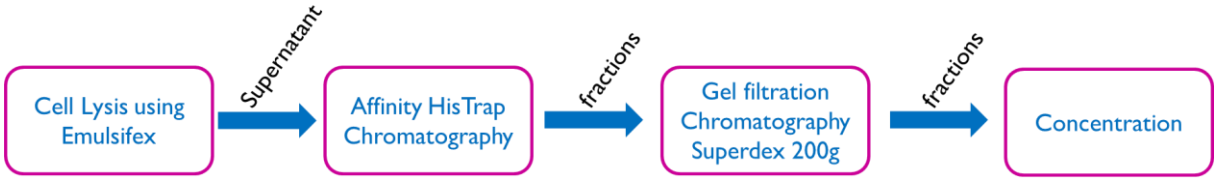


***Figure 5.3:* Diagram of the purification steps of the LESV protein**

The elution profile shows two main peaks with elution volumes of 77ml and 60ml (Figure 5.4) both containing LESV (Figure 5.4).

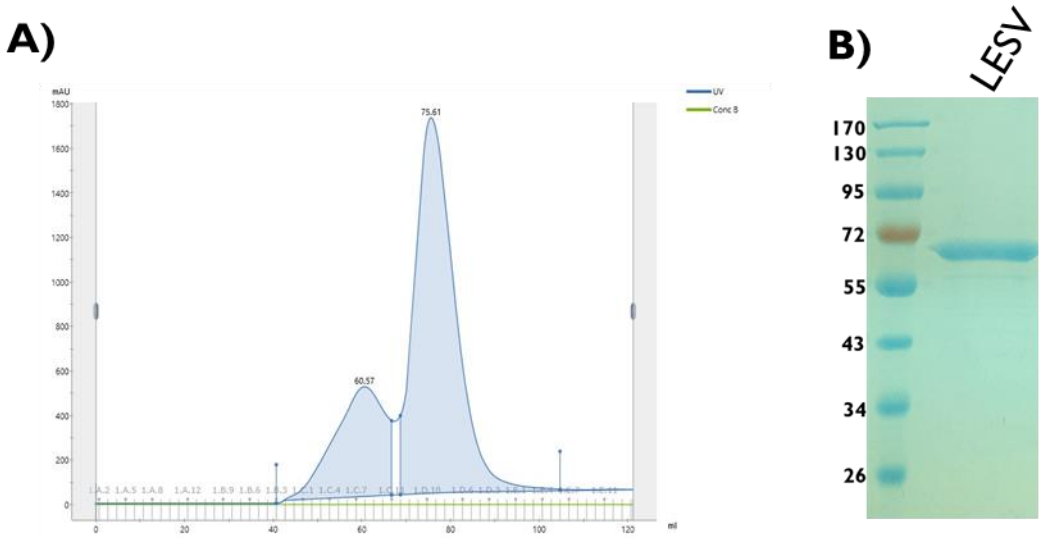


Figure 5.4: ***A) Purification profile of LESV by size exclusion chromatography Column*** volume: 120mL; Flow rate: 1mL/min. The blue curve represents the absorbance at 280 nm of the protein as a function of the volume. B) Analysis of the purified fractions using 12% SDS-PAGE gel in which Coomassie blue stained showing 2 µg of purified protein samples of LESV.

The peak with an elution volume of 77ml corresponds to the monomeric form of LESV that has been used for structural studies. LESV was then concentrated using Vivaspin concentrator. The protein was concentrated to a maximum concentration between 20 and 25mg/ml.

5.2. Structural study of ESV1 and LESV

One of the objectives of my PhD project is to resolve the molecular structure of the *Arabidopsis thaliana* ESV1 and LESV proteins at the atomic level using an integrative structural biology approach, mainly combining X-ray crystallography for stable domains and/or full-length proteins, small angle X-ray scattering to analyze disordered regions and dynamic organization of the domains, and circular dichroism to estimate the composition of the protein's secondary structure.

Since the ESV1 and LESV proteins have no known homologs, we decided to start our study by studying their structure by SAXS with the aim of identifying their different domains and their organization in order to obtain more information about the protein structure in solution. In order to analyze the structural organization and to complete the AlphaFold predictions of the ESV1

and LESV protein structures, we first examined the newly accessible AlphaFold2 prediction models for the ESV1 and LESV protein structures, and we experimentally validated these two AlphaFold2 prediction models using two structural biology approaches small angle X-ray scattering (SAXS) and circular dichroism (CD) techniques in order to evaluate their secondary structure content.

5.2.1. AlphaFold2 models: ESV1 and LESV

Recently, two molecular models for the ESV1 and LESV proteins from *Arabidopsis thaliana* were available in the alphafold database (Jumper et al. 2021). AlphaFold is a very powerful tool that reveals the tertiary structure models of both proteins in which they share a common β-sheet at the C-terminal region, a highly conserved region rich in tryptophan residues (Figure 5.5). These tryptophan-rich regions in both ESV1 and LESV proteins are confidently predicted with a value of about pLDDT >90 as we can see in Figure 5.5.

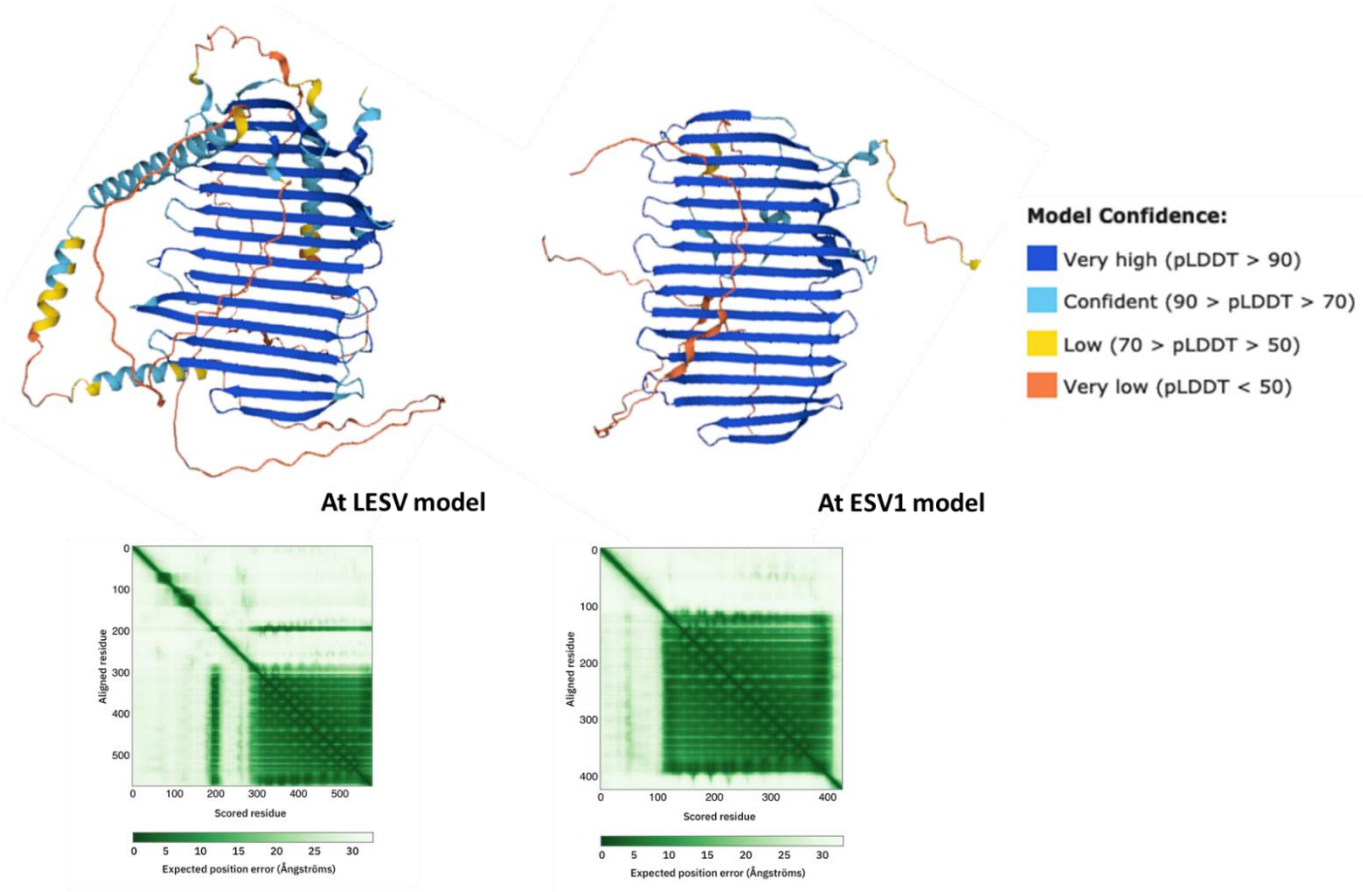


Figure 5.5: AlphaFold2 structure models predicted for full-length of *Arabidopsis* ESV1 (right: F4I9G2**) and LESV (left:** Q5EAH9**) proteins. A)** Structural elements with a confidence score of pLDDT > 90 are colored with blue color, elements with 90 > pLDDT < 70 are colored with cyan color and those with pLDDT<70 are colored with yellow color. **B)** Plots of predicted positional errors in Å for ESV1 (right) and LESV (left).

The N-terminal (including the supposed transit peptides) and the C-terminal of these β-sheets have lower confidence scores. The majority of these sections are likely disordered and exhibit poor conservation in the ESV1 protein in addition to the proline-rich region as we can see in Figure 5.5 (Feike et al. 2016), while numerous α-helices are predicted with high confidence (pLDDT > 70) in both the N- and C-terminal of the β-sheet in the LESV protein. In addition, a short helical region is predicted with high confidence (pLDDT > 90) at the C-terminus of the LESV protein (Figure 5.5). In addition to the prediction of the C-terminal region, AlphaFold2 has no prediction for the N-terminal domains.

The tryptophan-rich regions in both proteins exhibit unique conformations, forming elongated, twisted planar sheets that are around 70 Å long and consist of 16 antiparallel β-strands (Figure

5.6). In these antiparallel β-strands, the conserved tryptophan residues are arranged along two lines parallel to each other and to the axis of the sheet separated by approximately 10Å (Figure 5.6).

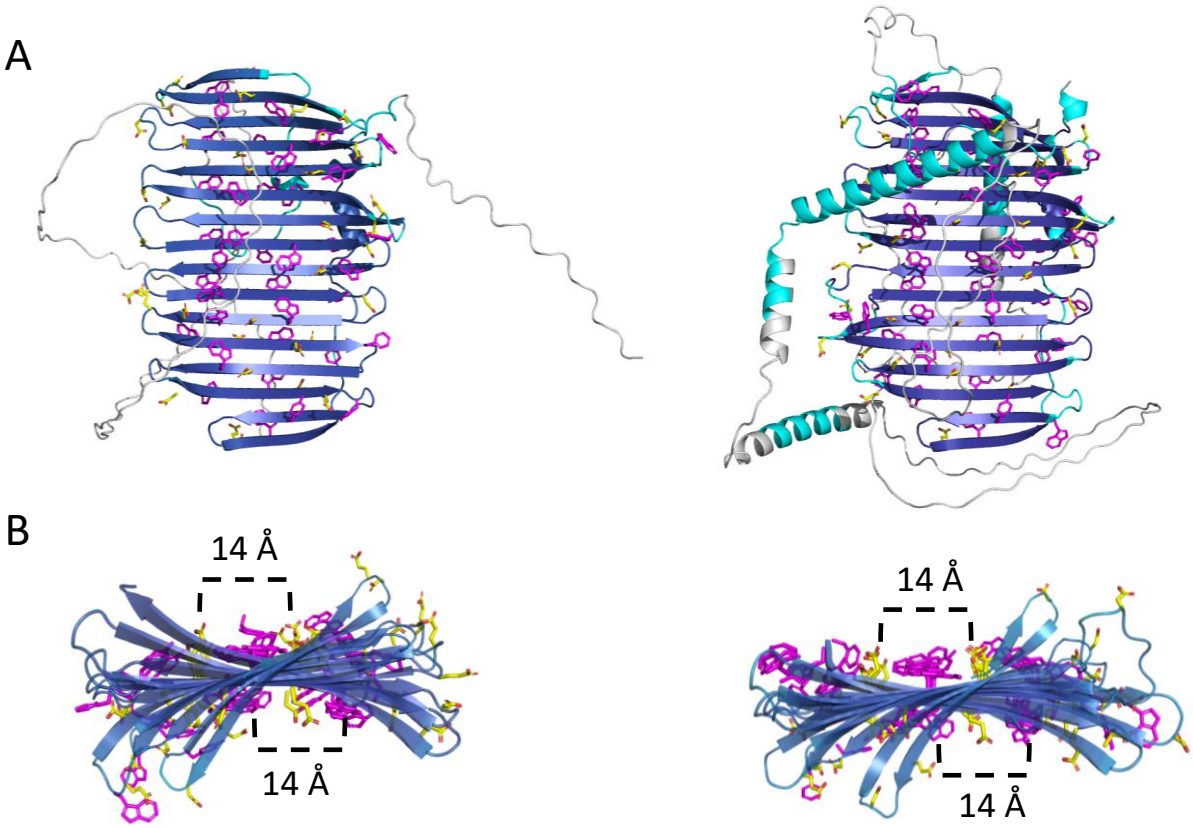


Figure 5.6: Structures and conservation of ESV1 and LESV proteins. (**A**) AlphaFold predictions of ESV1 and LESV (Uniprot identifiers F4I9G2 and Q5EAH9, respectively). Cartoon representations of the full proteins; aromatic (Trp, Tyr, Phe) and acidic residues are highlighted in pink and yellow, respectively (side chains shown as sticks; for better visualization, their alpha carbons are colored as the cartoon backbone). (**B**) Isolated β-sheet plane regions of ESV1 and LESV, reoriented along the x-axis to reveal the view along the long axis of the planes.

Since the aromatic and acidic residues in the β-sheets of ESV1 and LESV proteins are highly conserved, we looked at their spatial arrangement along the β-sheets. We found that the aromatic residues in the β-sheets form two lines, approximately 14 Å apart, spanning both sides of the β-sheets and perpendicular to the backbone β-strands. In addition, the acidic amino acids predominantly occupy the spaces between these aromatic lines, similarly arranged in two lines

spaced about 14 Å apart on both sides of the β-sheet. The distribution and orientation of the aromatic and acidic amino acids across the β-sheet is highly similar, but not identical between the two proteins (Figure 5.6).

5.2.2. The N-terminal domain of LESV contains helices whose folding and position depend on the environment of the protein.

To expand upon the Alphafold models for the ESV1 and LESV proteins described above, which are available in the Alphafold platform database. Independently of these two models, we recalculated a set of five models for the two proteins using Alphafold 2 by calculating the models for the N-terminally truncated constructs generated for expression of the two proteins used in the CD and SAXS experiments. We found that the predicted high-confidence regions with a pLDDT > 90 for the ESV1 protein between amino acids 142 to 395, and for the LESV protein between amino acids 318 to 573 remained consistent across the five generated models and consisted primarily of the tryptophan-rich, β-stranded C-terminal region described above. This result indicates a high level of stability for this domain, which is expected to be a new carbohydrate binding domain. In these five models, we focused on the predicted structures for the less conserved domains of the ESV1 and LESV proteins. As previously reported, the regions near the C-terminal β-sheet of ESV1 (46 amino acid residues at the N terminus and the polyproline region at the C terminus), which are poorly conserved, are predicted to be disordered. In the C-terminal region of the β-sheet, located on one side of the β-sheet, an α-helical region from residues 378 to 395 is predicted with a high confidence score in its structure and relative position to the β-sheet. This helix is also present in the LESV protein from residues 555 to 578 with the same confidence level as in ESV1. Both regions contain four conserved amino acids.

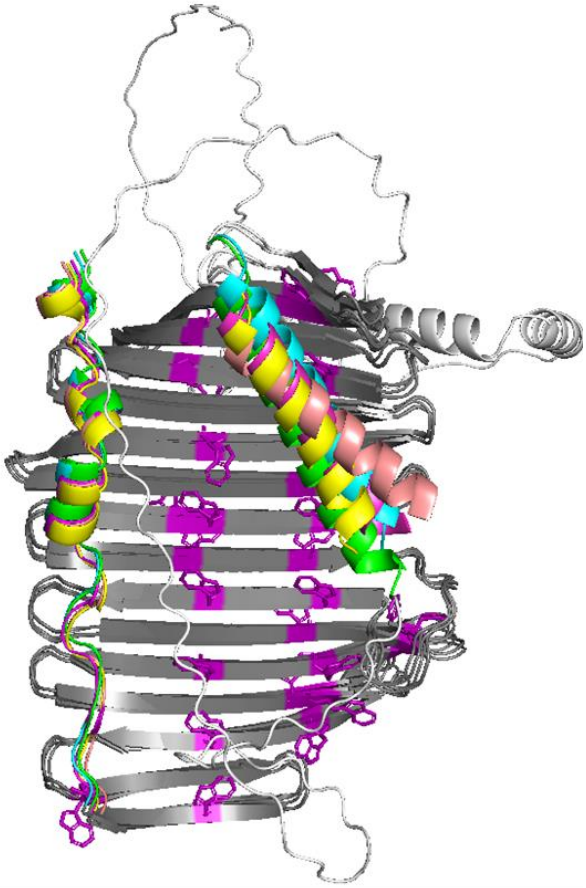


Figure 5.7: AlphaFold2 prediction model for *Arabidopsis* LESV. A) Structural elements with a confidence score of pLDDT > 90 are colored with dark grey color, the helices of the 5 models are colored light grey, salmon, cyan, yellow and green. The common helix with ESV1 is on the left, the long helix specific for LESV is on the right. The side chains of aromatic amino acids (Tyr, Trp, Phe) are shown in violet at the C-terminus of the β-sheet.

The N-terminal region of LESV is predicted with low to extremely low confidence. Nevertheless, three helical regions located in a conserved island were found to be predicted with pLDDT > 70. Only one of them forms a long helix from residues 245 to 273 whose position relative to the β-sheet, is predicted with high score of confidence (Figure 5.7). This helix is located on the same side of the β-sheet as the C-terminal helix conserved in both proteins. It is predicted with the same confidence score in all Alphafold-generated models, but its position may differ from model to model, suggesting that it may adjust according to the environment of the protein. For the other helices, although they are predicted with pLDDT > 70, their positions are not equivalent between the five models calculated. This result may indicate that the N-

terminal domain of LESV is disordered and susceptible to helix folding under certain conditions.

5.3. Determination of the molecular envelop of purified ESV1 & LESV by SAXS

In order to gain insight into the three-dimensional shapes of the proteins in solution we have used small angle X-ray scattering (SAXS). SAXS data were collected at the SWING beamline of the Synchrotron SOLEIL (Gif sur Yvette). Scattering data were recorded and several curves of scattered intensity (I(q)) as a function of the scattering vector (q) were obtained.

5.3.1. SAXS data for ESV1 protein

The ESV1 protein exhibited a smooth SAXS scattering profile, and analysis by Guinier approximation plots confirmed that it was not aggregated (Figure 5.8A). Examination of the protein particle size (Table 5.1) and interatomic distance distributions (Figure 5.8B) revealed that the ESV1 is composed of structured domains extended by a more dynamic elongated domain.

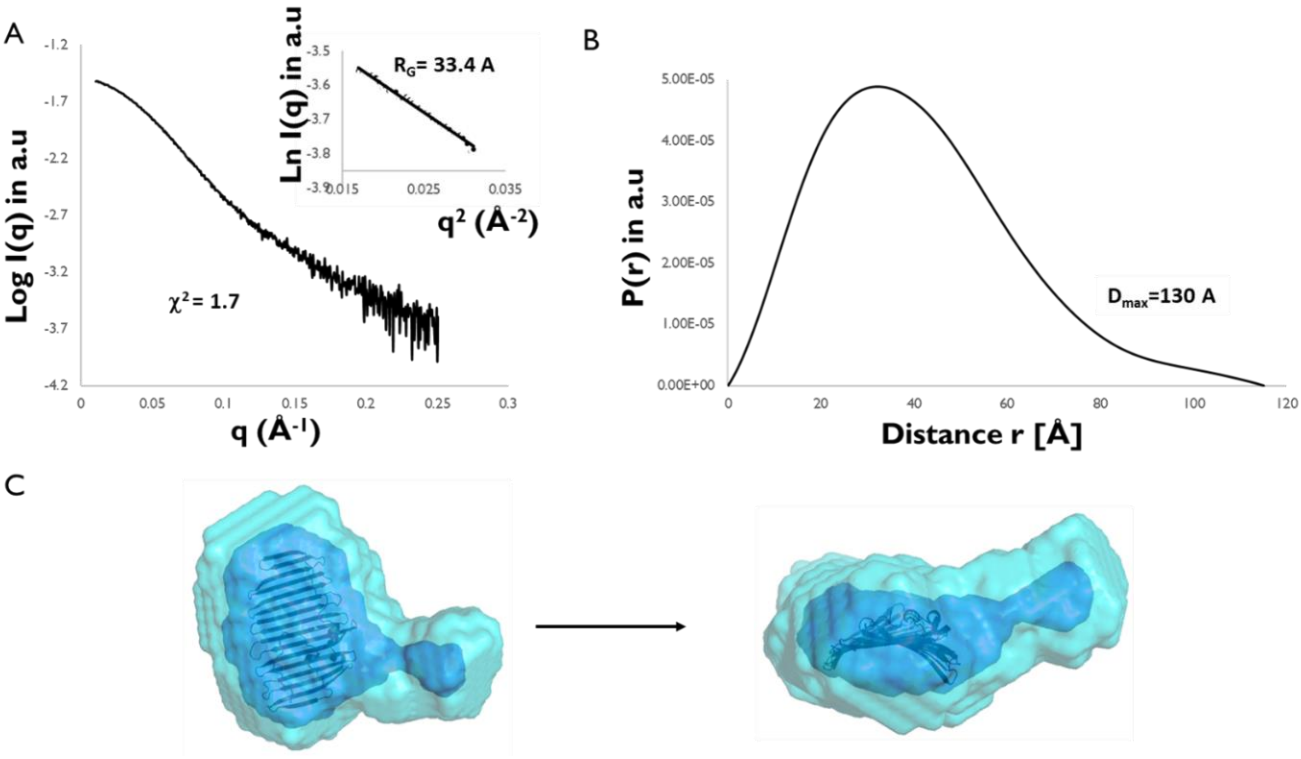


Figure 5.8: ***SAXS data and ab initio model for ESV1 protein*** **A)** Experimental data plotted as a function of the scattering vector q with the Guinier plot illustrated as an inset **B)** Illustrates the distance distribution function, and **C)** Average *ab initio* envelope predicted by DAMMIF/DAMAVER (from 10 *ab initio* protein models computed by DAMMIF, shown in light blue) superposed with the most typical shape filtered by DAMFILT (dark blue). The proteins' structural parts modeled with a confidence score of pLDDT ≥ 90 (cartoon) are manually fitted into the molecular envelope.

Table 5.1: ***SAXS structural parameters for ESV1 protein***

Beamline SWING, SOLEIL	**At ESV1**
I(0) (cm^{-1})	0.017 +/- 0.00005
R_g (Å)	33.40 +/-0.19
q_{min} ($Å^{-1}$)	0.0032
qR_g max	1.01
Coefficient of correlation, R^2	0.991
P(r) analysis	
I(0) (cm^{-1})	0.018 +/-0.000025
R_g (Å)	34.7 +/-0.02
q range $_{min}$ ($Å^{-1}$)	0.0032 to 0.25
Porod volume ($Å^3$)	103015
Dmax	130

The scattering curve obtained shows that ESV1 is probably formed by a rather flattened discoid domain extended by small dynamic regions (Figure 5.8B) in agreement with the model generated by Alphafold.

Ten *ab initio* molecular envelopes were calculated for ESV1 protein with DAMMIF (Daniel Franke et Svergun 2009) fitting the experimental data with χ^2 =1.7 for the ESV1 protein and then averaged and filtered with DAMAVER/DAMFILT (Volkov et Svergun 2003) . The result (Figure 5.8C) shows an oblate domain whose dimensions and shapes are consistent with the shape of the conserved β-sheet of ESV1 AlphaFold model.

5.3.2. Validation of AlphaFold ESV1 model and modelling of the N-terminus

To verify the compatibility of the ESV1 model with the experimental SAXS data, in addition to understanding and identifying the N-terminal domain of ESV1, we performed an *ab initio* modelling. We used the BUNCH program by integrating the high confidence structure of the C-terminal domain available in the Alphafold database with our SAXS data. BUNCH is a program designed to model the missing parts of a protein that are not involved in the predicted β-sheet using SAXS data (Petoukhov et Svergun 2005).The resulting model fits the SAXS data

with a high degree of confidence χ^2= 3.4, (Figure 5.9) and suggests that both the N-terminal and proline-rich regions are disordered and located on either side of the beta domain leaving the latter accessible for the binding of amylopectin molecules.

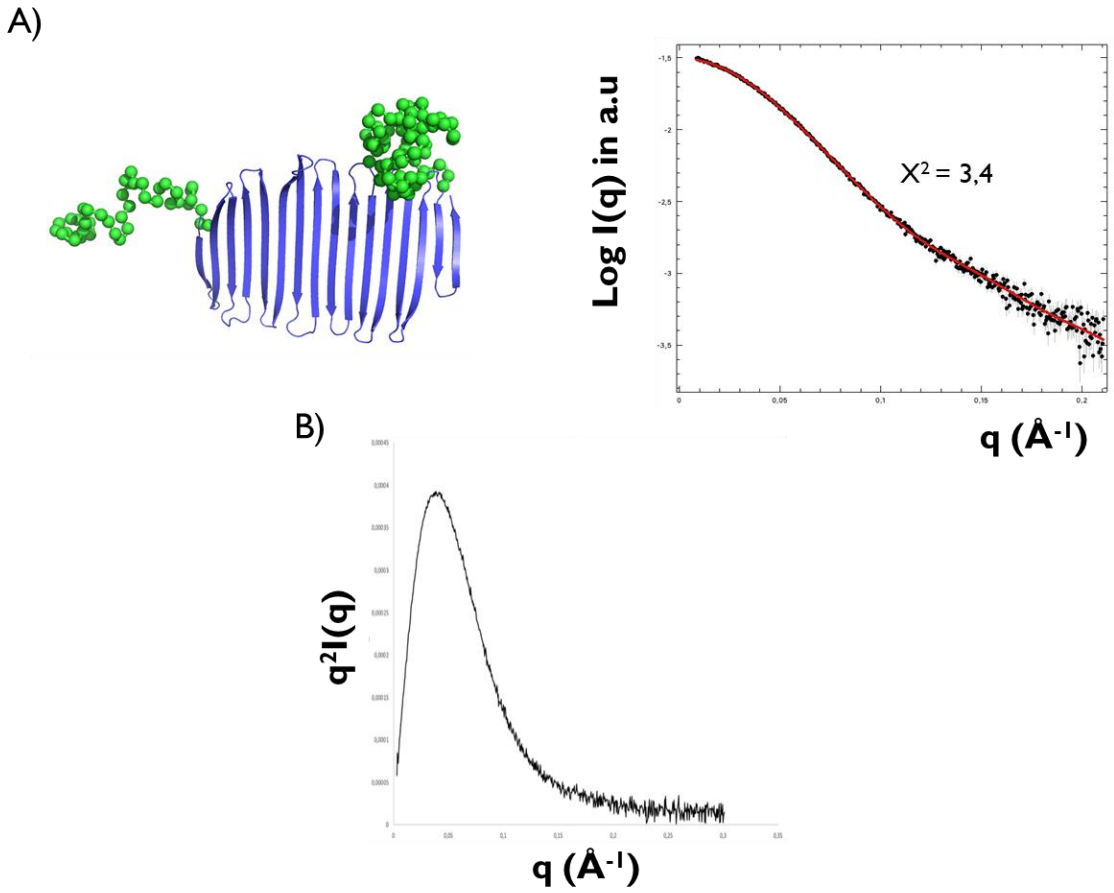


Figure 5.9: ***SAXS model of ESV1 domains.*** A) The left panel shows the optimal positions and orientations of the conserved β-sheet models of ESV1 in blue color, and the probable conformations of the N and C terminal domains in green color. The fit of the experimental scattering data (black dots) and the theoretical scattering patterns of the models are represented on the right panel. B) Kratky plot of scattering data for the ESV1 protein.

Kratky plot shows a typical bell-shape finishing to $q^2I(q)$ close to 0 typical of folded protein (Figure 5.9) suggesting only few unstructured regions in the protein.

5.3.3. SAXS data for LESV protein

The LESV protein exhibited a smooth SAXS scattering profile, and analysis by Guinier approximation plots confirmed that it was not aggregated (Figure 5.10A). Examination of protein particle size (Table 5.2) and interatomic distance distribution (Figure 5.10B) revealed that LESV protein consists of structured domains extended by a more dynamic elongated domain.

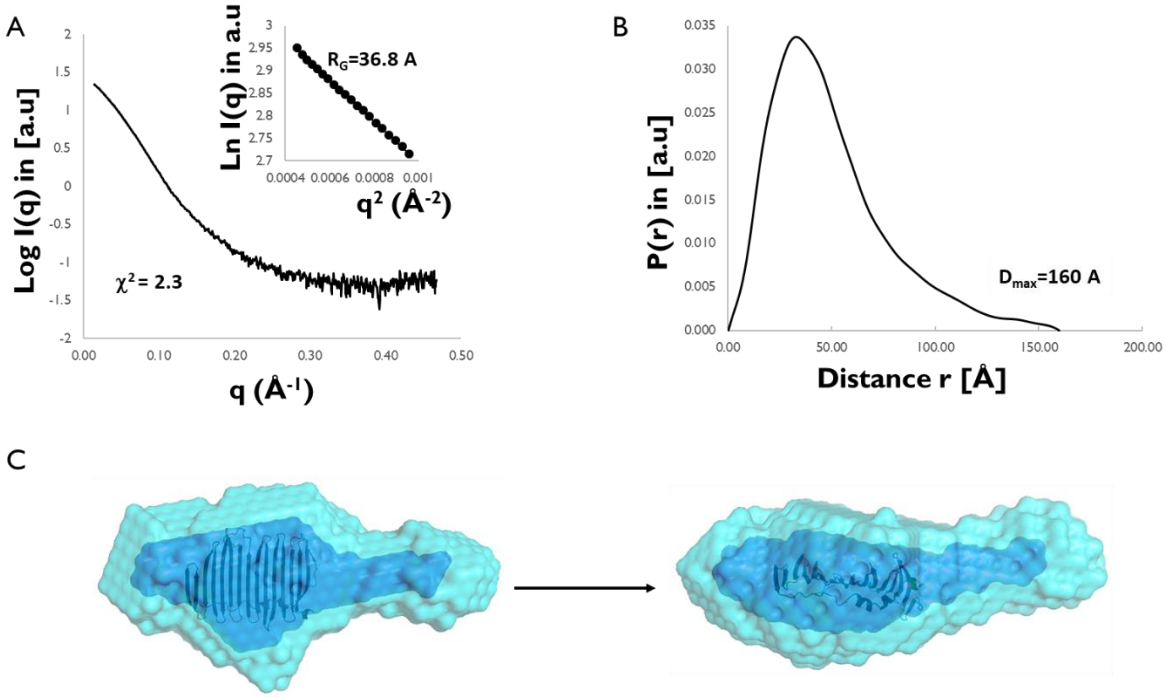


Figure 5.10: ***SAXS data and ab initio model for LESV protein*** **A)** Experimental data plotted as a function of the scattering vector q with the Guinier plot is illustrated as an inset **B)** illustrates the distance distribution function, and **C)** Average *ab initio* envelope predicted by DAMMIF/DAMAVER (from 10 *ab initio* LESV protein models computed by DAMMIF, shown in light blue) superposed with the most typical shape filtered by DAMFILT (dark blue). The proteins' structural parts modeled with a confidence score of pLDDT ≥ 90 (cartoon) are manually fitted into the molecular envelope.

Table 5.2: SAXS structural parameters for LESV protein

Beamline SWING, SOLEIL	**At LESV**
I(0) (cm^{-1})	23.31 +/-0.033
R_g (Å)	36.79 +/-0.08
q_{min} ($Å^{-1}$)	0.014
qR_g max	1.18
Coefficient of correlation, R^2	0.999
P(r) analysis	
I(0) (cm^{-1})	23.82 +/-0.034
R_g (Å)	39.71 +/-0.102
q range $_{min}$ ($Å^{-1}$)	0.014 to 0.4
Porod volume ($Å^3$)	121099
Dmax	160

We have used the same approach used as for ESV1 to study the position of the N-terminal domain of LESV. Ten *ab initio* molecular envelopes were calculated for LESV protein with DAMMIF (Daniel Franke et Svergun 2009). Its molecular envelope fits the experimental data with χ^2 =2.3, and the average model was filtered with DAMAVER/DAMFILT (Volkov et Svergun 2003) as shown in Figure 5.10C which represents the most typical model.

LESV has a large, oblate globular domain whose size and shape are compatible with the β-sheet domain model, although it is larger than the main domain of ESV1. This domain is flanked by a small, rather dynamic extension whose size cannot correspond to the entire N-terminal domain of the protein, suggesting that it does not form a structurated domain apart of the C-terminus domain.

5.3.4. LESV model and modelling of the N-terminus

To better visualize and localize the N-terminal domain of LESV, we performed an *ab initio* modelling based on the structure of the high-resolution C-terminal domain model available in the Alphafold database with our SAXS data using BUNCH. The resulting model, which closely aligns with the SAXS data with a high degree of confidence (χ^2= 3.8) as shown in Figure 5.11, suggests that the ends of the C and N-terminal domains between 50 and 80 residues are likely

to be disordered and emerge from the overall structure. The rest of the domain, however, is organized around the β-sheet in structures that could be compatible with the helices predicted by Alphafold.

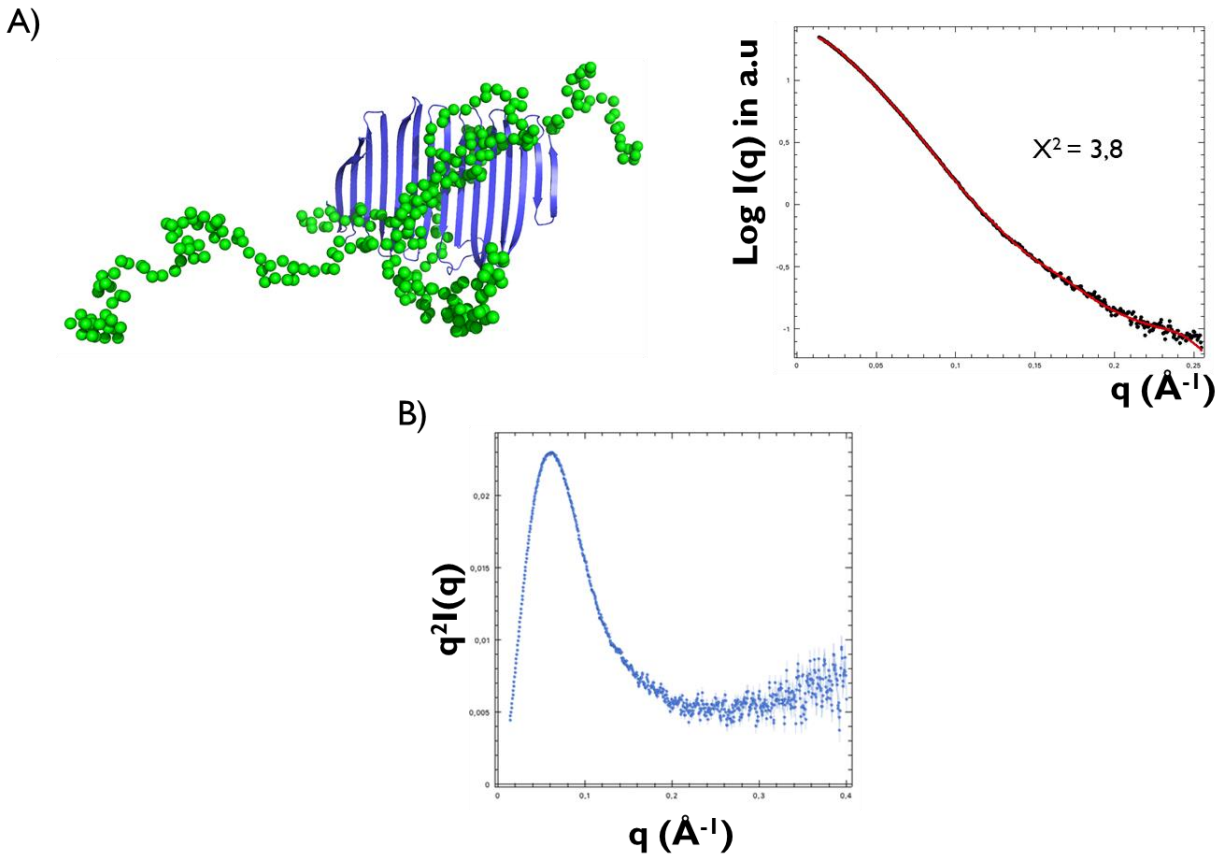


Figure 5.11: ***SAXS model of LESV domains.*** A) The left panel shows the optimal positions and orientations of the conserved β-sheet models of LESV in blue color, and the probable conformations of the N and C terminal domains in green color. The fit of the experimental scattering data (black dots) and the theoretical scattering patterns of the models (red line) are represented on the right panel. B) Kratky plot of scattering data for the LESV protein.

The presence of this domain around, or at least close to, the β-sheet is likely to affect glucan binding and could explain the differences in function between the two proteins described in (C. Liu et al. 2023). Kratky plot shows for the LESV protein a bell-shape profile that does not come back to $q^2I(q)=0$ which is typical of a folded protein having some disordered regions (Figure 5.11b).

5.3.5. Determination of the molecular envelope of the LESV protein with α-cyclodextrins by SAXS

Effect of interaction with alpha cyclodextrin

It's known that LESV interacts with starch (Feike et al. 2016). As starch components (amylopectin and amylose) cannot be prepared as homogeneous solution they are not good candidates for both crystallization and SAXS experiments that require monodisperse solution.
Cyclodextrin are macrocyclic rings of glucose (glc) subunits joined by α-1,4-glycosidic bonds that can mimic the helical organization of the amylopectin. For this study, we used α-CD (6 glc subunits). SAXS data have been collected for LESV incubated with α-CD in a molecular ratio 1:2.
After incubation with alpha cyclodextrin, the HPLC profile showed a single peak with an elution volume analogue to the protein alone.

Ab initio modelling from SAXS data using DAMMIF program (Figure 5.12) (Daniel Franke et Svergun 2009) shows that the particles have a rather spherical shape with a short elongated part. SAXS structural parameters seem to indicate the presence of interdomain dynamic (Table 5.3).

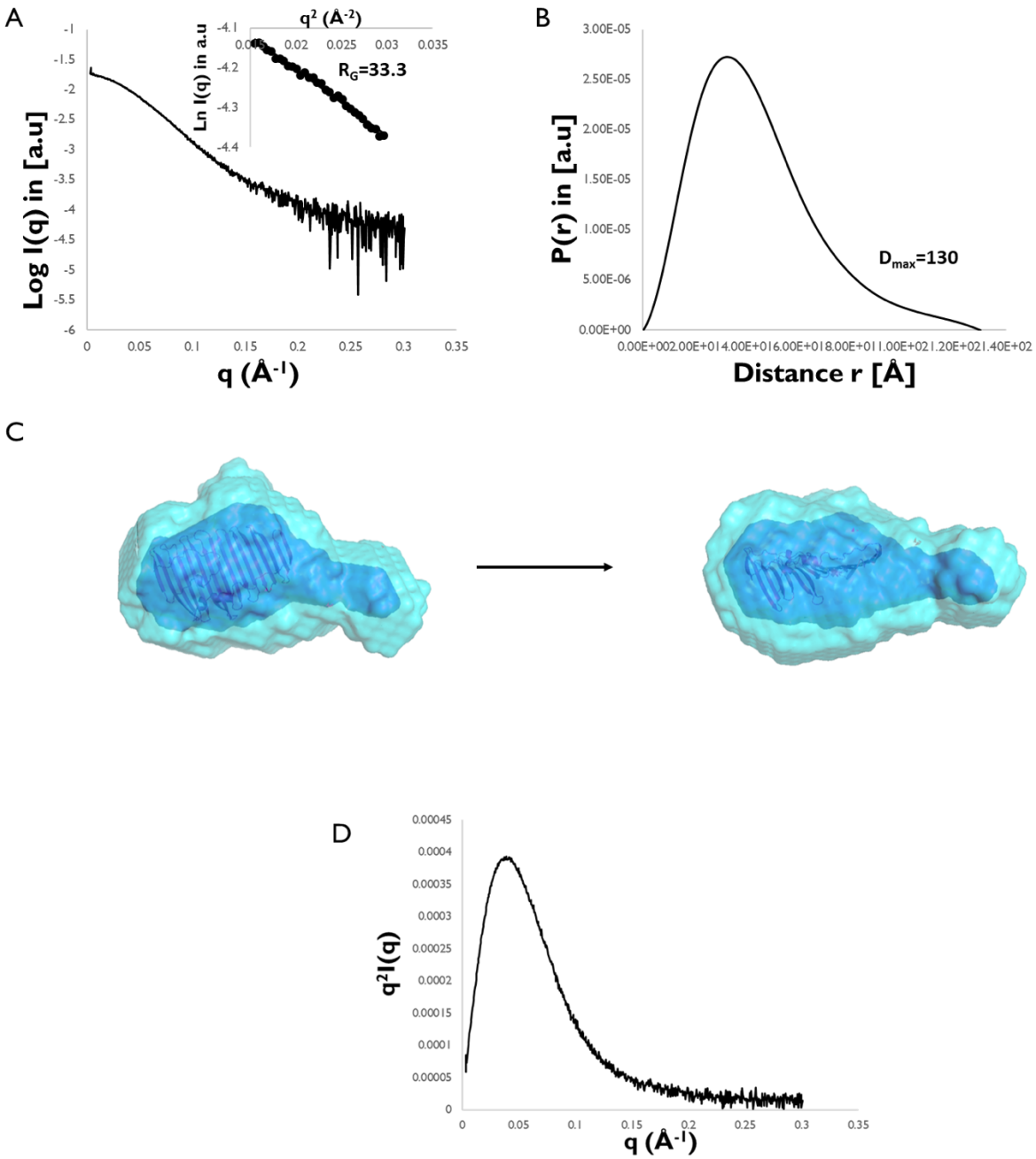


Figure 5.12: ***SAXS data and ab initio model for LESV protein with α-cyclodextrin*** **A)** Experimental data plotted as a function of the scattering vector q with the Guinier plot is illustrated as an inset **B)** illustrates the distance distribution function, and **C)** Average *ab initio* envelope predicted by DAMMIF/DAMAVER (from 10 *ab initio* LESV protein models computed by DAMMIF, shown in light blue) superposed with the most typical shape filtered by DAMFILT (dark blue). The proteins' structural parts modeled with a confidence score of pLDDT ≥ 90 (cartoon) are manually fitted into the molecular envelope. D) Kratky plot of scattering data for the LESV protein with α-cyclodextrin.

Table 5.3: SAXS structural parameters for LESV alone and in complex with α-cyclodextrin

	LESV	**LESV AlphaCD**
I(0) (cm^{-1})	23.31 +/-0.033	0.017 +/- 0.00005
R_g (Å)	36.79 +/-0.08	33.40 +/-0.19
q_{min} ($Å^{-1}$)	0.014	0.0032
qR_g max	1.18	1.01
Coefficient of correlation, R^2	0.999	0.991
P(r) analysis		
I(0) (cm^{-1})	23.82 +/-0.034	0.018 +/-0.000025
R_g (Å)	39.71 +/-0.102	34.7 +/-0.02
q range $_{min}$ ($Å^{-1}$)	0.014 to 0.4	0.0032 to 0.25
Porod volume ($Å^3$)	121099	103015
Dmax	160	130

Analysis of SAXS data analysis reveals that particles size and shape are different of the protein alone (Table 5.3) and that de interdomain dynamics effects are lower than the protein alone. Contrarily to what observed for LESV alone, the Kratky plot of the complex LESV/α-cyclodextrin shows a bell-shaped curve coming back to 0 for high values of q, typical to a folded protein.

This result shows that LESV can bind α-cyclodextrin and that this binding induces likely structural reorganisation of the N-terminal domain's unfolded regions around the β-sheet domain leading to a more globular and structurated protein that will be a better candidate for protein crystallisation.

5.4. Functional study of ESV1 and LESV

5.4.1. ESV1 and LESV interact differently with α-1,4-linked glucose polymers

To better understand the specificities of ESV1 and LESV and their interaction with starch glucans, we performed several experiments in solution. The heterogeneity of amylose and amylopectin solutions as well as their high viscosity did not allow them to be used for classical structural biology approaches (SAXS or X-ray crystallography) or SPR.

As it is known that both proteins bind to starch granules, my hypothesis was to check if the affinity was specifically directed to amylose or amylopectin. Secondly, since amylose and amylopectin have low solubility in water I have checked the affinity of the proteins to some more soluble and homogeneous derivatives (α-CD, β-limit dextrin) and that could be used for structural characterization.

To further our understanding of the specific characteristics of ESV1 and LESV and their interactions with starch glucans, we performed several experiments in solution. We studied the interaction of different glucans with ESV1 and LESV proteins using an electrophoretic mobility shift assay (EMSA), by observing the influence of increasing concentrations of these glucans on the electrophoretic mobility of ESV1 and LESV under non-denaturing conditions. This influence is estimated by the migration shift of these proteins compared to a control protein that has no affinity for the glucans.

First, ESV1 (1µg) and LESV (1µg) were analyzed on a native gel containing increasing concentrations of the glucans ranging from 0 to 0.5% amylose, and amylopectin in the separating gel (Figure 5.13), in the presence of a reference protein that has no affinity for the components of starch.

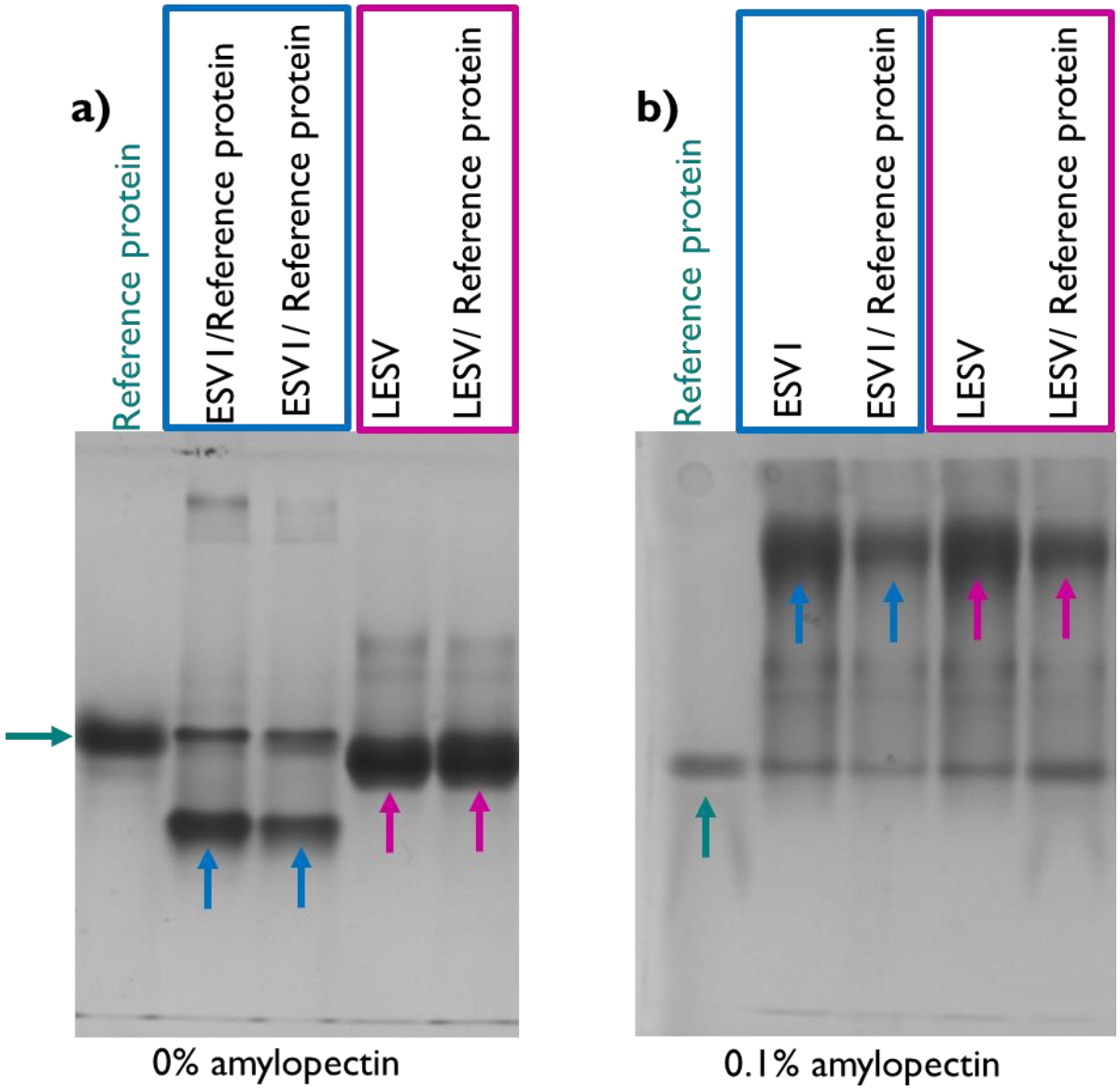


Figure 5.13: ***Analysis of ESV1 and LESV binding to amylopectin by EMSA gel with a concentration of 0%, and 0,1% of potential ligands.*** Separation of both proteins using native gels, separation gels were supplemented with 0,3%, 0.5% (not shown) of amylopectin. Here, 1µg of ESV1 and LESV were loaded per lane. As a further control 1µg of the reference protein was separated on native gels. The gels were stained with bromophenol blue. For each gel, the green arrow indicates the migration of the reference protein. The blue and the pink arrows indicate the position of the shifted ESV1 and LESV proteins respectively.

First, with amylopectin the results show a significant shift that increases with increasing concentration of amylopectin ranging from 0 to 0.5% for both proteins ESV1 and LESV as we can see in Figure 5.13. From 0.1% amylopectin, the migration of both proteins is significantly delayed compared to the control, with both bands aligning at the same level of the gel, just below the concentration gel. This observation suggests a strong specific affinity interaction between the two proteins and amylopectin.

In contrast, in native gels containing amylose in the same concentration ranges, nor or very weak differences were found for LESV and ESV1 compared to the glucan-free gel. The few differences that are seen are not increased by amylose concentration (Figure 5.14). This result suggests that LESV and ESV1 have no or very weak non-specific affinity for amylose under the conditions tested.

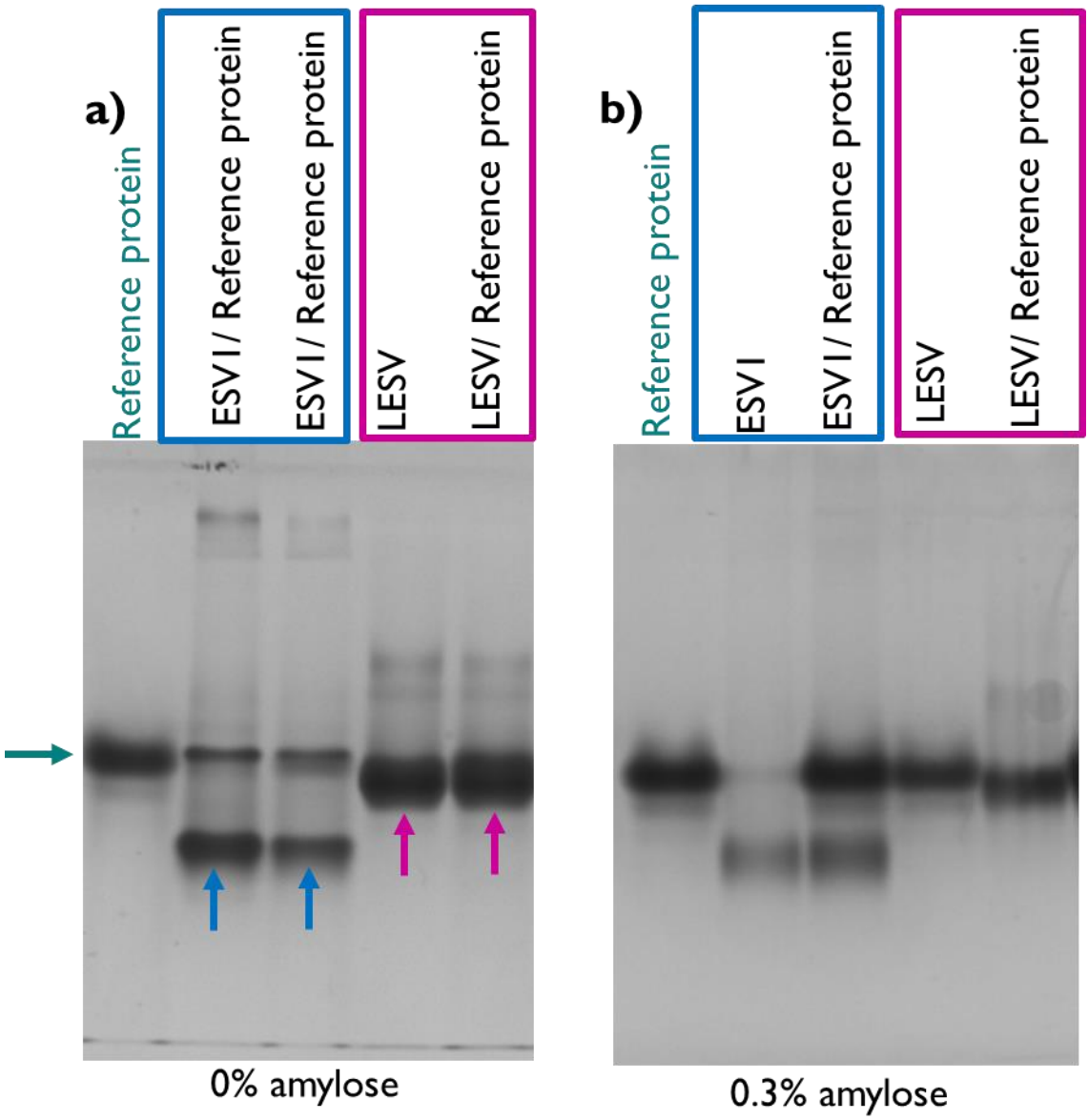


Figure 5.14: ***Analysis of ESV1 and LESV binding to amylose by EMSA gel with a concentration of 0%, 0,1% (not shown), 0,3% of potential ligands.*** Separation of both proteins using native gels, separation gels were supplemented with 0%, 0,1%, 0,3% of amylose. Here, 1µg of ESV1 and LESV were loaded per lane. As a further control 1µg of the reference protein was separated on native gels. The gels were stained with bromophenol blue. For each gel, the green arrow indicates the migration of the reference protein. The blue and the pink arrows indicate the position of the shifted ESV1 and LESV proteins respectively.

In order to understand the reasons for the differences in affinity between the two components of starch amylose and amylopectin, we decided to test the affinity of ESV1 and LESV for polyglucans with different characteristics (Figure 5.15). To do this, we repeated the same experiments with A600 (a mixture of α-1,4-linked glucose linear chains with DPs between 5 and 50 glucose residues), pullulan (glucan chains formed by maltotriose units linked by α-1,6-linkages), β-limit dextrin and finally glycogen (α-1,4-linked glucose chains poly-branched by α-1-6-linkages) in the separation gel. Analysis of the results shows that neither ESV1 nor LESV have any affinity for A600 (Supplementary Data S2) or pullulan (Supplementary Data S3). However, a significant shift in the migration was observed for ESV1 and LESV proteins when β-limit dextrin was included in the native gel (Figure 5.15).

Both amylopectin and β-limit dextrin are branched chains composed of α-1,4-glycosidic linkages with the branch chains linked to the main chain by α-1,6-glycosidic linkages, whereas amylose, pullulan and A600 are linear polysaccharides composed of α-1,4-glycosidic linkages.

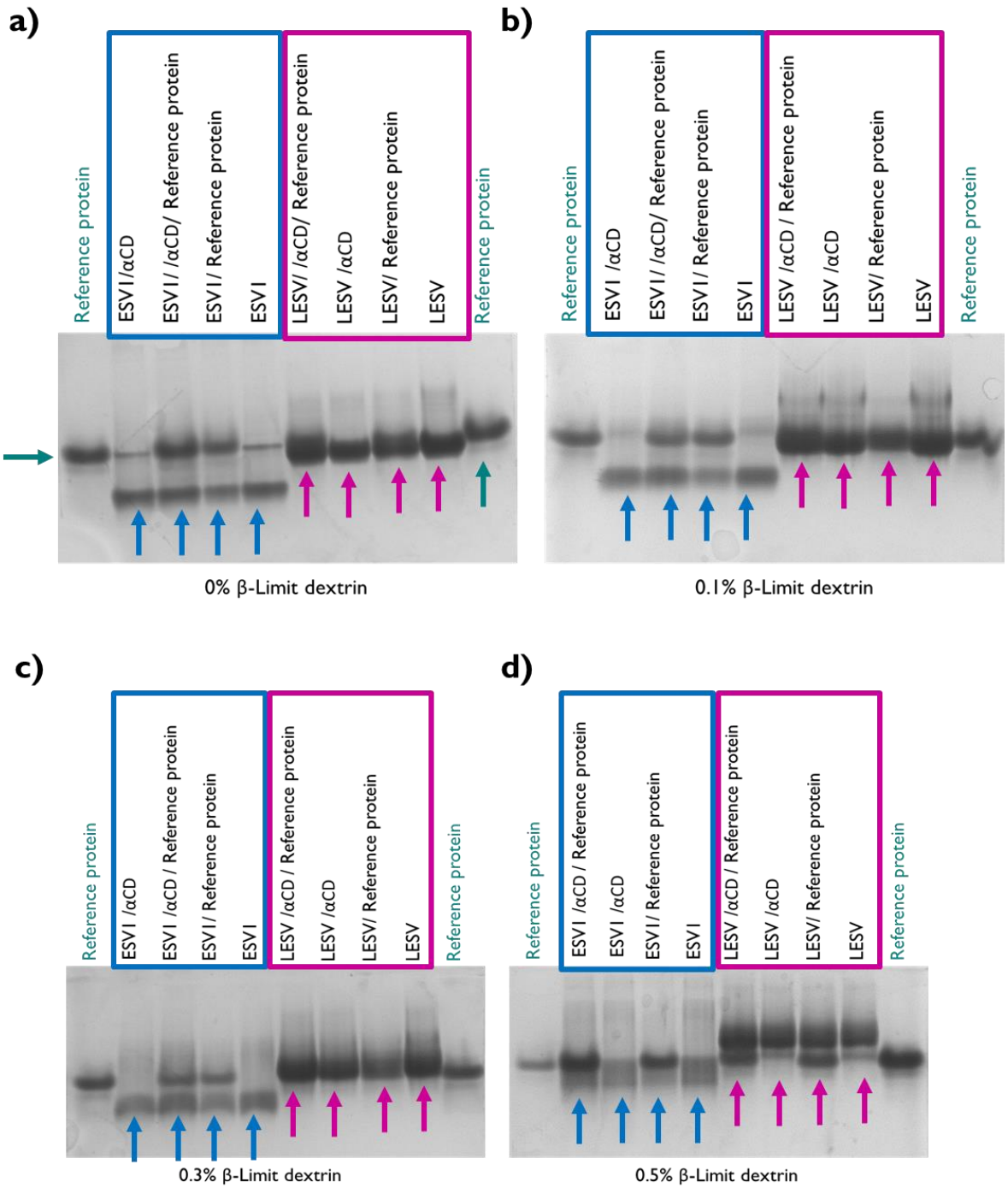


Figure 5.15: ***EMSA gels analyzing the interaction of ESV1 or LESV protein with β-limit dextrin at concentrations of 0%, 0,1%, 0,3%, 0,5%.*** Here, 1µg of ESV1 and LESV were loaded per lane. As a further control 1µg of the reference protein was separated on native gels. The gels were stained with bromophenol blue. For each gel, the green arrow indicates the migration of the reference protein. The blue and the pink arrows indicate the position of the shifted ESV1 and LESV proteins respectively.

On the other hand, in the gel containing 0.1% glycogen, a shift of the LESV protein band is clearly observed compared to the reference protein band (Figure 5.16). This shift is accentuated more and more when the glycogen concentration is increased to 0.3% and 0.5% respectively compared to the glucan-free gel, indicating an affinity of LESV for this branched glucan. This observation suggests a strong specific affinity interaction between LESV protein and glycogen.

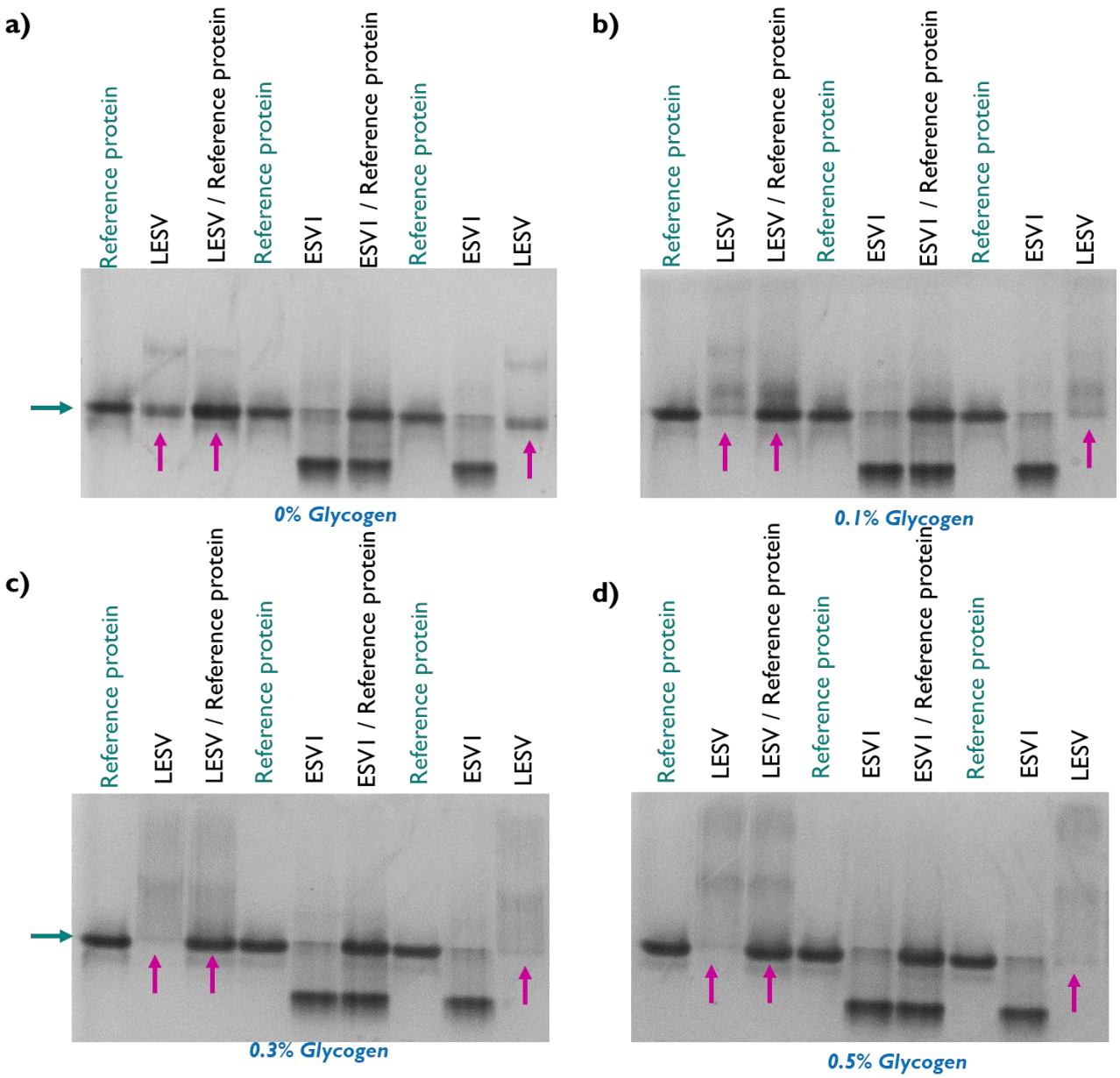


***Figure 5.16:* Analysis of ESV1 and LESV interaction with glycogen with a concentration of 0%, 0,1%, 0,3%, 0,5%.** Separation of both proteins using native gels, separation gels were supplemented with 0% to 0.5% of glycogen. Here, 1µg of ESV1 and LESV were loaded per lane. As a further control 1µg of the reference protein was separated on native gels. The gels

were stained with bromophenol blue. For each gel, the green arrow indicates the migration of the reference protein. The blue and the pink arrows indicate the position of the shifted ESV1 and LESV proteins respectively.

In contrast, with ESV1 with the same concentration ranges, no differences were found for the electrophoretic properties of ESV1 with 0.1%, 0.3% or 0.5% glycogen compared to the glucan-free gel (Figure 5.16), no delay in the migration. This result suggests that there is no interaction between ESV1 and glycogen meaning no affinity for this polyglucan.

My results show that ESV1 and LESV from *Arabidopsis thaliana* show a strong affinity toward amylopectin and a moderate affinity β-limit dextrin but not for linear chains such as amylose. This suggests that the affinity of LESV and ESV1 for starch is not due to the recognition of linear glucans but likely to the additional helices in the N-terminal domain of LESV, or to the particular 3D structure of amylopectin organized in double helices.

5.4.2. Conformational changes upon amylopectin binding to LESV protein and the appearance of α-helices

In order to analyse structural content of the ESV1 and LESV protein structures and to identify putative structural modifications in proteins' structure upon glucan binding, we have studied both proteins using circular dichroism spectroscopy which allows us to obtain information about secondary structure elements (Figure 5.17) (Kelly, Jess, et Price 2005). CD spectroscopy measures the difference in the absorption of right-handed polarized light versus left-handed polarized light which arise due to structural asymmetry. It is an excellent biophysical tool for rapidly estimating the secondary content (% helix, sheet, turns, etc.) and the tertiary structural organization of the protein using the far UV (180-260 nm) and near UV (250-330 nm). In general, the CD signal at 215 nm indicates the sheet content and the CD signal at 208 nm and 222 nm are used to estimate the helical content (Greenfield 2006).
In addition, it is particularly useful for determining and monitoring conformational changes in the proteins structure that happen upon ligand binding or upon changes in the protein environment, in our case due to protein-ligand interactions in our case.

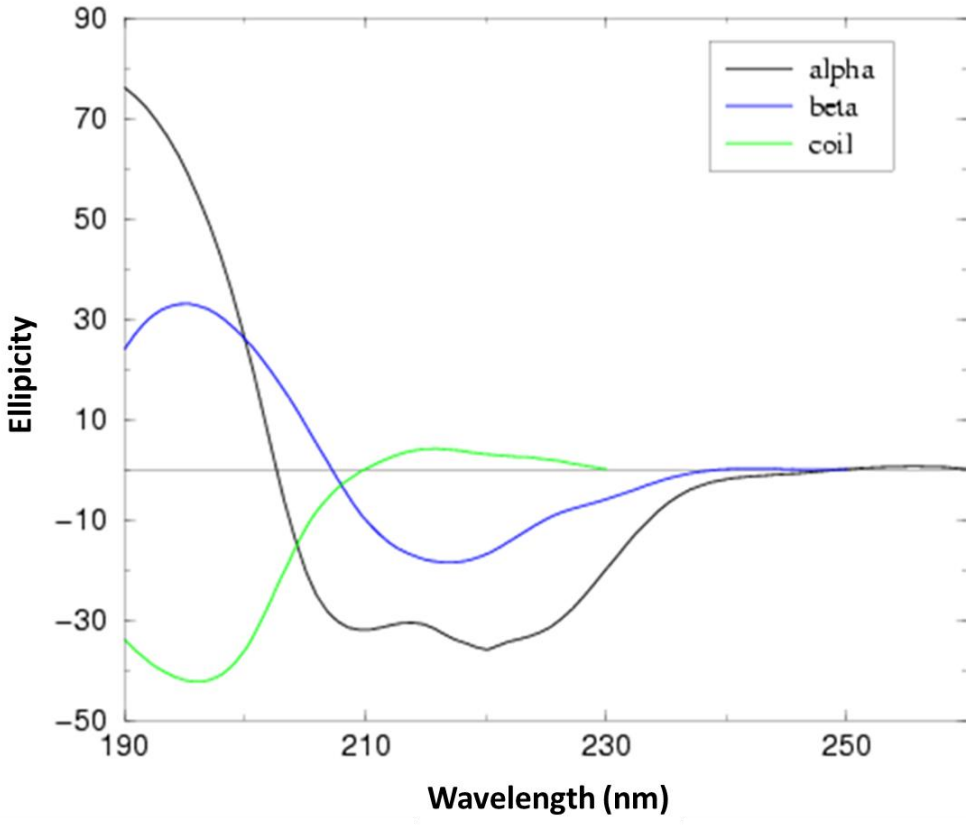


Figure 5.17: ***Illustrative CD Spectra for the three UV spectra signature of protein secondary structures***: α-helical protein (black line), β-sheet protein (blue line), and random coil protein (green line) (Wei, Thyparambil, et Latour 2014).

It is known that the disordered regions in proteins may play an important functional role. As the folding of the N-terminal regions of the LESV protein has not been predicted by Alphafold but only few helices, we wanted to estimate the structure of this domain and check if LESV undergoes conformational changes upon binding of the amylopectin due to its interaction with amylopectin.

We collected CD spectra using circular dichroism spectroscopy with a wavelength between 170 and 270nm. The result is shown in Figure 5.18. in which the secondary element composition was estimated using Bestsel (Micsonai et al. 2022). As we can see in Figure 5.18 both proteins are structured, with differences in the composition of the secondary structure. For ESV1, the pattern of CD spectra is consistent with a folded protein with a strong positive band at λ=196nm and a single negative band at λ=220nm, characteristic of all β-proteins. For LESV, the spectrum reveals a global folding of β-strands and α-helices, in which it shows a strong maximum at λ=192nm and a minimum at λ=216nm, signifying the presence of beta structures. However,

unlike ESV1, the spectrum also shows two shoulders at λ=210 and λ=222nm, demonstrating the simultaneous presence of α-helices. These results align with models generated by Alphafold that predict the folding of both proteins into a conserved C-terminal β-sheet domain and the presence of a few α-helices in the N-terminal part of LESV.

After analyzing ESV1 and LESV proteins by circular dichroism (CD) spectroscopy, we found that both proteins are predominantly β-sheets, in addition, LESV also contains α-helices (Figure 5.18), as predicted in the AlphaFold models.

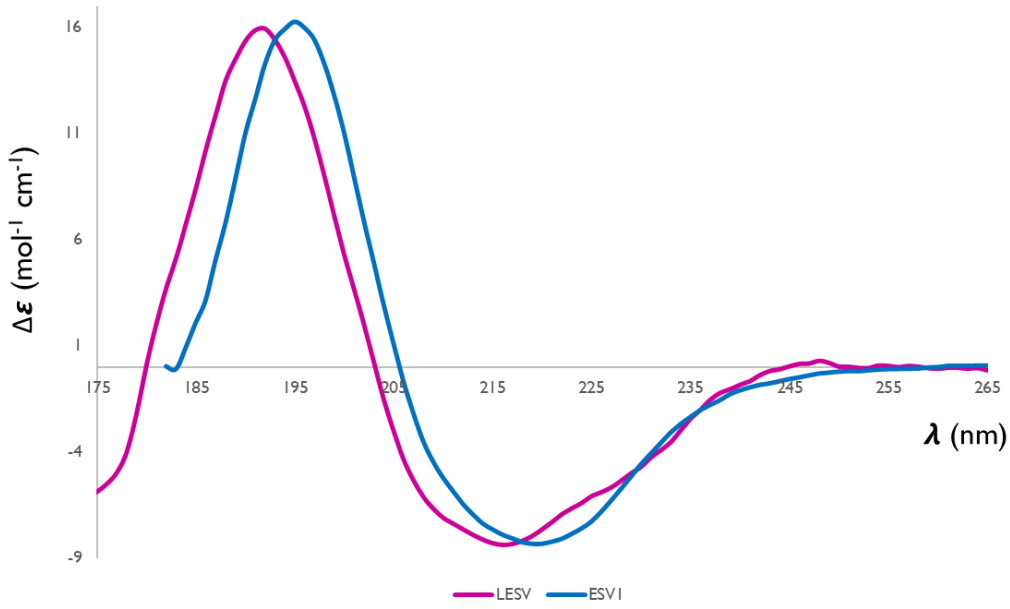


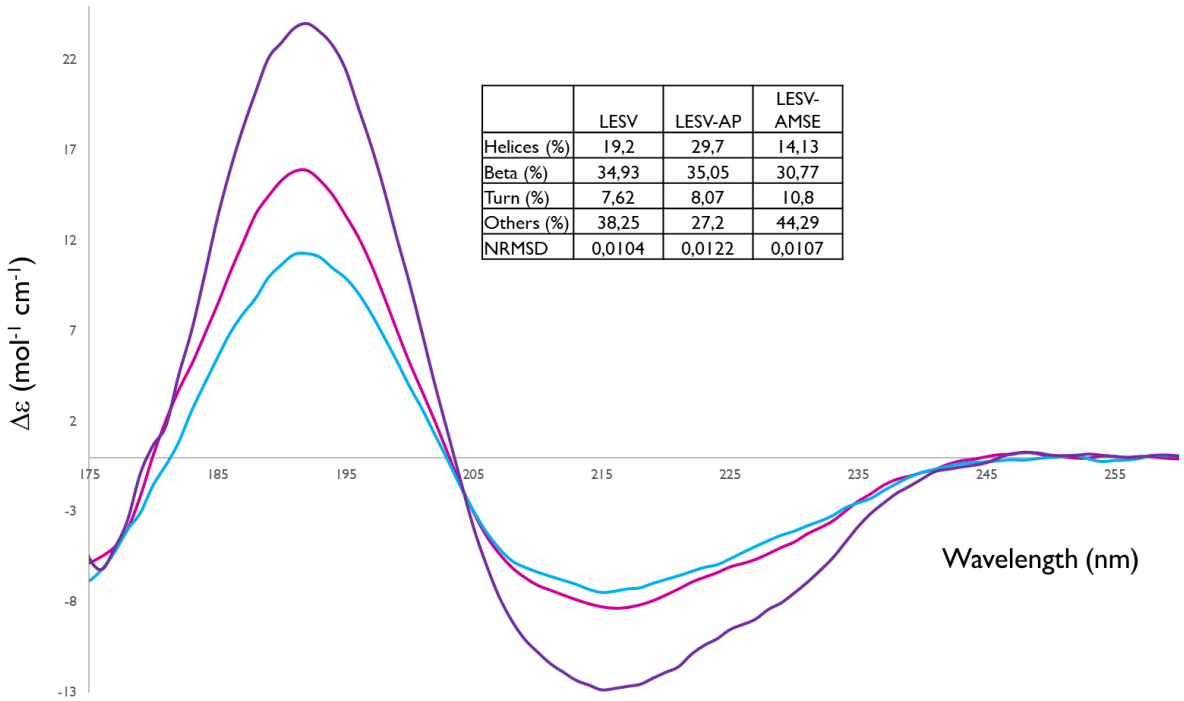
Figure 5.18: ***SRCD spectra for the proteins secondary structure content of LESV alone (violet color) and ESV1 alone (blue color).*** These spectra are averages of three separate measurements.

To further investigate the nature of the interaction between ESV1 and LESV with glucans, in particular to uncover the conformational changes of the proteins during complex formation, we used synchrotron radiation circular dichroism (SR-CD) study. Circular dichroism (CD) is the preferred technique for examining protein structure because it allows visualization of the content of secondary structural elements. CD also facilitates the study of interactions between proteins and their ligands, especially when these interactions trigger structural alterations. SRCD enhances the capabilities of standard CD spectroscopy by providing a wide spectral

range, improved signal-to-noise ratio, and accelerated data acquisition, even in the presence of absorbing elements such as buffers and salts (Hussain et al., 2012; Hussain et al., 2018).

Quantitative analysis of CD spectra also facilitates the prediction of the secondary structure content of a protein. Therefore, we compared the spectra obtained for proteins alone and in complex with amylose and amylopectin.

To examine the interactions between the proteins (ESV1 and LESV) and polyglucans (amylose and amylopectin), 4 µl of solutions of each protein were combined with 1 µl of 1% amylose or 1% amylopectin solutions and incubated for 2 hours before measurement. For each spectrum, the composition of the secondary structure elements was determined using BestSel (Micsonai et al. 2015). The values obtained were compared to those acquired for the proteins alone.



	LESV	LESV-AP	LESV-AMSE
Helices (%)	19,2	29,7	14,13
Beta (%)	34,93	35,05	30,77
Turn (%)	7,62	8,07	10,8
Others (%)	38,25	27,2	44,29
NRMSD	0,0104	0,0122	0,0107

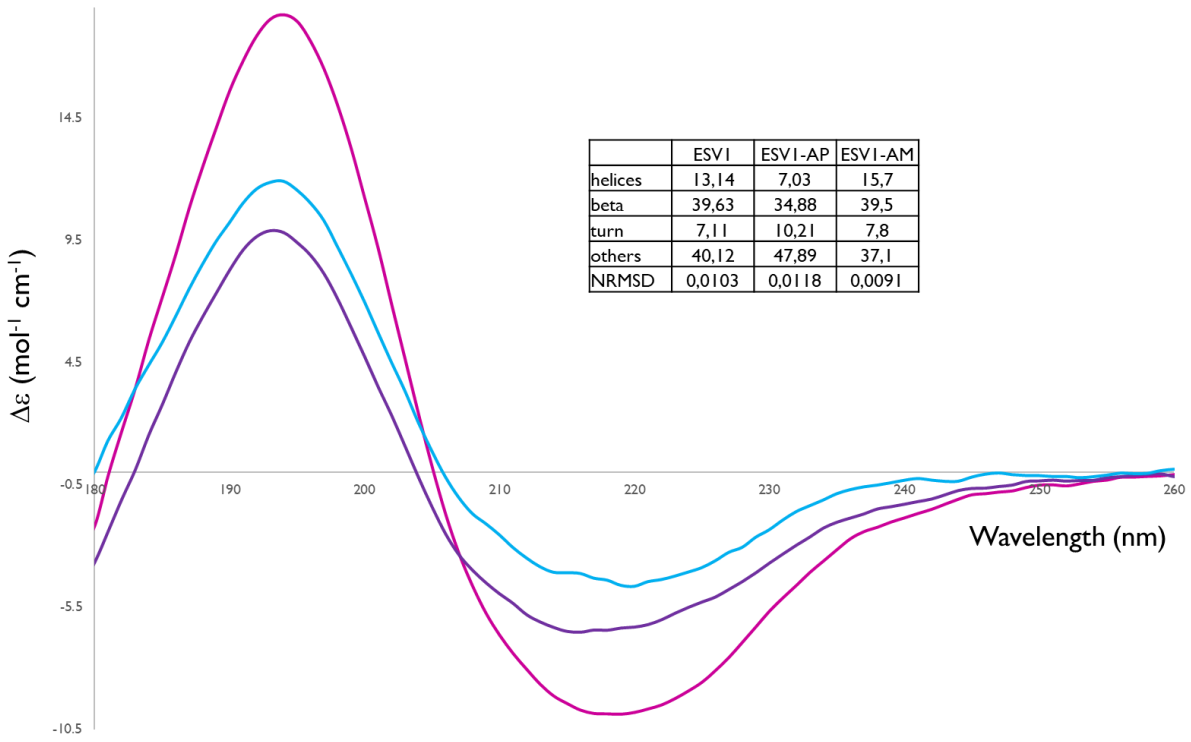
***Figure 5.19:* Superposition of the SR-CD spectra of LESV alone (Pink) and when mixed with amylopectin (violet) or amylose (blue) solutions.** The composition in secondary structural elements evaluated by BestSel are in inset.

The spectrum obtained for the LESV/amylose mixture largely reflects that of LESV alone, with peaks of slightly lower amplitude. This implies that the presence of amylose induces little or no conformational changes in the protein (Figure 5.19).

In contrast, the spectrum obtained for the LESV/amylopectin mixture shows much larger peaks than those obtained for the protein alone or in the presence of amylose. The positive peak at λ=192 nm is 60% larger than for the protein alone, indicating more pronounced structuring of the protein. At λ=208 and λ=215 nm, the molar ellipticity values, indicating the presence of α and β structures, are 50% lower than those observed for the protein alone or in the presence of amylose. More intriguingly, the molar ellipticity at λ=222 nm, indicating the presence of α-helices, is significantly lower (70%) than that observed for the protein alone or in the presence of amylose. This result demonstrates that amylopectin, unlike amylose, induces a significant conformational change in the LESV protein. This change suggests the formation of additional α-helices upon amylopectin binding.

To identify and quantify the conformational changes that LESV undergoes in the presence of amylose and amylopectin, the composition of secondary structural elements was analyzed using BestSel (Micsonai et al. 2015). The results, shown in the inset of Figure 5.19, corroborate the analysis of the CD spectra. The composition of the β-strands and turns is equivalent whether LESV is alone or in the presence of amylose or amylopectin. This indicates that the beta domain structure is not altered by the presence of polyglucans. LESV in the presence of amylose appears to have a slightly lower number of α-helices and strands than the protein alone. Conversely, a higher number of α-helices is observed when the protein is in the presence of amylopectin. This result confirms that the LESV protein interacts with amylopectin and that this interaction induces the formation of α-helices, most likely localized in the N-terminal domain.

The same analysis was performed for ESV1 in figure 5.20.



	ESV1	ESV1-AP	ESV1-AM
helices	13,14	7,03	15,7
beta	39,63	34,88	39,5
turn	7,11	10,21	7,8
others	40,12	47,89	37,1
NRMSD	0,0103	0,0118	0,0091

***Figure 5.20:* shows the superposition of the SR-CD spectra of ESV1 alone (pink) and when mixed with amylopectin (violet) or amylose (blue) solutions.** The composition in secondary structural elements evaluated by BestSel are in inset.

The spectrum obtained for the ESV1/amylose mixture largely reflects that of ESV1 alone, with peaks of slightly lower amplitude (Figure 5.20).

The analysis reveals broadly similar spectra for ESV1 in the presence of either amylose or amylopectin, with what appears to be a slightly lower degree of structuring for the protein alone. BestSel analysis of secondary structural element composition, as shown in the inset of Figure 5.20, indicates a similar composition for ESV1 alone or in the presence of amylose. However, there is a slight reduction or modification of some strands and α-helices when the protein is in the presence of amylopectin. This result suggests that ESV1 interacts with amylopectin without inducing significant conformational changes.

5.4.3. Glucan binding affect the Melting temperature (TM) of LESV but not ESV1

For LESV protein

Detection of ligand binding can be enhanced by SR-CD controlled thermal denaturation studies. This technique is more sensitive than simple spectral differences because it allows the discovery of interactions that do not induce structural changes in proteins. Using this approach, it is possible to determine the melting temperature (Tm) of the unfolding transition. The variation of the CD signal was measured as a function of temperature for isolated proteins and in the presence of amylose or amylopectin, under the same conditions as for the constant temperature spectra.

Figure 5.21 shows the results obtained for LESV. Upon denaturation of the protein, two stable states of the protein were observed, whether isolated (Figure 5.21A) or in the presence of amylopectin (Figure 5.21B) or amylose (Figure 5.21C). This was indicated by an isosbestic point at λ=210 nm. As temperature increases, the intensity of the positive peak at λ=192 nm decreases and a hypsochromic effect (shift to lower wavelengths) is observed, reflecting significant alterations in the overall structure of the protein. The magnitude of the peak between λ=210 and λ=225 nm decreased sharply, indicating a reduction in α and β structures. Specifically, the decrease at λ=222 nm reflects a greater decrease in α-helix structures in the LESV amylopectin mixture. Simultaneously, a new negative band appeared at λ=203 nm, suggesting the formation of random coils and thus confirming denaturation of the protein. However, Figure 5.21D the spectrum in the presence of α-cyclodextrin shows a behavior which is almost identical to LESV alone, but even that this binding doesn't lead to the formation of new helices (CD), it is sufficient to stabilize disordered regions around the beta domain that we have seen it with the SAXS experiment above.

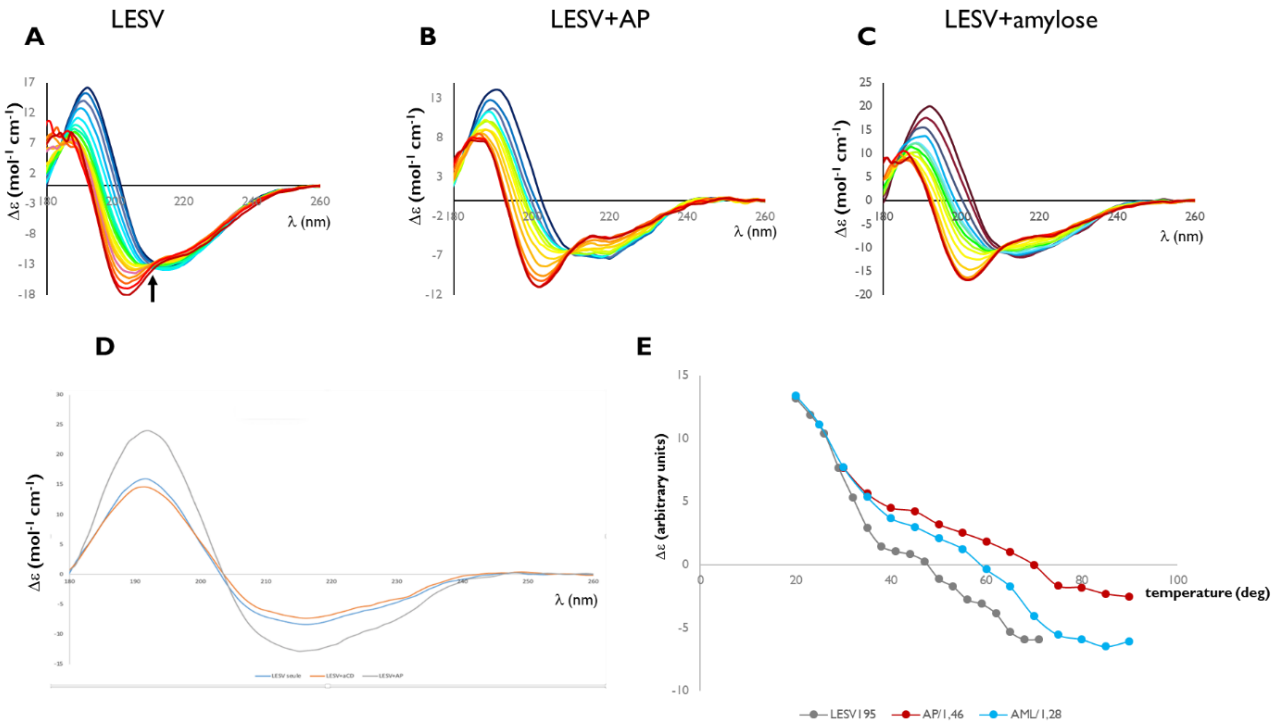


***Figure 5.21:* Thermal denaturation of LESV followed by SR-CD. Each plot represents consecutive scans on the protein collected at a set of temperature between 20 to 90°.** Scans are colored in a gradient from dark blue (first temperature) to dark red (last temperature) for a) LESV alone, b) LESV with amylopectin and c) LESV with amylose. d) shows the superposition of the SR-CD spectra of LESV alone (blue) and when mixed with amylopectin (grey) or α-cyclodextrin (orange) solutions. The composition in secondary structural elements evaluated by BestSel are in inset. e) thermal denaturation of LESV followed at λ =190nm. Curves corresponding of proteins alone, with amylopectin or amylose are colored grey, red and cyan respectively.

To assess the impact of the presence of glucans on protein stability, we determined the Tm of the mixtures by tracking the evolution of molar ellipticity as a function of temperature at λ=195nm. The curves obtained were normalized and are presented in Figure 5.21E. The experiment demonstrates that LESV denaturation is decelerated in the presence of amylose and even more so in the presence of amylopectin, with Tms of about 55°, 60°, and 65°, respectively. This result indicates that the presence of amylose and amylopectin stabilizes LESV, suggesting an interaction of the protein with both starch components. However, combining this result with

EMSA result, the interaction with amylopectin is likely much specifics than the interaction with amylose

For the ESV1 protein

The experiment was repeated with ESV1, and the results are analyzed in Figure 5.22D. It was noted that two isostatic points were present for the protein alone, with a first isostatic point at λ=210nm as observed for LESV and a second at λ=197nm. The same profile was observed for the protein in the presence of amylopectin. This observation confirms that denaturation of the protein alone or in the presence of amylopectin occurs in two different states. Conversely, only one state is observed in the presence of amylose, where only the isosbestic point at λ=201nm is present, suggesting that the presence of amylose may stabilize ESV1 in the initial denaturation step.

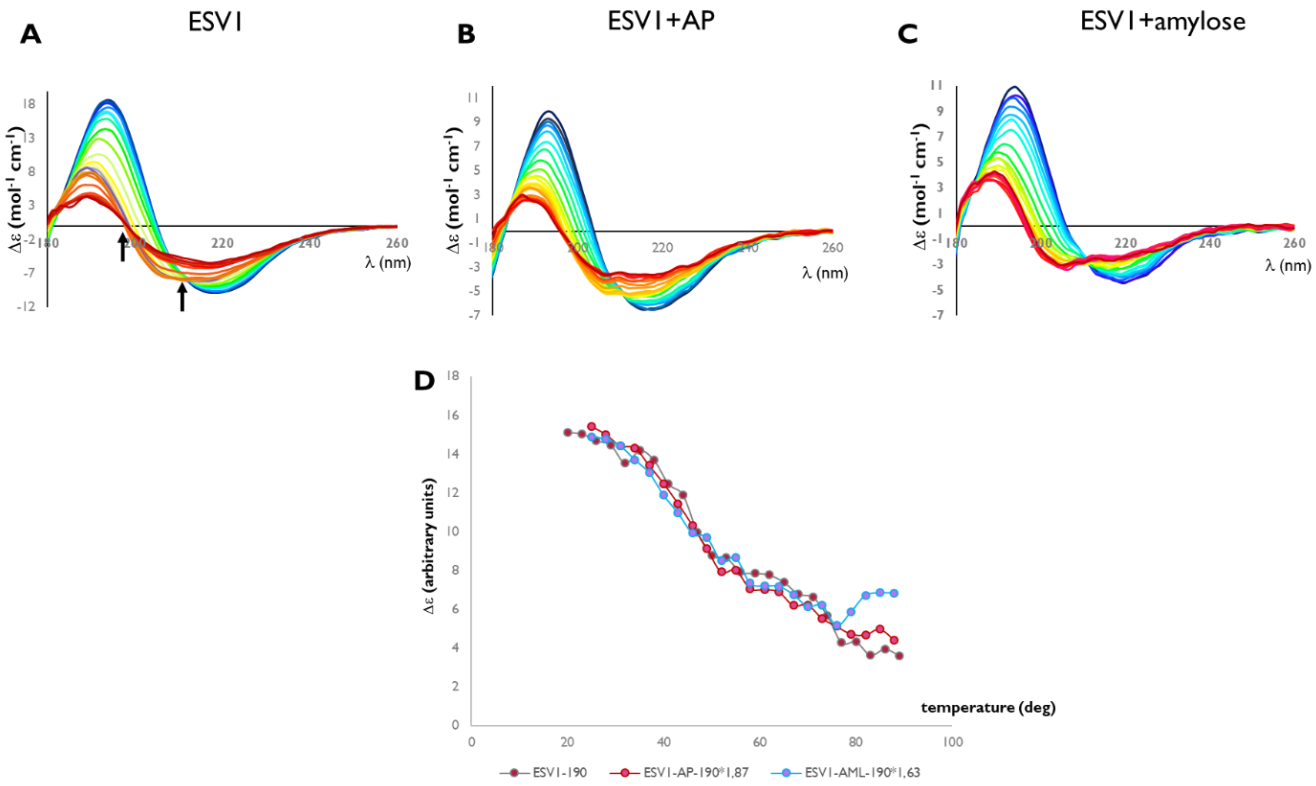


***Figure 5.22:* Thermal denaturation of ESV1 followed by SR-CD. Each plot represents consecutive scans on the protein collected at a set of temperature between 20 to 90°.** Scans are colored in a gradient from dark blue (first temperature) to dark red (last temperature) for a) ESV1 alone, b) ESV1 with amylopectin and c) ESV1 with amylose. d) thermal denaturation of

ESV1 followed at λ =195nm. Curves corresponding of proteins alone, with amylopectin or amylose are colored grey, red and cyan respectively.

In all assays, with increasing temperature, the peak observed at λ=193 nm decreases in intensity and undergoes a hypsochromic effect before stabilizing around λ=180nm. After the first denaturation step, this peak continues to decrease in intensity without undergoing a hypsochromic shift, signifying a change in the overall structure of the protein. The intensity of the negative peak observed at λ=220 nm decreases and also undergoes a hypsochromic shift (which stabilizes around λ=212nm) until the first denaturation state is reached at 43, 43, and 46 degrees for ESV1, ESV1/amylopectin, and ESV1/amylose, respectively. At higher temperatures, a different behavior is observed for the ESV1/amylose mixture. The magnitude of the negative peak stops decreasing, while a negative peak around λ=202 nm emerges, a peak that is not observed for ESV1 alone or in complex with amylopectin. For ESV1 alone and in complex with amylopectin, a second state is reached at about 75 degrees. Between the first and second states, the positive peak at λ=180nm continues to decrease in intensity without undergoing a hypsochromic change. Similarly, the peak at λ=212nm continues to decrease in magnitude without shifting.

Regarding the evolution of molar ellipticity as a function of temperature at λ=190nm, the obtained curve has been normalized and is presented in Figure 5.22D. The denaturation curves show the same pattern and can be superimposed, with an inferred Tm of about 50°C for ESV1. Therefore, unlike LESV, the presence of amylose or amylopectin does not impact the thermostability of ESV1, despite a demonstrated interaction with amylopectin.

5.4.4. ESV1 and LESV are found distributed across the entire surface of the starch granules.

To validate the results obtained in solution, we examined the binding of ESV1 and LESV to starch granules using UV fluorescence microscopy. This method, which allows the visualization of proteins by the fluorescence of their aromatic residues on starch granules, was used and described in detail in (Tawil et al. 2011).

Starch granules have a high concentration of GBSS, the enzyme that catalyzes the synthesis of amylose. For the experiment, we chose corn starch granules that had the GBSS gene deleted.

This allowed us to distinguish the fluorescence of the granule from that of the protein under investigation. The measurement was performed simultaneously in visible light and with excitation at λ=310 nm, which stimulates the tryptophans to be excited.

For 3D visualization, Z-scans (along the optical axis) were acquired with a step size of 300 nm over a Z range of 40 µm under the control of the µManager. However, a Z-scan refers to a method of acquiring multiple optical sections along the Z-axis (depth) of a sample. It involves systematically moving the focal plane of the microscope through the sample at different depths while capturing images. This technique allows us to obtain a series of images at different focal planes and reconstruct a three-dimensional representation of the sample.

We prepared a mixture in which 65µl of WAXY starch grains (300mg/ml) were incubated with 2µl of ESV1 (4mg/ml) or 2µl of LESV (10mg/ml) for 2 hours before measurement. Then a second mixture was prepared from the first by mixing 2.5µl of the first mixture with 2.5µl of the buffer, and then examined under the microscope.

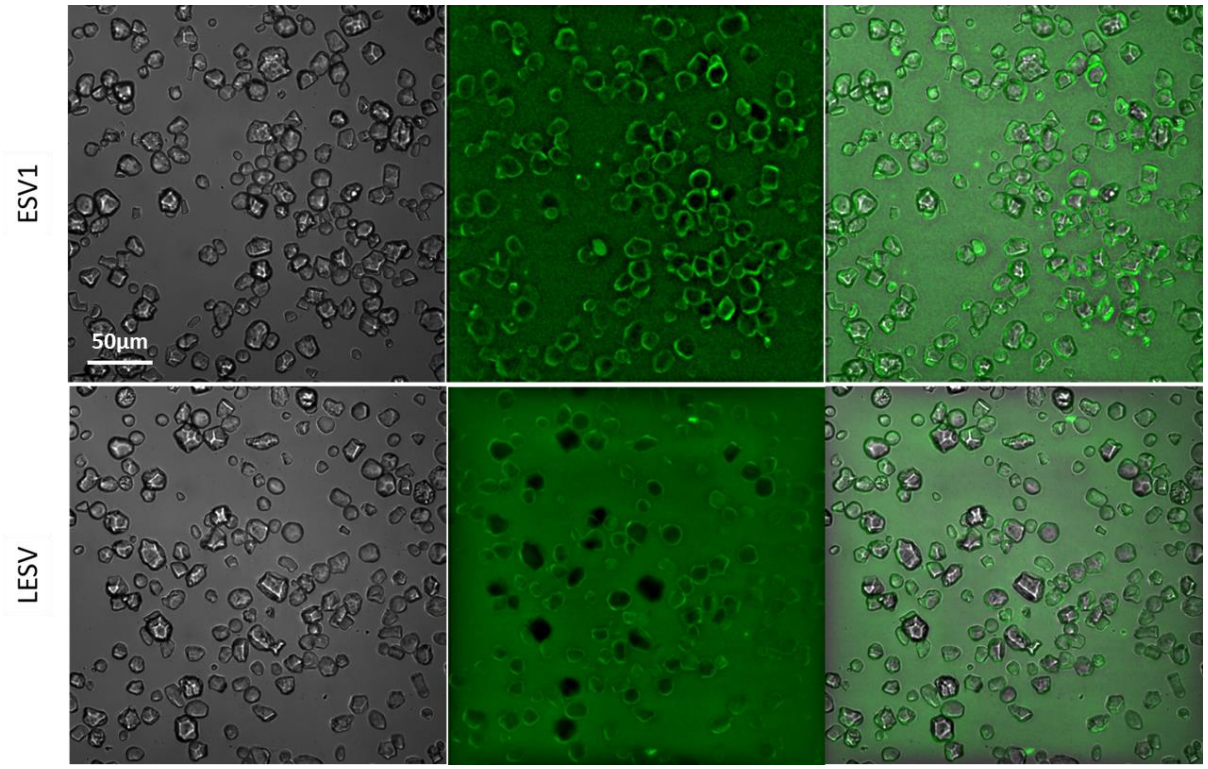


Figure 5.23: ***Starch granules seen under fluorescence microscopy*** showing the starch granules alone (to the left), starch granules with ESV1 or LESV (in the middle), and the superposition of both images to the right.

We secured the emission spectrum by using a filter to select a wavelength range between 329 and 351 nm. We performed two control experiments: the first with starch granules alone to confirm the absence of fluorescence, and the second with starch granules in the presence of BSA. As we can see in Figure 5.23,

As we can see in Figure 5.23, there is fluorescence around the starch granule in the image on the right compared to the control image on the left. This fluorescence indicates the presence of the ESV1 or LESV protein bound to the entire surface of starch granules demonstrating that both proteins are able to recognize and bind to the crystalline form of amylopectin. This indicates that both proteins are present on the surface of the starch granules, as we can see in the third image on the right.

5.4.5. The tryptophan-rich C-terminal domain of ESV1 and LESV is capable of binding two double helices of amylopectin on one side of the β-sheet

To deepen our understanding of the interaction between the novel sugar-binding domain (SBD) discovered on ESV1 and LESV, we simulated the complex with the conserved C-terminal domain of the proteins and a pair of amylopectin double helices using GOLD (Jones et al. 1997). First, we used the LESV model with the β-sheet and the well-positioned α-helix as described in section 5 and Figure 5.7.

In both protein models, we observed a polarity on either side of the C-terminal β-sheet, one side of which was occupied by a long helix (in LESV) or a long loop (in ESV1). These elements partially mask the aromatic or acidic amino acid residues that could interact with amylopectin (Supplementary data S4). In this study, we focused on the interaction of amylopectin only with the accessible side of the β-sheet. We did not consider the N-terminal domain, which has not yet been predicted.

We first performed a docking calculation involving a protein molecule and an amylopectin double helix centered on an aromatic stripe of LESV on the accessible side of the β-sheet. We obtained a satisfying result where a double helix of amylopectin binds to the β-sheet and aligns well with the aromatic stripe. We repeated this process with the protein binding a single chain of amylopectin, targeting the second aromatic stripe. We again obtained a good result, as shown in Figure 5.24. Here, two amylopectin double helices bind along the aromatic stripes and align

parallel to each other. Any attempts to add a third chain of amylopectin did not lead to any result.

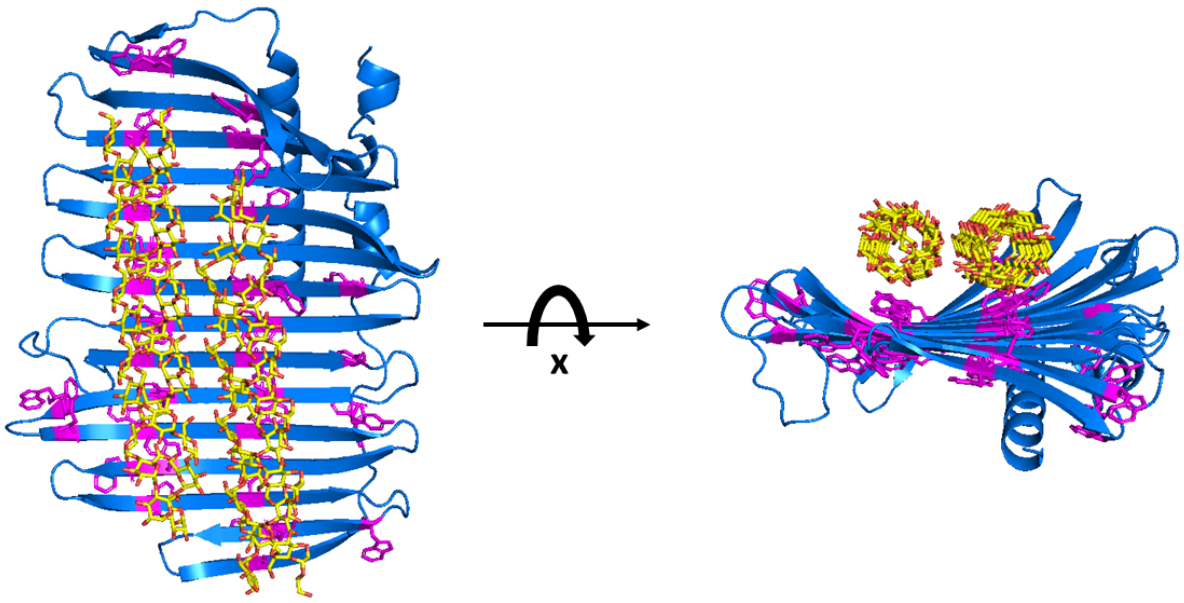


Figure 5.24: ***The molecular model shows the complex formed between the C-terminal domain of LESV and the double helices of amylopectin.*** The protein chain is cartoon-like and colored in blue. Aromatic residues are highlighted in magenta and their side chains are shown as sticks. The double helices of amylopectin are also shown as sticks, but color-coded according to atom type.

In this model, the amylopectin double helices are parallel and separated by approximately 10 Å (between the axes of the double helices). The arrangement of the double helices relative to each other in the calculated model is consistent with the configuration of amylopectin molecules reported for type A starch, which forms the starch granules in plant leaves (A. Imberty et al. 1988). This finding confirms that the C-terminal domain conserved in ESV1 and LESV is ideally suited for interaction with the semi crystalline form of amylopectin found in starch granules.

The amylopectin molecules interact with the protein domain *via* several typical protein-sugar interactions, as expected from the analysis of the primary sequence of the β-domain and the distribution of conserved amino acids in the AlphaFold model. The glucose rings interact via

hydrophobic stacking with the aromatic rings of the beta domain along the double helices. The conserved acidic residues in LESV and ESV1 also play an important role in the interaction by forming hydrogen bonds with the hydroxyl groups of the glucose residues.

5.5. X-Ray crystallography trials of ESV1 and LESV proteins

The goal of this project is to solve at the atomic level the molecular structure of LESV and ESV1 from *Arabidopsis thaliana* by an integrative structural biology approach, combining mainly X-ray crystallography (for stable domains and/or complete proteins), SAXS (to analyze disordered regions and dynamic domain organization) and CD (to estimate the secondary structure composition of the protein).

Since ESV1 and LESV having no known homologs, we decided to start our study by studying their structure by SAXS with the aim to identify their different domains and their organization in order to obtain more information on the structure of proteins in solution.
In order to analyse the structural organization and complete the AlphaFold predictions of the structures of the ESV1 and LESV proteins, we have used small angle X-ray scattering (SAXS), and circular dichroism.

Proteins crystallization

After confirming the quality of our protein samples, I have started crystallization trials with LESV, ESV1. To do that I used commercial kits (Table 5.4) and the Nano dispenser Mosquito robot using a protein concentration between 10 till 25 mg/ml for the LESV protein and a concentration between 4 & 8 mg/ml for the ESV1 protein.
Different crystallization commercial kits were used to perform vapour phase diffusion experiments in a sitting drop. It consists in establishing an equilibrium between the drop containing the protein/precipitant mixture and the well containing only the precipitant in a closed chamber. These kits composed of different solutions containing precipitating agents in different pH with different additives in order to influence the solubility of proteins. So, different parameters were varied to promote crystal formation, namely protein concentration, salt concentration, presence of cofactors and presence of substrate analogues (Table 5.4).

We didn't use amylose and amylopectin in our crystallographic assay because of the heterogeneity and high viscosity of the two components of starch. On the other hand, we used α-cyclodextrin and β-limit dextrin for our crystallographic assay based on our SAXS and EMSA results, which show that the binding of α-CD induces a structural change in the structure of LESV, leading to a stabilization of its structure, and that it could be a good candidate for co-crystallization with LESV.

Tests are performed using 11 different commercial kits (BCS Screen, CRYO, JCSG Plus, MIDAS, MemGold, PACT Premier, PGA Screen, Proplex, 3D structure, Memplus, MemGoldMeso) with and without ligands by mixing the protein solution with a precipitation solution from the kit at a concentration of 5.13 mM for the α-cyclodextrin with 200 µM LESV protein or with 200 µM ESV1 protein for example at an incubation temperature of 20°C as shown in Table 5.4:

Table 5.4: ***represents the different commercial kits used in crystallizations*** with varying concentrations of both proteins LESV and ESV1

Commercial Kits	Incubation T°C	Proteins	Ligand	Supplier
Cryo 1& 2	20°C	LESV ESV1	α-Cyclodextrin	Molecular dimensions
PGA Screen	20°C	LESV ESV1	α-Cyclodextrin	Molecular dimensions
PGA Screen	20°C	LESV ESV1	β-limit dextrin	Molecular dimensions
Cryo 1&2	20°C	LESV ESV1	β-limit dextrin	Molecular dimensions
Cryo 1&2	20°C	LESV ESV1	α-Cyclodextrin	Molecular dimensions
Cryo 1&2	20°C	LESV ESV1	-	Molecular dimensions
PGA Screen	20°C	LESV ESV1	-	Molecular dimensions
BCS	20°C	LESV ESV1	β-limit dextrin	Molecular dimensions
BCS	20°C	LESV ESV1	α-Cyclodextrin	Molecular dimensions

All the kit used are from molecular dimensions on this website https://www.moleculardimensions.com/products/all-crystallization-screens.

Despite the diversity of constructions used and the numerous conditions tested, very few conditions led to the formation of crystals.

All these conditions contain PEG and salts in the composition in a pH range 6,5-8,5. I decided to started experiments around these conditions in order to reproduce the crystals obtained by extending the concentration of precipitating agents, & by testing the effect of the pH. These crystals obtained were difficult to reproduce them again except for the condition of C1 (kit Cryo

1&2) and F7 (kit cryo 1&2). I used α-cyclodextrin and β-limit dextrin as ligands for ESV1 and LESV.

The crystals have been mounted in cryo-loop and frozen in liquid nitrogen for data collection on beamline Proxima 2 at synchrotron Soleil (Gif-sur-Yvette). I tested them at the synchrotron Soleil but all of them are salt.

6. Discussions

The aim of my project was to study the structure/function of ESV1 and LESV *Arabidopsis thaliana* in starch metabolism by solving their atomic structure and analyze their interaction with starch glucans.

By Alphafold, we obtained two molecular models for LESV and ESV1 that show that both proteins share a common β-sheet at the C-terminal tryptophane rich regions of the proteins and unpredicted fold for the N-terminal domains (Figure 5.24). This β-sheet is an original structure adapted for the fixation of long chains of glucans as we suspected by looking to the primary structure specifically by looking to the tryptophan rich region where the tryptophan residues are arranged along two lines parallel to each other, where we can manually fit a double-helix structure of amylopectin chain which is in agreement with my previous results on LESV and ESV1 interaction with amylopectin and β-limit dextrin. Both proteins are composed of a structured domain extended by a more dynamic elongated domain. And this structured domain has been confirmed by the SAXS of At ESV1.

Our findings indicate that the tryptophan-rich regions of both ESV1 and LESV proteins may serve as an uncharacterized carbohydrate interaction surface with the ability to bind structurated polyglucans.

The presence of the β-sheet structure obtained by alphafold has been confirmed by the circular dichroism (CD) and SAXS on ESV1 (composed mainly by the predicted β-sheet). In the same way LESV protein is folded predominately into β-sheets. The N-terminal region of LESV is not predicted by alphafold (except for a few helices), but we have shown using SAXS experiment that this region is not completely disordered although dynamic and organized around or close to the β-sheet.

The binding of LESV with amylopectin using circular dichroism induces a conformational change which show larger peak, in which has a positive peak at λ=192 nm approximately which suggest the formation of additional helices upon the binding of amylopectin and could these double helices align others helices propagating a phase transition resulting in crystalline lamella which highlight the importance of the presence of the amylopectin.

I showed that the interaction of the protein with starch performed *via* a specific interaction with amylopectin and β-limit dextrin but not for linear chains such as amylose, which is different

from what is published (Singh et al. 2022). Singh and his collaborators have demonstrated that LESV binds more strongly to amylose more than amylopectin. The difference may be due to different approaches: we worked on soluble molecules while they worked on starch granules. My study suggests that the affinity of LESV and ESV1 for starch is not due to the recognition of linear glucans but likely to the particular structure of amylopectin in double helices.

ESV1 and LESV bind specifically to amylopectin

In this study, we investigated the specific interaction of ESV1 and LESV with the different components of starch. The complexity and heterogeneity of the amylopectin and amylose macromolecules, which often form dense glucans in solution, limited the methods we could use. We initially chose an EMSA approach, which allowed us to study the behavior of both proteins with respect to each polyglucan.

Previous studies on the affinity of ESV1 and LESV for starch components (Malinova et al. 2018; Singh et al. 2022) focused on the interaction of ESV1 and LESV with insoluble starch glucans and/or different starch granules from mutant plants with altered starch composition. These studies suggest that ESV1 and LESV interact with starch granules, each with a specific affinity for amylopectin and amylose, respectively, but that this affinity does not depend on the protein/glucan ratio. However, these results are not fully consistent with the recent characterization of LESV and ESV1 and our current results (C. Liu et al. 2023).

To resolve this discrepancy, we performed a more detailed analysis of the interaction between ESV1 and LESV, taking into account various parameters. By examining the migration profiles of both proteins in the presence of different glucans, we were able to show that both proteins have a high affinity for amylopectin, an affinity that increases with the concentration of this glucan. We also found no differences in the behavior of the two proteins towards amylose. In particular, we did not observe any migration delay associated with amylose or other long linear polyglucans using this method.

Structural analysis of ESV1 and LESV in the presence of amylose and amylopectin by SR-CD revealed that only amylopectin induced significant conformational changes in LESV upon interaction, causing disordered regions of the protein's N-terminal domain to structure into α-helices. We did not observe any conformational changes in ESV1, which may be due to its reduced N-terminal domain and largely composed of a highly structured C-terminal conserved

domain that is unlikely to undergo conformational changes. However, we found that the presence of amylose moderately increased the TM of LESV, although less than that of amylopectin. All these observations suggest that while amylose can interact with LESV and ESV1, but this interaction, unlike that with amylopectin, is not highly specific as it doesn't induce any migration shift in EMSA but likely related to the substantial presence of aromatic amino acids in the structure of both proteins.

As the experiments we carried out were with amylopectin and amylose molecules in their solubilized form (obtained by chemical treatment), we wanted to verify that ESV1 and LESV were able to interact with amylopectin in its crystallized form. The results we obtained in fluorescence microscopy with maize waxy starch granules, which contain no amylose, clearly showed that ESV1 and LESV interact directly with amylopectin starch granules. In fact, the two proteins, identified by their fluorescence, accumulate on the entire surface of the starch granules.

LESV is able to bind to amylopectin during its biosynthesis

The two proteins, ESV1 and LESV, exhibit different behaviors toward glycogen, with only LESV demonstrating the ability to interact with glycogen. This finding is interesting for several reasons.

First, the difference in affinity of ESV1 for amylopectin and amylose does not appear to be related to the presence of branch points, but rather to the three-dimensional structure of the double helices. Second, a comparison was made between the structure of glycogen and amylopectin during their biosynthesis, before isoamylases remove the excess branch points (S. Ball et al. 1996). LESV would act either directly upstream or simultaneously with isoamylases during the biosynthesis of amylopectin chains. LESV could support the phase transition of double helices, although their number of branching points is not optimized.

The fact that LESV has an affinity for glycogen, which has a structural similarity to amylopectin chains prior to the action of isoamylases, is consistent with the hypothesized function of LESV. This is also consistent with the proposed role of ESV1, which would act downstream of LESV to stabilize newly formed starch granules and has an affinity only for the amylopectin chains in the crystalline phase of starch granules.

Both LESV and ESV1 share a common domain whose structure has been characterized as being particularly suitable for binding amylopectin double helices. The fact that ESV1, which consists mainly of this domain, doesn't interact with glycogen suggests that the interaction with the forming amylopectin is facilitated by the N-terminal domain of LESV. This gives LESV additional specificity, allowing it to bind to different glucans than ESV1.

C-terminal domain is designed to bind specifically (at least) two double helices of amylopectin

The models obtained for ESV1 and LESV *via* Alphafold provided a very reliable structure for the C-terminal tryptophan-rich domain, which is conserved between the two molecules. The other parts of the molecule were predicted without much reliability. The C-terminal domain folded into an original structure, forming a rather large oval (about 40 Å wide and 70Å long) antiparallel twisted β-sheet. On this β-sheet, the aromatic and acidic residues, organized in repeated sequences identified during the analysis of the protein sequences, form parallel lines equidistant from each other and parallel to the axis of the β-sheet. The side chains of these amino acids point alternately to both sides of the β-sheet, suggesting a bipolarity of the ß-sheet with respect to the attachment of amylopectin chains. However, one of the faces of the β-sheet is occupied by a conserved-C-terminal helix on the side of the β-sheet and a long α-helix or a long loop in LESV and ESV1 respectively. Based on our current data, even though we have observed protein conformational changes in the presence of amylopectin, we do not know if these parts can move to release the aromatic and acidic residue lines for amylopectin molecule binding. If that were the case, and if ESV1 and LESV could bind amylopectin on both sides of their β-sheet, we would expect to find assemblies with sandwich-like alignments of proteins and amylopectin. However, we have never observed such structures during the interaction experiments. Therefore, we computed models where only the "free" side interacted with double helices of amylopectin. We have been able to demonstrate that LESV and ESV1 can bind amylopectin in both its soluble form and its organized form within starch grains. Furthermore, we have shown that this interaction occurs through the common domain shared by LESV and ESV1, which is organized in a β-sheet structure using conserved aromatic and acidic residues. Moreover, we have shown that the unique and unprecedented structure of this domain enables it to bind two parallel double helices of amylopectin in the arrangement found in the crystalline phases of starch grains, thereby enabling its function in the organization and maintenance of starch grains in plants.

The N-terminal domain allows regulation of LESV specificity towards starch components.

While the structure of the N-terminal domain of LESV remains undefined, we've been able to determine that it consists primarily of disordered regions and α-helices located near the C-terminal domain. We've also found that interaction with amylopectin triggers the conversion of these disordered regions into α-helices. In ESV1, this domain is significantly reduced, but is still predicted to be unstructured, similar to the polyproline tail at the C-terminus. The presence of these regions likely thwarted our attempts to obtain crystals. Since amylopectin is not monodisperse in solution, it isn't suitable for stabilizing the proteins, which led us to search for analogues that could facilitate their crystallization.

While studying the structure of the C-terminal domains has provided insight into how both proteins interact with amylopectin, a complete structural analysis, especially for LESV, would allow us to define the role of the N-terminal domain. Although we've established that this domain likely contributes to the distinct specificity between the two proteins, a complete understanding remains elusive.

Given the presence of multiple helices in the N-terminal domain of LESV and its role in amylopectin biosynthesis, it's conceivable that LESV may also interact with other proteins. This concept is consistent with what's been documented for starch synthases (SSs) and branching enzymes (BEs) involved in starch biosynthesis (Ahmed, Tetlow et al. 2015, Crofts, Abe et al. 2015).

7. Conclusion and Perspectives

My research focused on studying the effects of the previously uncharacterized ESV1 and LESV proteins. The goal was to gain a better understanding of the structures and functions of these proteins in relation to starch biosynthesis. I found many interesting results.

First using a combination of structural and functional approaches, we could describe for the first time a new glucan binding domain devoted to the binding of amylopectin and further to starch granule.

Then we showed that beside this domain, conserved in both ESV1 et LESV, the N-terminal domain of LESV undergoes structural modification upon binding of amylopectin and particularly that some unfolded regions become structurated in α-helices. On the other hand, this N-terminal domain allows LESV protein to bind amylopectin during its biosynthesis, which is not the case for ESV1, which can only interact with amylopectin in its final form. So, ESV1 may intervene downstream the action of LESV to stabilize starch granules.

These results, combined with the functional results obtained by our collaborators allow us to describe the function of ESV1 and LESV, in which we can complete the mechanisms we proposed in our first publication (Figure 6). The function of LESV can start during the biosynthesis of amylopectin, probably concomitantly with the action of DBE, since their action has a cumulative effect in the formation of insoluble glucans. After the action of DBE, linked LESV can help the double helix organization in the formation of crystals nuclei. Then this organization nucleates and this organization spreads throughout the amylopectin molecules leading to the formation of crystalline lamellae probably helped by ESV1 molecules. Then ESV1 molecules bind to the newly formed crystalline lamellae, stabilizing them and preventing early degradation before night.

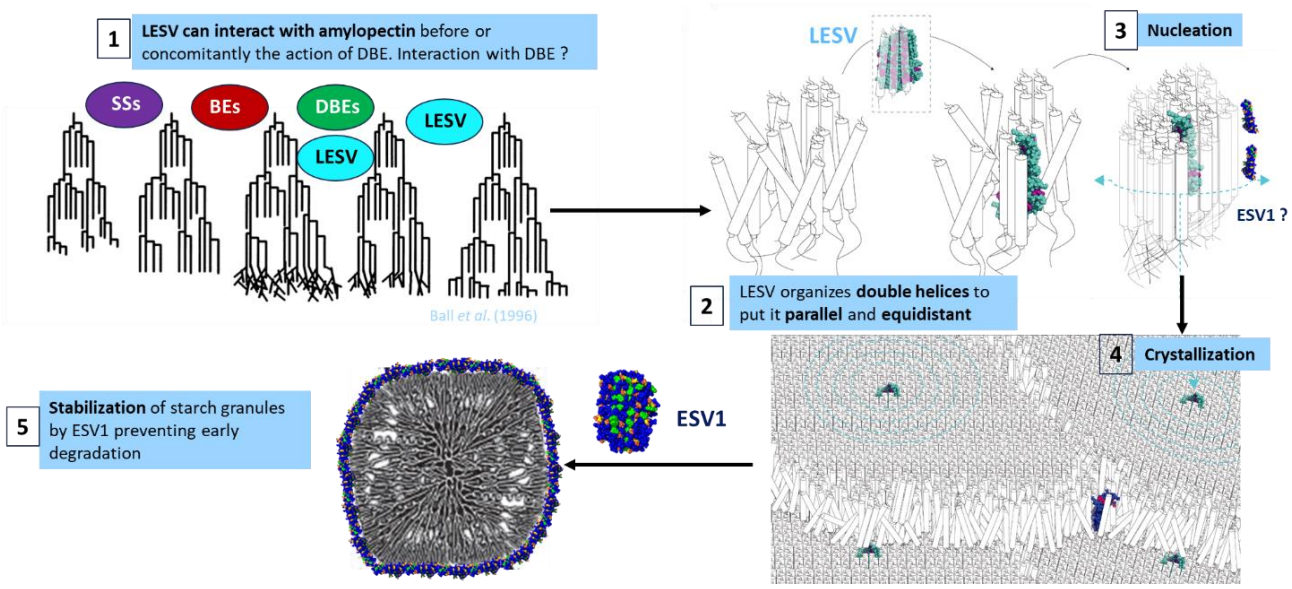


Figure 6: ***ESV1 and LESV proposed mechanism***

LESV may be a potential organizer of starch granule matrix, while ESV1 may act as a protective shield of amylopectin structure against degrading enzymes. These results are very important in the sense that they change the paradigm that until now described grain formation as an autonomous mechanism.

These results showed that the LESV and ESV1 are very important proteins in which their binding to amylopectin lead to the formation of additional helices and then could these double helices align others helices propagating a phase transition resulting in crystalline lamella which highlight the important function of both proteins LESV and ESV1 in the biosynthesis and in the maintenance of the starch granules.

Perspectives:

Further structural characterization of ESV1 and LESV by X-ray crystallography or Cryo Electron Microscopy would provide atomic-level insights into the protein-starch interactions and help to understand the binding mechanism, or by CryoEM.

The function of LESV and ESV1 has been studied in transient starch, it would be important to do the same analysis for reserve starch. As the metabolism of these two forms of starch are known to be different.

As ESV1 and LESV are conserved among the plants it would be interesting to check if their function is strictly conserved or slightly different in other species.

The exact mechanism of action is not yet described, further investigations are needed to understand the molecular process particularly putative interaction between ESV1 and LESV and other proteins and or enzymes involved in starch metabolism.

After having highlighted the function of ESV1 and LESV further exploration would be helpful. the potential application of ESV1 and LESV in biotechnological processes involving starch manipulation or engineering.

References

Abt, Melanie R., Barbara Pfister, Mayank Sharma, Simona Eicke, Léo Bürgy, Isabel Neale, David Seung, et Samuel C. Zeeman. 2020. « STARCH SYNTHASE5, a Noncanonical Starch Synthase-Like Protein, Promotes Starch Granule Initiation in Arabidopsis[OPEN] ». *The Plant Cell* 32 (8): 2543- 65. https://doi.org/10.1105/tpc.19.00946.

Adeva-Andany, María M., Manuel González-Lucán, Cristóbal Donapetry-García, Carlos Fernández-Fernández, et Eva Ameneiros-Rodríguez. 2016. « Glycogen metabolism in humans ». *BBA Clinical* 5 (février): 85- 100. https://doi.org/10.1016/j.bbacli.2016.02.001.

Alabadí, David, Tokitaka Oyama, Marcelo J. Yanovsky, Franklin G. Harmon, Paloma Más, et Steve A. Kay. 2001. « Reciprocal Regulation Between *TOC1* and *LHY* / *CCA1* Within the *Arabidopsis* Circadian Clock ». *Science* 293 (5531): 880- 83. https://doi.org/10.1126/science.1061320.

Alonso-Casajús, Nora, David Dauvillée, Alejandro Miguel Viale, Francisco José Muñoz, Edurne Baroja-Fernández, María Teresa Morán-Zorzano, Gustavo Eydallin, Steven Ball, et Javier Pozueta-Romero. 2006. « Glycogen Phosphorylase, the Product of the glgP Gene, Catalyzes Glycogen Breakdown by Removing Glucose Units from the Nonreducing Ends in Escherichia coli ». *Journal of Bacteriology* 188 (14): 5266- 72. https://doi.org/10.1128/JB.01566-05.

Ambigaipalan, P., R. Hoover, E. Donner, Q. Liu, S. Jaiswal, R. Chibbar, K. K. M. Nantanga, et K. Seetharaman. 2011. « Structure of Faba Bean, Black Bean and Pinto Bean Starches at Different Levels of Granule Organization and Their Physicochemical Properties ». *Food Research International* 44 (9): 2962- 74. https://doi.org/10.1016/j.foodres.2011.07.006.

Baker, Andrew A., Mervyn J. Miles, et William Helbert. 2001. « Internal Structure of the Starch Granule Revealed by AFM ». *Carbohydrate Research* 330 (2): 249- 56. https://doi.org/10.1016/S0008-6215(00)00275-5.

Ball, Steven G., et Matthew K. Morell. 2003. « From Bacterial Glycogen to Starch: Understanding the Biogenesis of the Plant Starch Granule ». *Annual Review of Plant Biology* 54 (1): 207- 33. https://doi.org/10.1146/annurev.arplant.54.031902.134927.

Ball, Steven G, Marion H.B.J van de Wal, et Richard G.F Visser. 1998. « Progress in Understanding the Biosynthesis of Amylose ». *Trends in Plant Science* 3 (12): 462- 67. https://doi.org/10.1016/S1360-1385(98)01342-9.

Ball, Steven, Han-Ping Guan, Martha James, Alan Myers, Peter Keeling, Gregory Mouille, Alain Buléon, Paul Colonna, et Jack Preiss. 1996. « From Glycogen to Amylopectin: A Model for the Biogenesis of the Plant Starch Granule ». *Cell* 86 (3): 349- 52. https://doi.org/10.1016/S0092-8674(00)80107-5.

Bayer, Roman G., Simon Stael, Edina Csaszar, et Markus Teige. 2011. « Mining the Soluble Chloroplast Proteome by Affinity Chromatography ». *PROTEOMICS* 11 (7): 1287- 99. https://doi.org/10.1002/pmic.201000495.

Bertoft, E. 1991. « Investigation of the Fine Structure of Alpha-Dextrins Derived from Amylopectin and Their Relation to the Structure of Waxy-Maize Starch ». *Carbohydrate Research* 212 (juin): 229- 44. https://doi.org/10.1016/0008-6215(91)84060-r.

Bertoft, Eric. 2013. « On the Building Block and Backbone Concepts of Amylopectin Structure ». *Cereal Chemistry* 90 (4): 294- 311. https://doi.org/10.1094/CCHEM-01-13-0004-FI.

———. 2017. « Understanding Starch Structure: Recent Progress ». *Agronomy* 7 (3): 56. https://doi.org/10.3390/agronomy7030056.

Bertoft, Eric, Kristine Koch, et Per Åman. 2012a. « Building Block Organisation of Clusters in Amylopectin from Different Structural Types ». *International Journal of Biological Macromolecules* 50 (5): 1212- 23. https://doi.org/10.1016/j.ijbiomac.2012.03.004.

———. 2012b. « Structure of Building Blocks in Amylopectins ». *Carbohydrate Research* 361 (novembre): 105- 13. https://doi.org/10.1016/j.carres.2012.08.012.

Bertoft, Eric, Kuakoon Piyachomkwan, Pathama Chatakanonda, et Klanarong Sriroth. 2008. « Internal Unit Chain Composition in Amylopectins ». *Carbohydrate Polymers* 74 (3): 527- 43. https://doi.org/10.1016/j.carbpol.2008.04.011.

Bertoft, Eric, Qin Zhu, Helena Andtfolk, et Marina Jungner. 1999. « Structural Heterogeneity in Waxy-Rice Starch ». *Carbohydrate Polymers* 38 (4): 349- 59. https://doi.org/10.1016/S0144-8617(98)00124-6.

Blauth, S. L., Y. Yao, J. D. Klucinec, J. C. Shannon, D. B. Thompson, et M. J. Guilitinan. 2001. « Identification of Mutator Insertional Mutants of Starch-Branching Enzyme 2a in Corn ». *Plant Physiology* 125 (3): 1396- 1405. https://doi.org/10.1104/pp.125.3.1396.

Blennow, Andreas, Anders K. Sjöland, Roger Andersson, et Per Kristiansson. 2005. « The Distribution of Elements in the Native Starch Granule as Studied by Particle-Induced X-Ray Emission and Complementary Methods ». *Analytical Biochemistry* 347 (2): 327- 29. https://doi.org/10.1016/j.ab.2005.06.020.

Bogracheva, T. Ya., V. J. Morris, S. G. Ring, et C. L. Hedley. 1998. « The Granular Structure of C-Type Pea Starch and Its Role in Gelatinization ». *Biopolymers* 45 (4): 323- 32. https://doi.org/10.1002/(SICI)1097-0282(19980405)45:4<323::AID-BIP6>3.0.CO;2-N.

Borovsky, Dov, Eric E. Smith, et William J. Whelan. 1976. « On The Mechanism of Amylose Branching by Potato Q-Enzyme ». *European Journal of Biochemistry* 62 (2): 307- 12. https://doi.org/10.1111/j.1432-1033.1976.tb10162.x.

Borovsky, Dov, Eric E. Smith, William J. Whelan, Dexter French, et Shoichi Kikumoto. 1979. « The Mechanism of Q-Enzyme Action and Its Influence on the Structure of Amylopectin ». *Archives*

of Biochemistry and Biophysics 198 (2): 627- 31. https://doi.org/10.1016/0003-9861(79)90540-X.

Boyer, Charles D., et Kang-Chien Liu. 1983. « Starch and Water-Soluble Polysaccharides from Sugary Endosperm of Sorghum ». *Phytochemistry* 22 (11): 2513- 15. https://doi.org/10.1016/0031-9422(83)80151-4.

Boyer, Charles D, et Jack Preiss. 1981. « Evidence for Independent Genetic Control of the Multiple Forms of Maize Endosperm Branching Enzymes and Starch Synthases1 2 » 67.

Bresolin, Nicole, Zhongyi Li, Behjat Kosar-Hashemi, Ian Tetlow, Manash Chatterjee, Sadequr Rahman, Matthew Morell, et Crispin Howitt. 2006. « Characterisation of disproportionating enzyme from wheat endosperm ». *Planta* 224 (juillet): 20- 31. https://doi.org/10.1007/s00425-005-0187-7.

Brummell, David A., Lyn M. Watson, Jun Zhou, Marian J. McKenzie, Ian C. Hallett, Lyall Simmons, Margaret Carpenter, et Gail M. Timmerman-Vaughan. 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 (avril): 28. https://doi.org/10.1186/s12896-015-0143-y.

Brust, Henrike, Tanja Lehmann, Christophe D'Hulst, et Joerg Fettke. 2014. « Analysis of the Functional Interaction of Arabidopsis Starch Synthase and Branching Enzyme Isoforms Reveals That the Cooperative Action of SSI and BEs Results in Glucans with Polymodal Chain Length Distribution Similar to Amylopectin ». *PLOS ONE* 9 (7): e102364. https://doi.org/10.1371/journal.pone.0102364.

Brust, Henrike, Slawomir Orzechowski, et Joerg Fettke. 2020. « Starch and Glycogen Analyses: Methods and Techniques ». *Biomolecules* 10 (7): 1020. https://doi.org/10.3390/biom10071020.

Buléon, A., P. Colonna, V. Planchot, et S. Ball. 1998. « Starch Granules: Structure and Biosynthesis ». *International Journal of Biological Macromolecules* 23 (2): 85- 112. https://doi.org/10.1016/S0141-8130(98)00040-3.

Buleon, Alain, Paul Colonna, Veronique Planchot, et S. Ball. 1998. « Starch Granules :Structure and Biosynthesis ». *International Journal of Biological Macromolecules* 23: 85.

Buleon, Alain, Bruno Pontoire, C. Riekel, H. Chanzy, William Helbert, et R. Vuong. 1997. « Crystalline Ultrastructure of Starch Granules Revealed by Synchrotron Radiation Microdiffraction Mapping ». *Macromolecules* 30 (13): 3952.

Burwinkel, B, H D Bakker, E Herschkovitz, S W Moses, Y S Shin, et M W Kilimann. 1998. « Mutations in the liver glycogen phosphorylase gene (PYGL) underlying glycogenosis type VI. » *American Journal of Human Genetics* 62 (4): 785-91.

Buschiazzo, Alejandro, Juan E Ugalde, Marcelo E Guerin, William Shepard, Rodolfo A Ugalde, et Pedro M Alzari. 2004. « Crystal structure of glycogen synthase: homologous enzymes catalyze glycogen synthesis and degradation ». *The EMBO Journal* 23 (16): 3196-3205. https://doi.org/10.1038/sj.emboj.7600324.

Busi, Maria V., Nicolas Palopoli, Hugo A. Valdez, Maria S. Fornasari, Nahuel Z. Wayllace, Diego F. Gomez-Casati, Gustavo Parisi, et Rodolfo A. Ugalde. 2008. « Functional and Structural Characterization of the Catalytic Domain of the Starch Synthase III from Arabidopsis Thaliana ». *Proteins: Structure, Function, and Bioinformatics* 70 (1): 31-40. https://doi.org/10.1002/prot.21469.

Cenci, Ugo, Malika Chabi, Mathieu Ducatez, Catherine Tirtiaux, Jennifer Nirmal-Raj, Yoshinori Utsumi, Daiki Kobayashi, et al. 2013. « Convergent Evolution of Polysaccharide Debranching Defines a Common Mechanism for Starch Accumulation in Cyanobacteria and Plants ». *The Plant Cell* 25 (10): 3961-75. https://doi.org/10.1105/tpc.113.118174.

Cenci, Ugo, Felix Nitschke, Martin Steup, Berge A. Minassian, Christophe Colleoni, et Steven G. Ball. 2014. « Transition from Glycogen to Starch Metabolism in Archaeplastida ». *Trends in Plant Science* 19 (1): 18-28. https://doi.org/10.1016/j.tplants.2013.08.004.

Chanzy, Henri, Jean-Luc Putaux, Danielle Dupeyre, Richard Davies, Manfred Burghammer, Suzelei Montanari, et Christian Riekel. 2006. « Morphological and Structural Aspects of the Giant Starch Granules from Phajus Grandifolius ». *Journal of Structural Biology* 154 (1): 100-110. https://doi.org/10.1016/j.jsb.2005.11.007.

Chen, Pei, Long Yu, Ling Chen, et Xiaoxi Li. 2006. « Morphology and Microstructure of Maize Starches with Different Amylose/Amylopectin Content ». *Starch - Stärke* 58 (12): 611-15. https://doi.org/10.1002/star.200500529.

Commuri, Padmavathi D., et Peter L. Keeling. 2001. « Chain-Length Specificities of Maize Starch Synthase I Enzyme: Studies of Glucan Affinity and Catalytic Properties ». *The Plant Journal* 25 (5): 475-86. https://doi.org/10.1046/j.1365-313x.2001.00955.x.

Coultate, 2001. s. d. *FOOD The Chemistry of Its Components, Fourth Edition*.

Critchley, J. H., S. C. Zeeman, T. Takaha, A. M. Smith, et S. M. Smith. 2001. « A Critical Role for Disproportionating Enzyme in Starch Breakdown Is Revealed by a Knock-out Mutation in Arabidopsis ». *The Plant Journal: For Cell and Molecular Biology* 26 (1): 89-100. https://doi.org/10.1046/j.1365-313x.2001.01012.x.

Dauvillée, David, Christophe Colleoni, Gregory Mouille, Matthew K. Morell, Christophe d'Hulst, Fabrice Wattebled, Luc Liénard, et al. 2001. « Biochemical Characterization of Wild-Type and Mutant Isoamylases of Chlamydomonas reinhardtii Supports a Function of the Multimeric Enzyme Organization in Amylopectin Maturation1 ». *Plant Physiology* 125 (4): 1723-31. https://doi.org/10.1104/pp.125.4.1723.

David, G., et J. Pérez. 2009. « Combined Sampler Robot and High-Performance Liquid Chromatography: A Fully Automated System for Biological Small-Angle X-Ray Scattering Experiments at the Synchrotron SOLEIL SWING Beamline ». *Journal of Applied Crystallography* 42 (5): 892-900. https://doi.org/10.1107/S0021889809029288.

Delatte, Thierry, Martine Trevisan, Mary L. Parker, et Samuel C. Zeeman. 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 (6): 815-30. https://doi.org/10.1111/j.1365-313X.2005.02348.x.

Denver, K., L. M. Barber, R. Burton, C. L. Hedley, C. M. Hylton, S. Johnson, D. A. Jones, et al. 1995. « The Isolation and Characterization of Novel Low-Amylose Mutants of Pisum Sativum L. » *Plant, Cell and Environment* 18 (9): 1019-26. https://doi.org/10.1111/j.1365-3040.1995.tb00612.x.

Denyer, K, D Waite, S Motawia, B L Møller, et A M Smith. 1999. « Granule-bound starch synthase I in isolated starch granules elongates malto-oligosaccharides processively. » *Biochemical Journal* 340 (Pt 1): 183-91.

Denyer, Kay, Christopher Sidebottom, Christopher M. Hylton, et Alison M. Smith. 1993. « Soluble Isoforms of Starch Synthase and Starch-Branching Enzyme Also Occur within Starch Granules in Developing Pea Embryos ». *The Plant Journal* 4 (1): 191-98. https://doi.org/10.1046/j.1365-313X.1993.04010191.x.

Dhital, Sushil, Ashok K. Shrestha, Jovin Hasjim, et Michael J. Gidley. 2011. « Physicochemical and Structural Properties of Maize and Potato Starches as a Function of Granule Size ». *Journal of Agricultural and Food Chemistry* 59 (18): 10151-61. https://doi.org/10.1021/jf202293s.

Donald, A. M. 2001. « Plasticization and Self Assembly in the Starch Granule ». *Cereal Chemistry Journal* 78 (3): 307-14. https://doi.org/10.1094/CCHEM.2001.78.3.307.

Doutch, James, et Elliot P. Gilbert. 2013. « Characterisation of Large Scale Structures in Starch Granules via Small-Angle Neutron and X-Ray Scattering ». *Carbohydrate Polymers* 91 (1): 444-51. https://doi.org/10.1016/j.carbpol.2012.08.002.

Drummond, George S., Eric E. Smith, et William J. Whelan. 1972. « Purification and Properties of Potato ». *European Journal of Biochemistry* 26 (2): 168-76. https://doi.org/10.1111/j.1432-1033.1972.tb01753.x.

Dumez, Sylvain, Fabrice Wattebled, David Dauvillee, David Delvalle, Véronique Planchot, Steven G. Ball, et Christophe D'Hulst. 2006. « Mutants of Arabidopsis Lacking Starch Branching Enzyme II Substitute Plastidial Starch Synthesis by Cytoplasmic Maltose Accumulation ». *The Plant Cell* 18 (10): 2694-2709. https://doi.org/10.1105/tpc.105.037671.

Dvonch, William., et Roy L. Whistler. 1949. « WATER-SOLUBLE POLYSACCHARIDES OF SWEET CORN ». *Journal of Biological Chemistry* 181 (2): 889-95. https://doi.org/10.1016/S0021-9258(18)56612-6.

Edelstein, Arthur, Nenad Amodaj, Karl Hoover, Ron Vale, et Nico Stuurman. 2010. « Computer Control of Microscopes Using MManager ». *Current Protocols in Molecular Biology* Chapter 14 (octobre): Unit14.20. https://doi.org/10.1002/0471142727.mb1420s92.

Edner, Christoph, Jing Li, Tanja Albrecht, Sebastian Mahlow, Mahdi Hejazi, Hasnain Hussain, Fatma Kaplan, et al. 2007. « Glucan, Water Dikinase Activity Stimulates Breakdown of Starch Granules by Plastidial Beta-Amylases ». *Plant Physiology* 145 (1): 17-28. https://doi.org/10.1104/pp.107.104224.

Edwards, Anne, Alip Borthakur, Stephen Bornemann, Julien Venail, Kay Denyer, Darren Waite, Dan Fulton, Alison Smith, et Cathie Martin. 1999. « Specificity of Starch Synthase Isoforms from Potato ». *European Journal of Biochemistry* 266 (3): 724-36. https://doi.org/10.1046/j.1432-1327.1999.00861.x.

Emanuelsson, Olof, Henrik Nielsen, et Gunnar Von Heijne. 1999. « ChloroP, a Neural Network-Based Method for Predicting Chloroplast Transit Peptides and Their Cleavage Sites ». *Protein Science* 8 (5): 978-84. https://doi.org/10.1110/ps.8.5.978.

Fasahat, Parviz, Sadequr Rahman, et Wickneswari Ratnam. 2014. « Genetic Controls on Starch Amylose Content in Wheat and Rice Grains ». *Journal of Genetics* 93 (1): 279-92. https://doi.org/10.1007/s12041-014-0325-8.

Feike, Doreen, David Seung, Alexander Graf, Sylvain Bischof, Tamaryn Ellick, Mario Coiro, Sebastian Soyk, et al. 2016. « The Starch Granule-Associated Protein EARLY STARVATION1 Is Required for the Control of Starch Degradation in Arabidopsis thaliana Leaves[OPEN] ». *The Plant Cell* 28 (6): 1472-89. https://doi.org/10.1105/tpc.16.00011.

Fisher, Dane K, Ming Gao, Kyung-Nam Kim, Charles D Boyer, et Mark J Guiltinan. 1996. « Allelic Analysis of the Maize Amylose-Exfender Locus Suggests That Lndependent Cenes Encode Starch-Branching Enzymes Ila and Llb' ».

Franke, D., M. V. Petoukhov, P. V. Konarev, A. Panjkovich, A. Tuukkanen, H. D. T. Mertens, A. G. Kikhney, et al. 2017. « ATSAS 2.8: A Comprehensive Data Analysis Suite for Small-Angle Scattering from Macromolecular Solutions ». *Journal of Applied Crystallography* 50 (4): 1212-25. https://doi.org/10.1107/S1600576717007786.

Franke, Daniel, et Dmitri I. Svergun. 2009. « DAMMIF, a program for rapid ab-initio shape determination in small-angle scattering ». *Journal of Applied Crystallography* 42 (Pt 2): 342-46. https://doi.org/10.1107/S0021889809000338.

French, Dexter. 1972. « Fine Structure of Starch and its Relationship to the Organization of Starch Granules ». *Journal of the Japanese Society of Starch Science* 19 (1): 8-25. https://doi.org/10.5458/jag1972.19.8.

Fujita, Naoko, Mayumi Yoshida, Noriko Asakura, Takashi Ohdan, Akio Miyao, Hirohiko Hirochika, et Yasunori Nakamura. 2006. « Function and Characterization of Starch Synthase I Using Mutants in Rice ». *Plant Physiology* 140 (3): 1070-84. https://doi.org/10.1104/pp.105.071845.

Fulton, Daniel C., Michaela Stettler, Tabea Mettler, Cara K. Vaughan, Jing Li, Perigio Francisco, Manuel Gil, et al. 2008. « Beta-AMYLASE4, a Noncatalytic Protein Required for Starch Breakdown, Acts Upstream of Three Active Beta-Amylases in Arabidopsis Chloroplasts ». *The Plant Cell* 20 (4): 1040-58. https://doi.org/10.1105/tpc.107.056507.

Furukawa, K, M Tagaya, M Inouye, J Preiss, et T Fukui. 1990. « Identification of Lysine 15 at the Active Site in Escherichia Coli Glycogen Synthase. Conservation of Lys-X-Gly-Gly Sequence in the Bacterial and Mammalian Enzymes. » *Journal of Biological Chemistry* 265 (4): 2086-90. https://doi.org/10.1016/S0021-9258(19)39943-0.

Gallant, Daniel J., Brigitte Bouchet, et Paul M. Baldwin. 1997. « Microscopy of Starch: Evidence of a New Level of Granule Organization ». *Carbohydrate Polymers* 32 (3): 177-91. https://doi.org/10.1016/S0144-8617(97)00008-8.

Geddes, R., C. T. Greenwood, et S. Mackenzie. 1965. « Studies on the Biosynthesis of Starch Granules: Part III. The Properties of the Components of Starches from the Growing Potato Tuber ». *Carbohydrate Research* 1 (1): 71-82. https://doi.org/10.1016/S0008-6215(00)80215-3.

Geigenberger, Peter. 2011. « Regulation of Starch Biosynthesis in Response to a Fluctuating Environment1 ». *Plant Physiology* 155 (4): 1566-77. https://doi.org/10.1104/pp.110.170399.

Gérard, Catherine, Véronique Planchot, Paul Colonna, et Eric Bertoft. 2000. « Relationship between Branching Density and Crystalline Structure of A- and B-Type Maize Mutant Starches ». *Carbohydrate Research* 326 (2): 130-44. https://doi.org/10.1016/S0008-6215(00)00025-2.

Gernat, Ch., S. Radosta, G. Damaschun, et F. Schierbaum. 1990. « Supramolecular Structure of Legume Starches Revealed by X-Ray Scattering ». *Starch - Stärke* 42 (5): 175-78. https://doi.org/10.1002/star.19900420504.

Gibbons, Brian J., Peter J. Roach, et Thomas D. Hurley. 2002. « Crystal Structure of the Autocatalytic Initiator of Glycogen Biosynthesis, Glycogenin ». *Journal of Molecular Biology* 319 (2): 463-77. https://doi.org/10.1016/S0022-2836(02)00305-4.

Gibon, Yves, Eva-Theresa Pyl, Ronan Sulpice, John E. Lunn, Melanie Höhne, Manuela Günther, et Mark Stitt. 2009. « Adjustment of Growth, Starch Turnover, Protein Content and Central Metabolism to a Decrease of the Carbon Supply When Arabidopsis Is Grown in Very Short

Photoperiods ». *Plant, Cell & Environment* 32 (7): 859-74. https://doi.org/10.1111/j.1365-3040.2009.01965.x.

Gibon, Yves, Bjoern Usadel, Oliver E. Blaesing, Beate Kamlage, Melanie Hoehne, Richard Trethewey, et Mark Stitt. 2006. « Integration of metabolite with transcript and enzyme activity profiling during diurnal cycles in Arabidopsisrosettes ». *Genome Biology* 7 (8): R76. https://doi.org/10.1186/gb-2006-7-8-r76.

Gidley, Michael J., et Paul V. Bulpin. 1987. « Crystallisation of Malto-Oligosaccharides as Models of the Crystalline Forms of Starch: Minimum Chain-Length Requirement for the Formation of Double Helices ». *Carbohydrate Research* 161 (2): 291-300. https://doi.org/10.1016/S0008-6215(00)90086-7.

Goldstein, Avi, George Annor, Jean-Luc Putaux, Kim H. Hebelstrup, Andreas Blennow, et Eric Bertoft. 2016. « Impact of Full Range of Amylose Contents on the Architecture of Starch Granules ». *International Journal of Biological Macromolecules* 89 (août): 305-18. https://doi.org/10.1016/j.ijbiomac.2016.04.053.

Graf, Alexander, Armin Schlereth, Mark Stitt, et Alison M. Smith. 2010. « Circadian control of carbohydrate availability for growth in Arabidopsis plants at night ». *Proceedings of the National Academy of Sciences of the United States of America* 107 (20): 9458-63. https://doi.org/10.1073/pnas.0914299107.

Greenfield, Norma J. 2006. « Using Circular Dichroism Spectra to Estimate Protein Secondary Structure ». *Nature Protocols* 1 (6): 2876-90. https://doi.org/10.1038/nprot.2006.202.

Gregory R. Ziegler, *, and John A. Creek, et James Runt. 2005. « Spherulitic Crystallization in Starch as a Model for Starch Granule Initiation ». Research-article. ACS Publications. American Chemical Society. World. 25 mars 2005. https://doi.org/10.1021/bm049214p.

Grimaud, Florent, Hélène Rogniaux, Martha G. James, Alan M. Myers, et Véronique Planchot. 2008. « Proteome and phosphoproteome analysis of starch granule-associated proteins from normal maize and mutants affected in starch biosynthesis ». *Journal of Experimental Botany* 59 (12): 3395-3406. https://doi.org/10.1093/jxb/ern198.

Guan, Han Ping, et Jack Preiss. 1993. « Differentiation of the Properties of the Branching Lsozymes from Maize (Zea Mays)' ».

Hanashiro, I., K. Itoh, Y. Kuratomi, M. Yamazaki, T. Igarashi, J.-i. Matsugasako, et Y. Takeda. 2008. « Granule-Bound Starch Synthase I Is Responsible for Biosynthesis of Extra-Long Unit Chains of Amylopectin in Rice ». *Plant and Cell Physiology* 49 (6): 925-33. https://doi.org/10.1093/pcp/pcn066.

Hanashiro, Isao, Jun-ichi Matsugasako, Tamami Egashira, et Yasuhito Takeda. 2005. « Structural Characterization of Long Unit-Chains of Amylopectin ». *Journal of Applied Glycoscience* 52 (3): 233-37. https://doi.org/10.5458/jag.52.233.

Hanson, Kenneth R., et Neil A. McHale. 1988. « A Starchless Mutant of Nicotiana sylvestris Containing a Modified Plastid Phosphoglucomutase ». *Plant Physiology* 88 (3): 838-44.

Haworth, W. N., E. L. Hirst, et F. A. Isherwood. 1937. « 129. Polysaccharides. Part XXIII. Determination of the Chain Length of Glycogen ». *Journal of the Chemical Society (Resumed)*, n° 0 (janvier): 577-81. https://doi.org/10.1039/JR9370000577.

Hennen-Bierwagen, Tracie A., Fushan Liu, Rebekah S. Marsh, Seungtaek Kim, Qinglei Gan, Ian J. Tetlow, Michael J. Emes, Martha G. James, et Alan M. Myers. 2008. « Starch Biosynthetic Enzymes from Developing Maize Endosperm Associate in Multisubunit Complexes ». *Plant Physiology* 146 (4): 1892-1908. https://doi.org/10.1104/pp.108.116285.

Hizukuri, Susumu. 1986. « Polymodal Distribution of the Chain Lengths of Amylopectins, and Its Significance ». *Carbohydrate Research* 147 (2): 342-47. https://doi.org/10.1016/S0008-6215(00)90643-8.

Hizukuri, Susumu, Shiro Tabata, Kagoshima, et Ziro Nikuni. 1970. « Studies on Starch Phosphate Part 1. Estimation of glucose-6-phosphate residues in starch and the presence of other bound phosphate(s) ». *Starch - Stärke* 22 (10): 338-43. https://doi.org/10.1002/star.19700221004.

Hizukuri, Susumu, Yasuhito Takeda, Michiko Yasuda, et Ayako Suzuki. 1981. « Multi-Branched Nature of Amylose and the Action of Debranching Enzymes ». *Carbohydrate Research* 94 (2): 205-13. https://doi.org/10.1016/S0008-6215(00)80718-1.

Hobson, P. N., W. J. Whelan, et Stanley Peat. 1951. « 333. The Enzymic Synthesis and Degradation of Starch. Part XIV. R-Enzyme ». *Journal of the Chemical Society (Resumed)*, n° 0 (janvier): 1451-59. https://doi.org/10.1039/JR9510001451.

Hurley, Thomas D., Chad Walls, John R. Bennett, Peter J. Roach, et Mu Wang. 2006. « Direct detection of glycogenin reaction products during glycogen initiation ». *Biochemical and biophysical research communications* 348 (2): 374-78. https://doi.org/10.1016/j.bbrc.2006.07.106.

Hussain, Rohanah, Tamas Javorfi, et Giuliano Siligardi. 2012. « 8.23 Spectroscopic Analysis: Synchrotron Radiation Circular Dichroism ». *Comprehensive Chirality* 8 (septembre): 438-48. https://doi.org/10.1016/B978-0-08-095167-6.00841-7.

Hussain, Rohanah, Edoardo Longo, et Giuliano Siligardi. 2018. « UV-Denaturation Assay to Assess Protein Photostability and Ligand-Binding Interactions Using the High Photon Flux of Diamond B23 Beamline for SRCD ». *Molecules (Basel, Switzerland)* 23 (8): 1906. https://doi.org/10.3390/molecules23081906.

Hwang, Seon-Kap, Kaan Koper, Hikaru Satoh, et Thomas W. Okita. 2016. « Rice Endosperm Starch Phosphorylase (Pho1) Assembles with Disproportionating Enzyme (Dpe1) to Form a Protein Complex That Enhances Synthesis of Malto-oligosaccharides ». *The Journal of Biological Chemistry* 291 (38): 19994-7. https://doi.org/10.1074/jbc.M116.735449.

Imberty, A., H. Chanzy, S. Pérez, A. Buléon, et V. Tran. 1988. « The Double-Helical Nature of the Crystalline Part of A-Starch ». *Journal of Molecular Biology* 201 (2): 365-78. https://doi.org/10.1016/0022-2836(88)90144-1.

Imberty, Anne, Alain Buléon, Vinh Tran, et Serge Péerez. 1991. « Recent Advances in Knowledge of Starch Structure ». *Starch - Stärke* 43 (10): 375-84. https://doi.org/10.1002/star.19910431002.

Inouchi, N., D. V. Glover, et H. Fuwa. 1987. « Chain Length Distribution of Amylopectins of Several Single Mutants and the Normal Counterpart, and Sugary-1 Phytoglycogen in Maize (Zea mays L.) ». *Starch - Stärke* 39 (8): 259-66. https://doi.org/10.1002/star.19870390802.

James, M. G., D. S. Robertson, et A. M. Myers. 1995. « Characterization of the Maize Gene Sugary1, a Determinant of Starch Composition in Kernels ». *The Plant Cell* 7 (4): 417-29. https://doi.org/10.1105/tpc.7.4.417.

James, Martha G, Kay Denyer, et Alan M Myers. 2003. « Starch Synthesis in the Cereal Endosperm ». *Current Opinion in Plant Biology* 6 (3): 215-22. https://doi.org/10.1016/S1369-5266(03)00042-6.

Jamme, F., D. Bourquin, G. Tawil, A. Viksø-Nielsen, A. Buléon, et M. Réfrégiers. 2014. « 3D Imaging of Enzymes Working in Situ ». *Analytical Chemistry* 86 (11): 5265-70. https://doi.org/10.1021/ac403699h.

Jane, Jay-Lin, Tunyawat Kasemsuwan, Sharon Leas, Henzy Zobel, et John F. Robyt. 1994. « Anthology of Starch Granule Morphology by Scanning Electron Microscopy ». *Starch - Stärke* 46 (4): 121-29. https://doi.org/10.1002/star.19940460402.

Jane, Jay-lin, et James J. Shen. 1993. « Internal Structure of the Potato Starch Granule Revealed by Chemical Gelatinization ». *Carbohydrate Research* 247 (septembre): 279-90. https://doi.org/10.1016/0008-6215(93)84260-D.

Janeček, Štefan, Birte Svensson, et E. Ann MacGregor. 2011. « Structural and Evolutionary Aspects of Two Families of Non-Catalytic Domains Present in Starch and Glycogen Binding Proteins from Microbes, Plants and Animals ». *Enzyme and Microbial Technology* 49 (5): 429-40. https://doi.org/10.1016/j.enzmictec.2011.07.002.

Jenkins, P. J., et A. M. Donald. 1995. « The Influence of Amylose on Starch Granule Structure ». *International Journal of Biological Macromolecules* 17 (6): 315-21. https://doi.org/10.1016/0141-8130(96)81838-1.

Jenkins, Paul J., Ruth E. Cameron, et Athene M. Donald. 1993. « A Universal Feature in the Structure of Starch Granules from Different Botanical Sources ». *Starch - Stärke* 45 (12): 417-20. https://doi.org/10.1002/star.19930451202.

Jobling, Stephen A.., Gerhard P.. Schwall, Roger J.. Westcott, Christopher M.. Sidebottom, Martine Debet, Michael J.. Gidley, Roger Jeffcoat, et Richard Safford. 1999. « A Minor Form of Starch Branching Enzyme in Potato (Solanum Tuberosum L.) Tubers Has a Major Effect on Starch Structure: Cloning and Characterisation of Multiple Forms of SBE A ». *The Plant Journal* 18 (2): 163-71. https://doi.org/10.1046/j.1365-313X.1999.00441.x.

Jones, Gareth, Peter Willett, Robert C Glen, Andrew R Leach, et Robin Taylor. 1997. « Development and Validation of a Genetic Algorithm for Flexible Docking11Edited by F. E. Cohen ». *Journal of Molecular Biology* 267 (3): 727-48. https://doi.org/10.1006/jmbi.1996.0897.

Jumper, John, Richard Evans, Alexander Pritzel, Tim Green, Michael Figurnov, Olaf Ronneberger, Kathryn Tunyasuvunakool, et al. 2021. « Highly Accurate Protein Structure Prediction with AlphaFold ». *Nature* 596 (7873): 583-89. https://doi.org/10.1038/s41586-021-03819-2.

Källman, Anna, Eric Bertoft, Kristine Koch, Chuanxin Sun, Per Åman, et Roger Andersson. 2015. « Starch Structure in Developing Barley Endosperm ». *International Journal of Biological Macromolecules* 81 (novembre): 730-35. https://doi.org/10.1016/j.ijbiomac.2015.09.013.

Kelly, Sharon M., Thomas J. Jess, et Nicholas C. Price. 2005. « How to Study Proteins by Circular Dichroism ». *Biochimica et Biophysica Acta (BBA) - Proteins and Proteomics* 1751 (2): 119-39. https://doi.org/10.1016/j.bbapap.2005.06.005.

Keurentjes, Joost J. B., Ronan Sulpice, Yves Gibon, Marie-Caroline Steinhauser, Jingyuan Fu, Maarten Koornneef, Mark Stitt, et Dick Vreugdenhil. 2008. « Integrative Analyses of Genetic Variation in Enzyme Activities of Primary Carbohydrate Metabolism Reveal Distinct Modes of Regulation in Arabidopsis Thaliana ». *Genome Biology* 9 (8): R129. https://doi.org/10.1186/gb-2008-9-8-r129.

Khalid, Saud, Long Yu, Linghan Meng, Hongsheng Liu, Amjad Ali, et Ling Chen. 2017. « Poly(Lactic Acid)/Starch Composites: Effect of Microstructure and Morphology of Starch Granules on Performance ». *Journal of Applied Polymer Science* 134 (46): 45504. https://doi.org/10.1002/app.45504.

Klucinec, Jeffrey D., et Donald B. Thompson. 2002. « Structure of Amylopectins from Ae-Containing Maize Starches ». *Cereal Chemistry* 79 (1): 19-23. https://doi.org/10.1094/CCHEM.2002.79.1.19.

Koornneef, Maarten, et David Meinke. 2010. « The Development of Arabidopsis as a Model Plant ». *The Plant Journal* 61 (6): 909-21. https://doi.org/10.1111/j.1365-313X.2009.04086.x.

Kossmann, Jens, Gernot Abel, Franziska Springer, James Lloyd, et Lothar Willmitzer. 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 (juin): 503-11. https://doi.org/10.1007/s004250050587.

Kötting, Oliver, Diana Santelia, Christoph Edner, Simona Eicke, Tina Marthaler, Matthew S. Gentry, Sylviane Comparot-Moss, et al. 2009. « STARCH-EXCESS4 Is a Laforin-like Phosphoglucan Phosphatase Required for Starch Degradation in Arabidopsis Thaliana ». *The Plant Cell* 21 (1): 334-46. https://doi.org/10.1105/tpc.108.064360.

Kozlov, S. S., A. V. Krivandin, Olga V. Shatalova, T. Noda, E. Bertoft, J. Fornal, et V. P. Yuryev. 2007. « Structureof Starches Extracted from Near-Isogenic Wheat Lines ». *Journal of Thermal Analysis and Calorimetry* 87 (2): 575-84. https://doi.org/10.1007/s10973-006-7880-z.

Kunz, H H, R E Häusler, J Fettke, K Herbst, P Niewiadomski, M Gierth, K Bell, M Steup, U-I Flügge, et A Schneider. 2010. « The Role of Plastidial Glucose-6-Phosphate/Phosphate Translocators in Vegetative Tissues of Arabidopsis Thaliana Mutants Impaired in Starch Biosynthesis ». *Plant Biology (Stuttgart, Germany)* 12 Suppl 1 (septembre): 115-28. https://doi.org/10.1111/j.1438-8677.2010.00349.x.

Lafiandra, Domenico, Gabriele Riccardi, et Peter R. Shewry. 2014. « Improving Cereal Grain Carbohydrates for Diet and Health ». *Journal of Cereal Science*, Cereal Science for Food Security,Nutrition and Sustainability, 59 (3): 312-26. https://doi.org/10.1016/j.jcs.2014.01.001.

Laohaphatanaleart, Kamlai, Kuakoon Piyachomkwan, Klanarong Sriroth, Vilai Santisopasri, et Eric Bertoft. 2009. « A Study of the Internal Structure in Cassava and Rice Amylopectin ». *Starch - Stärke* 61 (10): 557-69. https://doi.org/10.1002/star.200900154.

Larsson, Clas-Tomas, Per Hofvander, Jamshid Khoshnoodi, Bo Ek, Lars Rask, et Håkan Larsson. 1996. « Three Isoforms of Starch Synthase and Two Isoforms of Branching Enzyme Are Present in Potato Tuber Starch ». *Plant Science* 117 (1): 9-16. https://doi.org/10.1016/0168-9452(96)04408-1.

Leterrier, Marina, Lynn D Holappa, Karen E Broglie, et Diane M Beckles. 2008. « Cloning, Characterisation and Comparative Analysis of a Starch Synthase IV Gene in Wheat: Functional and Evolutionary Implications ». *BMC Plant Biology* 8 (1): 98. https://doi.org/10.1186/1471-2229-8-98.

Lin, B., H. Hiraiwa, C. J. Pan, R. C. Nordlie, et J. Y. Chou. 1999. « Type-1c Glycogen Storage Disease Is Not Caused by Mutations in the Glucose-6-Phosphate Transporter Gene ». *Human Genetics* 105 (5): 515-17. https://doi.org/10.1007/s004390051140.

Lin, Qiaohui, Binquan Huang, Mingxu Zhang, Xiaoli Zhang, Joshua Rivenbark, Ryan L. Lappe, Martha G. James, Alan M. Myers, et Tracie A. Hennen-Bierwagen. 2012. « Functional Interactions

between Starch Synthase III and Isoamylase-Type Starch-Debranching Enzyme in Maize Endosperm ». *Plant Physiology* 158 (2): 679-92. https://doi.org/10.1104/pp.111.189704.

Liu, Chun, Barbara Pfister, Rayan Osman, Maximilian Ritter, Arvid Heutinck, Mayank Sharma, Simona Eicke, et al. 2023. « LIKE EARLY STARVATION 1 and EARLY STARVATION 1 Promote and Stabilize Amylopectin Phase Transition in Starch Biosynthesis », 2023.

Liu, Fushan, Amina Makhmoudova, Elizabeth A. Lee, Robin Wait, Michael J. Emes, et Ian J. Tetlow. 2009. « The Amylose Extender Mutant of Maize Conditions Novel Protein-Protein Interactions between Starch Biosynthetic Enzymes in Amyloplasts ». *Journal of Experimental Botany* 60 (15): 4423-40. https://doi.org/10.1093/jxb/erp297.

Liu, Fushan, Nadya Romanova, Elizabeth A Lee, Regina Ahmed, Martin Evans, Elliot P Gilbert, Matthew K Morell, Michael J Emes, et Ian J Tetlow. 2012. « Glucan Affinity of Starch Synthase IIa Determines Binding of Starch Synthase I and Starch-Branching Enzyme IIb to Starch Granules ». *The Biochemical Journal* 448 (3): 373-87. https://doi.org/10.1042/bj20120573.

Lloyd, James R., Jens Kossmann, et Gerhard Ritte. 2005. « Leaf Starch Degradation Comes out of the Shadows ». *Trends in Plant Science* 10 (3): 130-37. https://doi.org/10.1016/j.tplants.2005.01.001.

Lohmeier-Vogel, Elke M, David Kerk, Mhairi Nimick, Susan Wrobel, Lori Vickerman, Douglas G Muench, et Greg BG Moorhead. 2008. « Arabidopsis At5g39790 encodes a chloroplast-localized, carbohydrate-binding, coiled-coil domain-containing putative scaffold protein ». *BMC Plant Biology* 8 (novembre): 120. https://doi.org/10.1186/1471-2229-8-120.

Lu, Yan, Jon M. Steichen, Sean E. Weise, et Thomas D. Sharkey. 2006. « Cellular and Organ Level Localization of Maltose in Maltose-Excess Arabidopsis Mutants ». *Planta* 224 (4): 935-43. https://doi.org/10.1007/s00425-006-0263-7.

Lundquist, Peter K., Otho Mantegazza, Anja Stefanski, Kai Stühler, et Andreas P. M. Weber. 2017. « Surveying the Oligomeric State of Arabidopsis Thaliana Chloroplasts ». *Molecular Plant* 10 (1): 197-211. https://doi.org/10.1016/j.molp.2016.10.011.

Mahlow, Sebastian, Sławomir Orzechowski, et Joerg Fettke. 2016. « Starch Phosphorylation: Insights and Perspectives ». *Cellular and Molecular Life Sciences: CMLS* 73 (14): 2753-64. https://doi.org/10.1007/s00018-016-2248-4.

Malinova, Irina, Saleh Alseekh, Regina Feil, Alisdair R. Fernie, Otto Baumann, Mark Aurel Schöttler, John E. Lunn, et Joerg Fettke. 2017. « Starch Synthase 4 and Plastidal Phosphorylase Differentially Affect Starch Granule Number and Morphology ». *Plant Physiology* 174 (1): 73-85. https://doi.org/10.1104/pp.16.01859.

Malinova, Irina, Sebastian Mahlow, Saleh Alseekh, Tom Orawetz, Alisdair R. Fernie, Otto Baumann, Martin Steup, et Joerg Fettke. 2014. « Double Knockout Mutants of Arabidopsis Grown under

Normal Conditions Reveal That the Plastidial Phosphorylase Isozyme Participates in Transitory Starch Metabolism ». *Plant Physiology* 164 (2): 907-21. https://doi.org/10.1104/pp.113.227843.

Malinova, Irina, Harendra Mahto, Felix Brandt, Shadha AL-Rawi, Hadeel Qasim, Henrike Brust, Mahdi Hejazi, et Joerg Fettke. 2018. « EARLY STARVATION1 Specifically Affects the Phosphorylation Action of Starch-Related Dikinases ». *The Plant Journal* 95 (1): 126-37. https://doi.org/10.1111/tpj.13937.

Malinova, Irina, Hadeel M. Qasim, Henrike Brust, et Joerg Fettke. 2018. « Parameters of Starch Granule Genesis in Chloroplasts of Arabidopsis Thaliana ». *Frontiers in Plant Science* 9: 761. https://doi.org/10.3389/fpls.2018.00761.

Manners, David J. 1989. « Recent Developments in Our Understanding of Amylopectin Structure ». *Carbohydrate Polymers* 11 (2): 87-112. https://doi.org/10.1016/0144-8617(89)90018-0.

———. 1991. « Recent Developments in Our Understanding of Glycogen Structure ». *Carbohydrate Polymers* 16 (1): 37-82. https://doi.org/10.1016/0144-8617(91)90071-J.

Meinke, D. W., J. M. Cherry, C. Dean, S. D. Rounsley, et M. Koornneef. 1998. « Arabidopsis Thaliana: A Model Plant for Genome Analysis ». *Science (New York, N.Y.)* 282 (5389): 662, 679-82. https://doi.org/10.1126/science.282.5389.662.

Micsonai, András, Éva Moussong, Frank Wien, Eszter Boros, Henrietta Vadászi, Nikoletta Murvai, Young-Ho Lee, et al. 2022. « BeStSel: Webserver for Secondary Structure and Fold Prediction for Protein CD Spectroscopy ». *Nucleic Acids Research* 50 (W1): W90-98. https://doi.org/10.1093/nar/gkac345.

Micsonai, András, Frank Wien, Linda Kernya, Young-Ho Lee, Yuji Goto, Matthieu Réfrégiers, et József Kardos. 2015. « Accurate secondary structure prediction and fold recognition for circular dichroism spectroscopy ». *Proceedings of the National Academy of Sciences* 112 (24): E3095-3103. https://doi.org/10.1073/pnas.1500851112.

Miles, A. J., Robert W. Janes, A. Brown, D. T. Clarke, J. C. Sutherland, Y. Tao, B. A. Wallace, et S. V. Hoffmann. 2008. « Light Flux Density Threshold at Which Protein Denaturation Is Induced by Synchrotron Radiation Circular Dichroism Beamlines ». *Journal of Synchrotron Radiation* 15 (Pt 4): 420-22. https://doi.org/10.1107/S0909049508009606.

Miles, Andrew J., et B. A. Wallace. 2018. « CDtoolX, a Downloadable Software Package for Processing and Analyses of Circular Dichroism Spectroscopic Data ». *Protein Science: A Publication of the Protein Society* 27 (9): 1717-22. https://doi.org/10.1002/pro.3474.

Miles, Andrew J., Frank Wien, et B. A. Wallace. 2004. « Redetermination of the Extinction Coefficient of Camphor-10-Sulfonic Acid, a Calibration Standard for Circular Dichroism Spectroscopy ». *Analytical Biochemistry* 335 (2): 338-39. https://doi.org/10.1016/j.ab.2004.08.035.

Møller, Marie Sofie, Anette Henriksen, et Birte Svensson. 2016. « Structure and Function of α-Glucan Debranching Enzymes ». *Cellular and Molecular Life Sciences: CMLS* 73 (14): 2619-41. https://doi.org/10.1007/s00018-016-2241-y.

Monroe, Jonathan D., Amanda R. Storm, Elizabeth M. Badley, Michael D. Lehman, Samantha M. Platt, Lauren K. Saunders, Jonathan M. Schmitz, et Catherine E. Torres. 2014. « β-Amylase1 and β-Amylase3 Are Plastidic Starch Hydrolases in Arabidopsis That Seem to Be Adapted for Different Thermal, PH, and Stress Conditions ». *Plant Physiology* 166 (4): 1748-63. https://doi.org/10.1104/pp.114.246421.

Morris, Daniel Luzon, et Carol Tilden Morris. 1939. « GLYCOGEN IN THE SEED OF ZEA MAYS (VARIETY GOLDEN BANTAM) ». *Journal of Biological Chemistry* 130 (2): 535-44. https://doi.org/10.1016/S0021-9258(18)73524-2.

Mouille, G., M. L. Maddelein, N. Libessart, P. Talaga, A. Decq, B. Delrue, et S. Ball. 1996. « Preamylopectin Processing: A Mandatory Step for Starch Biosynthesis in Plants ». *The Plant Cell* 8 (8): 1353-66. https://doi.org/10.1105/tpc.8.8.1353.

Mu-Forster, Chen, Rongmin Huang, Joseph R Powers, Robert W Harriman, Mary Knight, Ceorge W Singletary, et Bruce P Wasserman. s. d. « 7 Physical Association of Starch Biosynthetic Enz Mes with Starch Granules of Maize Endosperm ».

Mugford, Sam T., Olivier Fernandez, Jemima Brinton, Anna Flis, Nicole Krohn, Beatrice Encke, Regina Feil, et al. 2014. « Regulatory Properties of ADP Glucose Pyrophosphorylase Are Required for Adjustment of Leaf Starch Synthesis in Different Photoperiods ». *Plant Physiology* 166 (4): 1733-47. https://doi.org/10.1104/pp.114.247759.

Nakamura, Y., M. Ono, C. Utsumi, et M. Steup. 2012. « Functional Interaction Between Plastidial Starch Phosphorylase and Starch Branching Enzymes from Rice During the Synthesis of Branched Maltodextrins ». *Plant and Cell Physiology* 53 (5): 869-78. https://doi.org/10.1093/pcp/pcs030.

Nakamura, Yasunori. 1996. « Some Properties of Starch Debranching Enzymes and Their Possible Role in Amylopectin Biosynthesis ». *Plant Science* 121 (1): 1-18. https://doi.org/10.1016/S0168-9452(96)04504-9.

Nakamura, Yasunori, Masami Ono, Takayuki Sawada, Naoko Crofts, Naoko Fujita, et Martin Steup. 2017. « Characterization of the Functional Interactions of Plastidial Starch Phosphorylase and Starch Branching Enzymes from Rice Endosperm during Reserve Starch Biosynthesis ». *Plant Science* 264 (novembre): 83-95. https://doi.org/10.1016/j.plantsci.2017.09.002.

Nielsen, Tom H., Lone Baunsgaard, et Andreas Blennow. 2002. « Intermediary Glucan Structures Formed during Starch Granule Biosynthesis Are Enriched in Short Side Chains, a Dynamic Pulse Labeling Approach* ». *Journal of Biological Chemistry* 277 (23): 20249-55. https://doi.org/10.1074/jbc.M201866200.

Nitschke, Felix, Peixiang Wang, Peter Schmieder, Jean-Marie Girard, Donald E. Awrey, Tony Wang, Johan Israelian, et al. 2013. « Hyperphosphorylation of Glucosyl C6 Carbons and Altered Structure of Glycogen in the Neurodegenerative Epilepsy Lafora Disease ». *Cell Metabolism* 17 (5): 756-67. https://doi.org/10.1016/j.cmet.2013.04.006.

O'Neill, Ellis, et Robert Field. 2015. « Underpinning Starch Biology with in vitro Studies on Carbohydrate-Active Enzymes and Biosynthetic Glycomaterials ». *Frontiers in Bioengineering and Biotechnology* 3 (septembre): 136. https://doi.org/10.3389/fbioe.2015.00136.

Pal, Sunil Kumar, Magdalena Liput, Maria Piques, Hirofumi Ishihara, Toshihiro Obata, Marina C.M. Martins, Ronan Sulpice, et al. 2013. « Diurnal Changes of Polysome Loading Track Sucrose Content in the Rosette of Wild-Type Arabidopsis and the Starchless *Pgm* Mutant ». *Plant Physiology* 162 (3): 1246-65. https://doi.org/10.1104/pp.112.212258.

Palmer, T. Norman, Lynne E. Macaskie, et Kanwaljit K. Grewel. 1983. « The Unit-Chain Distribution Profiles of Branched (1→4)-α-d-Glucans ». *Carbohydrate Research* 114 (2): 338-42. https://doi.org/10.1016/0008-6215(83)88205-6.

Pan, D. D., et J. I. Jane. 2000. « Internal Structure of Normal Maize Starch Granules Revealed by Chemical Surface Gelatinization ». *Biomacromolecules* 1 (1): 126-32. https://doi.org/10.1021/bm990016l.

Perez Herrera, Mariana, Thava Vasanthan, et Ratnajothi Hoover. 2016. « Characterization of Maize Starch Nanoparticles Prepared by Acid Hydrolysis ». *Cereal Chemistry* 93 (3): 323-30. https://doi.org/10.1094/CCHEM-08-15-0175-R.

Pérez, Serge, Paul M. Baldwin, et Daniel J. Gallant. 2009. « Chapter 5 - Structural Features of Starch Granules I ». In *Starch (Third Edition)*, édité par James BeMiller et Roy Whistler, 149-92. Food Science and Technology. San Diego: Academic Press. https://doi.org/10.1016/B978-0-12-746275-2.00005-7.

Petoukhov, Maxim V., et Dmitri I. Svergun. 2005a. « Global Rigid Body Modeling of Macromolecular Complexes against Small-Angle Scattering Data ». *Biophysical Journal* 89 (2): 1237-50. https://doi.org/10.1529/biophysj.105.064154.

———. 2005b. « Global Rigid Body Modeling of Macromolecular Complexes against Small-Angle Scattering Data ». *Biophysical Journal* 89 (2): 1237-50. https://doi.org/10.1529/biophysj.105.064154.

Pfannemüller, Beate. 1987. « Influence of Chain Length of Short Monodisperse Amyloses on the Formation of A- and B-Type X-Ray Diffraction Patterns ». *International Journal of Biological Macromolecules* 9 (2): 105-8. https://doi.org/10.1016/0141-8130(87)90034-1.

Pfister, Barbara, Kuan-Jen Lu, Simona Eicke, Regina Feil, John E. Lunn, Sebastian Streb, et Samuel C. Zeeman. 2014. « Genetic Evidence That Chain Length and Branch Point Distributions Are Linked Determinants of Starch Granule Formation in Arabidopsis ». *Plant Physiology* 165 (4): 1457-74. https://doi.org/10.1104/pp.114.241455.

Pfister, Barbara, Antoni Sánchez-Ferrer, Ana Diaz, Kuanjen Lu, Caroline Otto, Mirko Holler, Farooque Razvi Shaik, Florence Meier, Raffaele Mezzenga, et Samuel C. Zeeman. 2016. « Recreating the Synthesis of Starch Granules in Yeast ». *ELife* 5 (novembre): e15552. https://doi.org/10.7554/eLife.15552.

Pfister, Barbara, et Samuel C. Zeeman. 2016. « Formation of starch in plant cells ». *Cellular and Molecular Life Sciences* 73: 2781-2807. https://doi.org/10.1007/s00018-016-2250-x.

Putaux, Jean-Luc, Gabrielle Potocki-Véronèse, Magali Remaud-Simeon, et Alain Buleon. 2006. « α-D-Glucan-Based Dendritic Nanoparticles Prepared by in Vitro Enzymatic Chain Extension of Glycogen ». *Biomacromolecules* 7 (6): 1720-28. https://doi.org/10.1021/bm050988v.

Refregiers, M., F. Wien, H.-P. Ta, L. Premvardhan, S. Bac, F. Jamme, V. Rouam, et al. 2012. « DISCO Synchrotron-Radiation Circular-Dichroism Endstation at SOLEIL ». *Journal of Synchrotron Radiation* 19 (5): 831-35. https://doi.org/10.1107/S0909049512030002.

Regina, Ahmed, Anthony Bird, David Topping, Sarah Bowden, Judy Freeman, Tina Barsby, Behjat Kosar-Hashemi, Zhongyi Li, Sadequr Rahman, et Matthew Morell. 2006. « High-Amylose Wheat Generated by RNA Interference Improves Indices of Large-Bowel Health in Rats ». *Proceedings of the National Academy of Sciences of the United States of America* 103 (10): 3546-51. https://doi.org/10.1073/pnas.0510737103.

Regina, Ahmed, Behjat Kosar-Hashemi, Zhongyi Li, Andrew Pedler, Yasuhiko Mukai, Maki Yamamoto, Kevin Gale, Peter J. Sharp, Matthew K. Morell, et Sadequr Rahman. 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 (5): 899-909. https://doi.org/10.1007/s00425-005-0032-z.

Regina, Ahmed, Behjat Kosar-Hashemi, Zhongyi Li, Lynette Rampling, Mark Cmiel, Maria C. Gianibelli, Christine Konik-Rose, Oscar Larroque, Sadequr Rahman, et Matthew K. Morell. 2004. « Multiple Isoforms of Starch Branching Enzyme-I in Wheat: Lack of the Major SBE-I Isoform Does Not Alter Starch Phenotype ». *Functional Plant Biology: FPB* 31 (6): 591-601. https://doi.org/10.1071/FP03193.

Ritte, Gerhard, Matthias Heydenreich, Sebastian Mahlow, Sophie Haebel, Oliver Kötting, et Martin Steup. 2006. « Phosphorylation of C6- and C3-Positions of Glucosyl Residues in Starch Is Catalysed by Distinct Dikinases ». *FEBS Letters* 580 (20): 4872-76. https://doi.org/10.1016/j.febslet.2006.07.085.

Roach, Peter J. 2002. « Glycogen and Its Metabolism ». *Current Molecular Medicine* 2 (2): 101-20. https://doi.org/10.2174/1566524024605761.

Roach, Peter J., Anna A. Depaoli-Roach, Thomas D. Hurley, et Vincent S. Tagliabracci. 2012. « Glycogen and its metabolism: some new developments and old themes ». *The Biochemical journal* 441 (3): 763-87. https://doi.org/10.1042/BJ20111416.

Robyt, John, et Dexter French. 1963. « Action Pattern and Specificity of an Amylase from Bacillus Subtilis ». *Archives of Biochemistry and Biophysics* 100 (3): 451-67. https://doi.org/10.1016/0003-9861(63)90112-7.

Saccomanno, Benedetta, Alan H. Chambers, Alec Hayes, Ian Mackay, Simon C. McWilliam, et Kay Trafford. 2017. « Starch Granule Morphology in Oat Endosperm ». *Journal of Cereal Science* 73 (janvier): 46-54. https://doi.org/10.1016/j.jcs.2016.10.011.

Samsudin, Hayati, et Norziah M. Hani. 2017. « Chapter 8 - Use of Starch in Food Packaging ». In *Starch-Based Materials in Food Packaging*, édité par Marcelo A. Villar, Silvia E. Barbosa, M. Alejandra García, Luciana A. Castillo, et Olivia V. López, 229-56. Academic Press. https://doi.org/10.1016/B978-0-12-809439-6.00008-X.

Santelia, Diana, Oliver Kötting, David Seung, Mario Schubert, Matthias Thalmann, Sylvain Bischof, David A. Meekins, et al. 2011. « The Phosphoglucan Phosphatase like Sex Four2 Dephosphorylates Starch at the C3-Position in Arabidopsis ». *The Plant Cell* 23 (11): 4096-4111. https://doi.org/10.1105/tpc.111.092155.

Satoh, Hikaru, Aiko Nishi, Kazuhiro Yamashita, Yoko Takemoto, Yasumasa Tanaka, Yuko Hosaka, Aya Sakurai, Naoko Fujita, et Yasunori Nakamura. 2003. « Starch-Branching Enzyme I-Deficient Mutation Specifically Affects the Structure and Properties of Starch in Rice Endosperm ». *Plant Physiology* 133 (3): 1111-21. https://doi.org/10.1104/pp.103.021527.

Schwall, G. P., R. Safford, R. J. Westcott, R. Jeffcoat, A. Tayal, Y. C. Shi, M. J. Gidley, et S. A. Jobling. 2000. « Production of Very-High-Amylose Potato Starch by Inhibition of SBE A and B ». *Nature Biotechnology* 18 (5): 551-54. https://doi.org/10.1038/75427.

Scialdone, Antonio, Sam T Mugford, Doreen Feike, Alastair Skeffington, Philippa Borrill, Alexander Graf, Alison M Smith, et Martin Howard. 2013. « Arabidopsis plants perform arithmetic division to prevent starvation at night ». Édité par Detlef Weigel. *eLife* 2 (juin): e00669. https://doi.org/10.7554/eLife.00669.

Seung, David, Julien Boudet, Jonathan Monroe, Tina B. Schreier, Laure C. David, Melanie Abt, Kuan-Jen Lu, Martina Zanella, et Samuel C. Zeeman. 2017. « Homologs of PROTEIN TARGETING TO STARCH Control Starch Granule Initiation in Arabidopsis Leaves ». *The Plant Cell* 29 (7): 1657-77. https://doi.org/10.1105/tpc.17.00222.

Seung, David, Alberto Echevarría-Poza, Burkhard Steuernagel, et Alison M. Smith. 2020. « Natural Polymorphisms in Arabidopsis Result in Wide Variation or Loss of the Amylose Component of Starch1[OPEN] ». *Plant Physiology* 182 (2): 870-81. https://doi.org/10.1104/pp.19.01062.

Seung, David, Tina B. Schreier, Léo Bürgy, Simona Eicke, et Samuel C. Zeeman. 2018. « Two Plastidial Coiled-Coil Proteins Are Essential for Normal Starch Granule Initiation in Arabidopsis[OPEN] ». *The Plant Cell* 30 (7): 1523-42. https://doi.org/10.1105/tpc.18.00219.

Seung, David, et Alison M. Smith. 2019. « Starch Granule Initiation and Morphogenesis-Progress in Arabidopsis and Cereals ». *Journal of Experimental Botany* 70 (3): 771-84. https://doi.org/10.1093/jxb/ery412.

Seung, David, Sebastian Soyk, Mario Coiro, Benjamin A. Maier, Simona Eicke, et Samuel C. Zeeman. 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 (2): e1002080. https://doi.org/10.1371/journal.pbio.1002080.

Shen, Xinyu, Eric Bertoft, Genyi Zhang, et Bruce R. Hamaker. 2013. « Iodine Binding to Explore the Conformational State of Internal Chains of Amylopectin ». *Carbohydrate Polymers* 98 (1): 778-83. https://doi.org/10.1016/j.carbpol.2013.06.050.

Shure, M., S. Wessler, et N. Fedoroff. 1983. « Molecular Identification and Isolation of the Waxy Locus in Maize ». *Cell* 35 (1): 225-33. https://doi.org/10.1016/0092-8674(83)90225-8.

Silver, Dylan M., Oliver Kötting, et Greg B. G. Moorhead. 2014. « Phosphoglucan Phosphatase Function Sheds Light on Starch Degradation ». *Trends in Plant Science* 19 (7): 471-78. https://doi.org/10.1016/j.tplants.2014.01.008.

Sim, Lyann, Sophie R. Beeren, Justin Findinier, David Dauvillée, Steven G. Ball, Anette Henriksen, et Monica M. Palcic. 2014. « Crystal Structure of the Chlamydomonas Starch Debranching Enzyme Isoamylase ISA1 Reveals Insights into the Mechanism of Branch Trimming and Complex Assembly* ». *Journal of Biological Chemistry* 289 (33): 22991-3. https://doi.org/10.1074/jbc.M114.565044.

Singh, Aakanksha, Julia Compart, Shadha Abduljaleel AL-Rawi, Harendra Mahto, Abubakar Musa Ahmad, et Joerg Fettke. 2022. « LIKE EARLY STARVATION 1 Alters the Glucan Structures at the Starch Granule Surface and Thereby Influences the Action of Both Starch-Synthesizing and Starch-Degrading Enzymes ». *The Plant Journal* 111 (3): 819-35. https://doi.org/10.1111/tpj.15855.

Smirnova, Julia, Alisdair R. Fernie, et Martin Steup. 2015. « Starch Degradation ». In *Starch*, édité par Yasunori Nakamura, 239-90. Tokyo: Springer Japan. https://doi.org/10.1007/978-4-431-55495-0_7.

Smith, Alison M., Samuel C. Zeeman, et Steven M. Smith. 2005. « Starch Degradation ». *Annual Review of Plant Biology* 56 (1): 73-98. https://doi.org/10.1146/annurev.arplant.56.032604.144257.

Smythe, Carl, et Philip Cohen. 1991. « The Discovery of Glycogenin and the Priming Mechanism for Glycogen Biogenesis ». *European Journal of Biochemistry* 200 (3): 625-31. https://doi.org/10.1111/j.1432-1033.1991.tb16225.x.

Song, Kijong, Dae-Woo Lee, Jeongheon Kim, Jaewook Kim, Hwanuk Guim, Keunhwa Kim, Jong-Seong Jeon, et Giltsu Choi. 2021. « EARLY STARVATION 1 Is a Functionally Conserved Protein

Promoting Gravitropic Responses in Plants by Forming Starch Granules ». *Frontiers in Plant Science* 12. https://www.frontiersin.org/articles/10.3389/fpls.2021.628948.

Stitt, Mark, et Samuel C Zeeman. 2012. « Starch Turnover: Pathways, Regulation and Role in Growth ». *Current Opinion in Plant Biology* 15 (3): 282-92. https://doi.org/10.1016/j.pbi.2012.03.016.

Streb, Sebastian, Thierry Delatte, Martin Umhang, Simona Eicke, Martine Schorderet, Didier Reinhardt, et Samuel C. Zeeman. 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 (12): 3448-66. https://doi.org/10.1105/tpc.108.063487.

Streb, Sebastian, et Samuel C. Zeeman. 2012. « Starch Metabolism in Arabidopsis ». *The Arabidopsis Book / American Society of Plant Biologists* 10 (septembre): e0160. https://doi.org/10.1199/tab.0160.

———. 2014. « 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* 9 (3): e92174. https://doi.org/10.1371/journal.pone.0092174.

Sulpice, Ronan, Eva-Theresa Pyl, Hirofumi Ishihara, Sandra Trenkamp, Matthias Steinfath, Hanna Witucka-Wall, Yves Gibon, et al. 2009. « Starch as a major integrator in the regulation of plant growth ». *Proceedings of the National Academy of Sciences of the United States of America* 106 (25): 10348-53. https://doi.org/10.1073/pnas.0903478106.

Svergun, D. I. 1992. « Determination of the Regularization Parameter in Indirect-Transform Methods Using Perceptual Criteria ». *Journal of Applied Crystallography* 25 (4): 495-503. https://doi.org/10.1107/S0021889892001663.

Svergun, Dmitri I., Maxim V. Petoukhov, et Michel H. J. Koch. 2001. « Determination of Domain Structure of Proteins from X-Ray Solution Scattering ». *Biophysical Journal* 80 (6): 2946-53. https://doi.org/10.1016/S0006-3495(01)76260-1.

Szydlowski, Nicolas, Paula Ragel, Sandy Raynaud, et M Mercedes Lucas. 2009. « Starch Granule Initiation in Arabidopsis Requires the Presence of Either Class IV or Class III Starch Synthases W ».

Takeda, Yasuhito, Susumu Hizukuri, Chieno Takeda, et Ayako Suzuki. 1987. « Structures of Branched Molecules of Amyloses of Various Origins, and Molar Fractions of Branched and Unbranched Molecules ». *Carbohydrate Research* 165 (1): 139-45. https://doi.org/10.1016/0008-6215(87)80089-7.

Tanaka, Naoki, Naoko Fujita, Aiko Nishi, Hikaru Satoh, Yuko Hosaka, Masashi Ugaki, Shinji Kawasaki, et Yasunori Nakamura. 2004. « The Structure of Starch Can Be Manipulated by Changing the Expression Levels of Starch Branching Enzyme IIb in Rice Endosperm ». *Plant Biotechnology Journal* 2 (6): 507-16. https://doi.org/10.1111/j.1467-7652.2004.00097.x.

Tang, Hanjun, Toshio Mitsunaga, et Yukio Kawamura. 2006. « Molecular Arrangement in Blocklets and Starch Granule Architecture ». *Carbohydrate Polymers* 63 (4): 555-60. https://doi.org/10.1016/j.carbpol.2005.10.016.

Tawil, Georges, Frédéric Jamme, Matthieu Réfrégiers, Anders Viksø-Nielsen, Paul Colonna, et Alain Buléon. 2011. « In Situ Tracking of Enzymatic Breakdown of Starch Granules by Synchrotron UV Fluorescence Microscopy ». *Analytical Chemistry* 83 (3): 989-93. https://doi.org/10.1021/ac1027512.

Tester, R. F., et W. R. Morrison. 1992. « Swelling and Gelatinization of Cereal Starches. III. Some Properties of Waxy and Normal Nonwaxy Barley Starches ». *Cereal Chemistry.* 69 (6): 654-58.

Tester, Richard, et W. Morrison. 1990. « Swelling and gelatinization of cereal starches. I. Effects of amylopectin, amylose and lipids ». *Cereal Chemistry* 67 (janvier): 551-57.

Tetlow, Ian J., Kim G. Beisel, Scott Cameron, Amina Makhmoudova, Fushan Liu, Nicole S. Bresolin, Robin Wait, Matthew K. Morell, et Michael J. Emes. 2008. « Analysis of Protein Complexes in Wheat Amyloplasts Reveals Functional Interactions among Starch Biosynthetic Enzymes ». *Plant Physiology* 146 (4): 1878-91. https://doi.org/10.1104/pp.108.116244.

Tetlow, Ian J., et Eric Bertoft. 2020a. « A Review of Starch Biosynthesis in Relation to the Building Block-Backbone Model ». *International Journal of Molecular Sciences* 21 (19): 7011. https://doi.org/10.3390/ijms21197011.

———. 2020b. « A Review of Starch Biosynthesis in Relation to the Building Block-Backbone Model ». *International Journal of Molecular Sciences* 21 (19): 7011. https://doi.org/10.3390/ijms21197011.

Tetlow, Ian J., et Michael J. Emes. 2014a. « A Review of Starch-Branching Enzymes and Their Role in Amylopectin Biosynthesis ». *IUBMB Life* 66 (8): 546-58. https://doi.org/10.1002/iub.1297.

———. 2014b. « A Review of Starch-Branching Enzymes and Their Role in Amylopectin Biosynthesis ». *IUBMB Life* 66 (8): 546-58. https://doi.org/10.1002/iub.1297.

———. 2017. « Starch Biosynthesis in the Developing Endosperms of Grasses and Cereals ». *Agronomy* 7 (4): 81. https://doi.org/10.3390/agronomy7040081.

Tiessen, Axel, Janneke H. M. Hendriks, Mark Stitt, Anja Branscheid, Yves Gibon, Eva M. Farré, et Peter Geigenberger. 2002. « Starch Synthesis in Potato Tubers Is Regulated by Post-Translational Redox Modification of ADP-Glucose Pyrophosphorylase: A Novel Regulatory Mechanism Linking Starch Synthesis to the Sucrose Supply ». *The Plant Cell* 14 (9): 2191-2213. https://doi.org/10.1105/tpc.003640.

Tsai, C. Y., F. Salamini, et O. E. Nelson. 1970. « Enzymes of Carbohydrate Metabolism in the Developing Endosperm of Maize 1 ». *Plant Physiology* 46 (2): 299-306.

Tunçay, Hande, Justin Findinier, Thierry Duchêne, Virginie Cogez, Charlotte Cousin, Gilles Peltier, Steven G. Ball, et David Dauvillée. 2013. « A Forward Genetic Approach in Chlamydomonas

Reinhardtii as a Strategy for Exploring Starch Catabolism ». *PloS One* 8 (9): e74763. https://doi.org/10.1371/journal.pone.0074763.

Tyynelä, Janna, et Alan Schulman. 1993. « An analysis of soluble starch synthase isozymes from the developing grains of normal and shx cv. Bomi barley (Hordeum vulgare) ». *Physiologia Plantarum* 89 (avril): 835-41. https://doi.org/10.1111/j.1399-3054.1993.tb05293.x.

Umeki, K., et T. Yamamoto. 1972. « Enzymatic Determination of Structure of Singly Branched Hexaose Dextrins Formed by Liquefying -Amylase of Bacillus Subtilis ». *Journal of Biochemistry* 72 (1): 101-9. https://doi.org/10.1093/oxfordjournals.jbchem.a129874.

Umemoto, Takayuki, M Yano, H Satoh, A Shomura, et Yasunori Nakamura. 2002. « Mapping of a gene responsible for the difference in amylopectin structure between ». *TAG. Theoretical and applied genetics. Theoretische und angewandte Genetik* 104 (janvier): 1-8. https://doi.org/10.1007/s001220200000.

Vamadevan, Varatharajan, Eric Bertoft, et Koushik Seetharaman. 2013. « On the Importance of Organization of Glucan Chains on Thermal Properties of Starch ». *Carbohydrate Polymers* 92 (2): 1653-59. https://doi.org/10.1016/j.carbpol.2012.11.003.

Vamadevan, Varatharajan, Ratnajothi Hoover, Eric Bertoft, et Koushik Seetharaman. 2014. « Hydrothermal Treatment and Iodine Binding Provide Insights into the Organization of Glucan Chains within the Semi-crystalline Lamellae of Corn Starch Granules ». Édité par C. Allen Bush. *Biopolymers* 101 (8): 871-85. https://doi.org/10.1002/bip.22468.

Vandromme, Camille, Corentin Spriet, David Dauvillée, Adeline Courseaux, Jean-Luc Putaux, Adeline Wychowski, Frédéric Krzewinski, Maud Facon, Christophe D'Hulst, et Fabrice Wattebled. 2019. « PII1: A Protein Involved in Starch Initiation That Determines Granule Number and Size in Arabidopsis Chloroplast ». *New Phytologist* 221 (1): 356-70. https://doi.org/10.1111/nph.15356.

Volkov, Vladimir V., et Dmitri I. Svergun. 2003. « Uniqueness of Ab Initio Shape Determination in Small-Angle Scattering ». *Journal of Applied Crystallography* 36 (3-1): 860-64. https://doi.org/10.1107/S0021889803000268.

Walters, Robin G., Douglas G. Ibrahim, Peter Horton, et Nicholas J. Kruger. 2004. « A Mutant of Arabidopsis Lacking the Triose-Phosphate/Phosphate Translocator Reveals Metabolic Regulation of Starch Breakdown in the Light ». *Plant Physiology* 135 (2): 891-906. https://doi.org/10.1104/pp.104.040469.

Wang, Juan, Pan Hu, Lingshang Lin, Zichun Chen, Qiaoquan Liu, et Cunxu Wei. 2018. « Gradually Decreasing Starch Branching Enzyme Expression Is Responsible for the Formation of Heterogeneous Starch Granules ». *Plant Physiology* 176 (1): 582-95. https://doi.org/10.1104/pp.17.01013.

Wang, Kai, Robert J. Henry, et Robert G. Gilbert. 2014. « Causal Relations Among Starch Biosynthesis, Structure, and Properties ». *Springer Science Reviews* 2 (1): 15-33. https://doi.org/10.1007/s40362-014-0016-0.

Waterschoot, Jasmien, Sara V. Gomand, Ellen Fierens, et Jan A. Delcour. 2015. « Production, Structure, Physicochemical and Functional Properties of Maize, Cassava, Wheat, Potato and Rice Starches ». *Starch - Stärke* 67 (1-2): 14-29. https://doi.org/10.1002/star.201300238.

Wattebled, Fabrice, Ying Dong, Sylvain Dumez, David Delvallé, Véronique Planchot, Pierre Berbezy, Darshna Vyas, et al. 2005. « Mutants of Arabidopsis Lacking a Chloroplastic Isoamylase Accumulate Phytoglycogen and an Abnormal Form of Amylopectin ». *Plant Physiology* 138 (1): 184-95. https://doi.org/10.1104/pp.105.059295.

Wattebled, Fabrice, Véronique Planchot, Ying Dong, Nicolas Szydlowski, Bruno Pontoire, Aline Devin, Steven Ball, et Christophe D'Hulst. 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 (3): 1309-23. https://doi.org/10.1104/pp.108.129379.

Wayllace, Nahuel Z., Hugo A. Valdez, Rodolfo A. Ugalde, Maria V. Busi, et Diego F. Gomez-Casati. 2010. « The Starch-Binding Capacity of the Noncatalytic SBD2 Region and the Interaction between the N- and C-Terminal Domains Are Involved in the Modulation of the Activity of Starch Synthase III from Arabidopsis Thaliana ». *The FEBS Journal* 277 (2): 428-40. https://doi.org/10.1111/j.1742-4658.2009.07495.x.

Wei, Yang, Aby Thyparambil, et Robert Latour. 2014. « Protein Helical Structure Determination Using CD Spectroscopy for Solutions with Strong Background Absorbance from 190-230 nm ». *Biochimica et biophysica acta* 1844 (octobre). https://doi.org/10.1016/j.bbapap.2014.10.001.

Weise, Sean E., Andreas P. M. Weber, et Thomas D. Sharkey. 2004. « Maltose Is the Major Form of Carbon Exported from the Chloroplast at Night ». *Planta* 218 (3): 474-82. https://doi.org/10.1007/s00425-003-1128-y.

Wikman, Jeanette, Andreas Blennow, Alain Buléon, Jean-Luc Putaux, Serge Pérez, Koushik Seetharaman, et Eric Bertoft. 2014. « Influence of Amylopectin Structure and Degree of Phosphorylation on the Molecular Composition of Potato Starch Lintners ». *Biopolymers* 101 (3): 257-71. https://doi.org/10.1002/bip.22344.

Wilkens, Casper, Jose Cuesta-Seijo, Monica Palcic, et Birte Svensson. 2014. « Selectivity of the surface binding site (SBS) on barley starch synthase I ». *Biologia* 69 (septembre): 1118-21. https://doi.org/10.2478/s11756-014-0418-0.

Wilson, Wayne A., Peter J. Roach, Manuel Montero, Edurne Baroja-Fernández, Francisco José Muñoz, Gustavo Eydallin, Alejandro M. Viale, et Javier Pozueta-Romero. 2010. « Regulation of glycogen metabolism in yeast and bacteria ». *FEMS microbiology reviews* 34 (6): 952-85. https://doi.org/10.1111/j.1574-6976.2010.00220.x.

Wong, K. S., et J. Jane. 1997. « Quantitative Analysis of Debranched Amylopectin by HPAEC-PAD with a Postcolumn Enzyme Reactor ». *Journal of Liquid Chromatography & Related Technologies* 20 (2): 297-310. https://doi.org/10.1080/10826079708010654.

Yu, Shiyao, Fangdong Zhang, Cheng Li, et Robert G. Gilbert. 2017. « Molecular Structural Differences between Maize Leaf and Endosperm Starches ». *Carbohydrate Polymers* 161 (avril): 10-15. https://doi.org/10.1016/j.carbpol.2016.12.064.

Yun, Min-Soo, Takayuki Umemoto, et Yasushi Kawagoe. 2011. « Rice Debranching Enzyme Isoamylase3 Facilitates Starch Metabolism and Affects Plastid Morphogenesis ». *Plant and Cell Physiology* 52 (6): 1068-82. https://doi.org/10.1093/pcp/pcr058.

Zeeman, Samuel C, Takayuki Umemoto, Wei-Ling Lue, Pui Au-Yeung, Cathie Martin, Alison M Smith, et Jychian Chen. 1998. « A Mutant of Arabidopsis Lacking a Chloroplastic Isoamylase Accumulates Both Starch and Phytoglycogen ».

Zeqiraj, Elton, Xiaojing Tang, Roger W. Hunter, Mar García-Rocha, Andrew Judd, Maria Deak, Alexander von Wilamowitz-Moellendorff, et al. 2014. « Structural basis for the recruitment of glycogen synthase by glycogenin ». *Proceedings of the National Academy of Sciences* 111 (28): E2831-40. https://doi.org/10.1073/pnas.1402926111.

Zhang, Xiaoli, Christophe Colleoni, Vlada Ratushna, Mirella Sirghie-Colleoni, Martha G. James, et Alan M. Myers. 2004. « Molecular Characterization Demonstrates That the Zea Mays Gene Sugary2 Codes for the Starch Synthase Isoform SSIIa ». *Plant Molecular Biology* 54 (6): 865-79. https://doi.org/10.1007/s11103-004-0312-1.

Zhang, Yachuan, Curtis Rempel, et Derek McLaren. 2014. « Chapter 16 - Thermoplastic Starch ». In *Innovations in Food Packaging (Second Edition)*, édité par Jung H. Han, 391-412. Food Science and Technology. San Diego: Academic Press. https://doi.org/10.1016/B978-0-12-394601-0.00016-3.

Zhu, Fan, Eric Bertoft, You Wang, Michael Emes, Ian Tetlow, et Koushik Seetharaman. 2015. « Structure of Arabidopsis Leaf Starch Is Markedly Altered Following Nocturnal Degradation ». *Carbohydrate Polymers* 117 (mars): 1002-13. https://doi.org/10.1016/j.carbpol.2014.09.092.

Zobel, H. F., A. D. French, et M. E. Hinkle. 1967. « X-Ray Diffraction of Oriented Amylose Fibers. II. Structure of V Amyloses ». *Biopolymers* 5 (9): 837-45. https://doi.org/10.1002/bip.1967.360050906.

Publication

LIKE EARLY STARVATION 1 and EARLY STARVATION 1 Promote and Stabilize Amylopectin Phase Transition in Starch Biosynthesis

Chun Liu, Barbara Pfister, Rayan Osman, Maximilian Ritter, Arvid Heutinck, Mayank Sharma, Simona Eicke, Michaela Fischer-Stettler, David Seung, Coralie Bompard, Melanie R. Abt, Samuel C. Zeeman.

Science Advances, 2023 In press

PLANT SCIENCES

LIKE EARLY STARVATION 1 and EARLY STARVATION 1 promote and stabilize amylopectin phase transition in starch biosynthesis

Chun Liu[1], Barbara Pfister[1], Rayan Osman[2], Maximilian Ritter[3], Arvid Heutinck[1], Mayank Sharma[1], Simona Eicke[1], Michaela Fischer-Stettler[1], David Seung[1]†, Coralie Bompard[2], Melanie R. Abt[1], Samuel C. Zeeman[1]*

Starch, the most abundant carbohydrate reserve in plants, primarily consists of the branched glucan amylopectin, which forms semi-crystalline granules. Phase transition from a soluble to an insoluble form depends on amylopectin architecture, requiring a compatible distribution of glucan chain lengths and a branch-point distribution. Here, we show that two starch-bound proteins, LIKE EARLY STARVATION 1 (LESV) and EARLY STARVATION 1 (ESV1), which have unusual carbohydrate-binding surfaces, promote the phase transition of amylopectin-like glucans, both in a heterologous yeast system expressing the starch biosynthetic machinery and in *Arabidopsis* plants. We propose a model wherein LESV serves as a nucleating role, with its carbohydrate-binding surfaces helping align glucan double helices to promote their phase transition into semi-crystalline lamellae, which are then stabilized by ESV1. Because both proteins are widely conserved, we suggest that protein-facilitated glucan crystallization may be a general and previously unrecognized feature of starch biosynthesis.

INTRODUCTION

Starch is the most abundant and widespread nonstructural carbohydrate in higher plants. In seeds, roots, and other perennating tissues, starch serves as a long-term carbohydrate reserve, fuelling germination or seasonal regrowth. This "storage starch" is a major constituent of our staple crops (e.g., cereals, potato, cassava, legumes, among others), accounting for half of human calorific intake, and is used extensively as an industrial commodity in the food, nonfood, and energy sectors (*1*). Starch is also synthesized in plant leaves during the day as a primary product of photosynthesis and is remobilized at night to sustain metabolism. This so-called "transitory starch" is very important for plant growth, as demonstrated by the reduced vigor of plants unable to make or degrade it effectively (*2*).

Starch is composed of two glucose polymers—amylopectin and amylose—and forms insoluble, semi-crystalline granules. The latter characteristic is conferred by the major component, amylopectin, a branched macromolecule consisting of linear chains of α-1,4–linked glucose units with α-1,6 branch points. The branched architecture is thought to give the molecule an overall tree-like appearance, featuring interconnected clusters of linear chain segments (*3*). Within this molecule, secondary and tertiary structures form, as pairs of adjacent chains entwine into double helices that align and pack into dense, crystalline lamellae. These crystalline lamellae alternate with amorphous lamellae harboring the branch points and the chains connecting the crystalline layers within the starch granule. Together, this regular arrangement of crystalline (~6 nm) and amorphous (~3 nm) layers is believed to underlie the 9- to 10-nm repeat structure that is commonly observed for plant starches (*3*, *4*).

Three major enzyme classes, working in an interdependent and concerted fashion, are involved in amylopectin biosynthesis: First, soluble starch synthases (SSs) elongate existing glucose chains by transferring the glucosyl moiety from the substrate adenosine diphosphate (ADP)–glucose (ADP-Glc), thereby forming an additional α-1,4 bond. Second, branching enzymes (BEs) create α-1,6 branches by transferring small glucan segments from an existing linear chain to the C6 position of another glucose unit elsewhere in the glucan molecule (*5*). Third, debranching enzymes of the isoamylase (ISA1) class hydrolyse some of the introduced branches (*6*). This presumably serves to remove surplus or misplaced branches that interfere with the formation of higher-order amylopectin structures and thereby facilitates amylopectin crystallization (*7*–*9*). This role of ISA1 is evident from the phenotypes of mutant plants lacking it; for example, in *Arabidopsis* (where ISA1 associates with a related, nonenzymatic ISA2 subunit to form a single heteromultimeric isoamylase), the loss of ISA1 activity causes a reduction in leaf starch together with the accumulation of substantial amounts of an aberrant soluble glucan called phytoglycogen (named according to its similarity to glycogen) in its place (*9*–*11*). Compared to amylopectin, phytoglycogen is enriched in short chains and has more branch points, which are located closer together, presumably because branches that remain would normally be removed by ISA1. Its suboptimal branching pattern is proposed to hinder the transition into a semi-crystalline state, and, as a consequence, phytoglycogen is soluble. Note although that besides phytoglycogen, *Arabidopsis isa1* mutants generate a variety of different glucans, including normal-looking starch granules, albeit with an altered amylopectin structure, in certain cell types (*9*, *12*).

Starch biosynthesis is highly conserved between plant species. Recently, the pathway from *Arabidopsis* was systematically reconstructed in the yeast *Saccharomyces cerevisiae* using a synthetic

[1]Institute of Molecular Plant Biology, ETH Zurich, Universitätstrasse 2, 8092 Zurich, Switzerland. [2]Université de Lille, CNRS, UMR 8576–UGSF–Unité de Glycobiologie Structurale et Fonctionnelle, Lille, France. [3]Institute for Building Materials, ETH Zurich, Stefano-Franscini-Platz 3, 8093 Zurich, Switzerland.
*Corresponding author. Email: szeeman@ethz.ch
†Present address: John Innes Centre, Norwich Research Park, Colney Lane, Norwich NR47UH, UK.

biology approach (*13*). After purging the yeast of its glycogen metabolic genes, combinations of the plant amylopectin biosynthetic enzymes were expressed, together with a bacterial enzyme providing the ADP-Glc substrate. Depending on the enzyme complement, the resulting strains produced different branched glucan structures. This allowed not only biosynthetic factors to be characterized in isolation and at greater depth but also actors and features promoting or limiting glucan phase transition to be identified. For example, strains expressing all SS and BE isoforms produced branched but soluble glucans, while strains expressing the same enzymes but in addition ISA1 and ISA2 also accumulated substantial amounts of starch-like insoluble glucans, reinforcing the key role of ISA1 in promoting the formation of semi-crystalline granules. While the insoluble glucans produced by this strain had many starch-like properties, some structural differences to plant amylopectin remained and not all glucans made the phase transition. One explanation for this is that the higher-order structuring of amylopectin may require more than just the known set of the starch biosynthetic enzymes.

In this study, we investigated whether two widely conserved, nonenzymatic plant proteins, LIKE EARLY STARVATION 1 (LESV) and EARLY STARVATION 1 (ESV1), could be involved in promoting the phase transition of amylopectin. Both proteins are found associated with starch granules from different species (*14*, *15*). In addition, *Arabidopsis esv1* mutants were identified in a genetic screen for plants that exhaust their starch reserves too fast at night and so display signs of early nighttime carbon starvation (*14*). Starch produced by *esv1* appears to have a normal glucan structure but seems more susceptible to degradation, such that it is turned over too fast not only at night but also even during the day, during net starch accumulation. In contrast, ESV1 overexpression causes excess starch to accumulate. The second protein, LESV, shares strong sequence homology to ESV1 in its C-terminal region, which is very rich in aromatic and acidic amino acid residues. The *lesv* mutant does not have altered nighttime starch degradation, but LESV overexpression causes the formation of many small starch granules and alters their turnover. Thus, the two proteins appear to be functionally related, albeit with distinct roles.

Here, we analyze the recently available AlphaFold predictions of the ESV1 and LESV protein structures and validate these predictions using small angle x-ray scattering (SAXS) and circular dichroism (CD) spectroscopy. We show that parts of the two proteins form a previously undescribed carbohydrate interaction surface capable of glucan binding. Using both a synthetic biology approach in yeast and in vivo experiments in *Arabidopsis*, we provide direct evidence that LESV promotes the phase transition of branched, amylopectin-like glucans. We propose a model whereby the carbohydrate interaction surface of LESV serves as a template upon which to align amylopectin double helices and that this nucleates the crystallization of lamellae, while ESV1 serves to stabilize this structure, once formed.

RESULTS

The tryptophan-rich regions of ESV1 and LESV are strongly conserved

ESV1 and LESV were previously described to lack functional protein domains. However, the C-terminal ends of both proteins share a large region rich in aromatic amino acids, especially tryptophan (Trp), and the acidic amino acids aspartate and glutamate (*14*). To deepen our understanding of the proteins' key characteristics, we retrieved orthologous sequences and assessed their conservation—an approach we deemed promising considering the recent demonstration of functional conservation for distantly related ESV1 homologs (*16*). Separate alignments of orthologous ESV1 and LESV proteins suggest that, in both cases, the entire Trp-rich regions identified earlier (*14*) are more stringently conserved than the remaining sequences, although additional regions of local conservation are revealed for LESV-like proteins (Fig. 1A). Within the strongly conserved C-terminal regions, the numerous Trp residues themselves are the most prominent sequence feature (fig. S1, A and B).

The Trp-rich regions fold into unusual putative carbohydrate-binding surfaces

We next modeled the proteins' tertiary structures, making use of the recently available AlphaFold platform (*17*, *18*). The Trp-rich regions of both ESV1 and LESV are predicted with high confidence to adopt peculiar conformations, folding into extended, twisted planar sheets, about 70 Å in length and consisting of 16 antiparallel β strands (Fig. 1B). The sequences N-terminal (including the putative transit peptides) and C-terminal to these β sheets are modeled with lower confidence scores; for ESV1, most of these parts are likely disordered and are poorly conserved (Fig. 1, A and B), including the proline-rich region previously noted (*14*). For LESV, several α helices are confidently [predicted local distance difference test (pLDDT) >70] predicted both N- and C-terminally of the β sheet. Because these are presumably flexibly connected to the β sheet via linker sequences, their position may vary depending on the protein's state. While the N- and C-terminal parts of LESV are overall less conserved than the Trp-rich region, islands of conservation overlap well with some of the predicted α helices. Furthermore, three other conserved regions, one featuring a Trp pair, lie within the putatively unstructured protein parts (Fig. 1A and fig. S1B).

Given the high density of aromatic and acidic amino acids in the β sheets of ESV1 and LESV, we analyzed their spatial distribution. The aromatic residues aligned into two main strips spaced roughly 14 Å apart and running across both sides of the β sheet, perpendicular to the backbone β strands (Fig. 1, B and C). The acidic residues are chiefly located in the interstices between these aromatic strips, similarly organized in two strips also spaced 14 Å apart on both faces of the β sheet. Overlaying the ESV1 and LESV structures indicated that the distribution and orientation of aromatic and acidic residues are predicted to be extremely similar but not identical (fig. S2A). Some aromatic residues, specific to one or the other protein, are located either on the β strands at the level of the aromatic strips or on the loops connecting them.

To support the AlphaFold models experimentally, recombinant ESV1 and LESV proteins (minus transit peptides) were purified (fig. S2B) and analyzed for their secondary structure content by CD spectroscopy. Not only both proteins predominantly consist of β sheets, but LESV also had a certain fraction of α helices (fig. S2C), again consistent with the AlphaFold models. We next implemented SAXS to gather information on the three-dimensional (3D) shape of the proteins in solution. Both ESV1 and LESV showed smooth SAXS profiles, and analysis of the Guinier approximation plots confirmed that the proteins were nonaggregated (fig. S3, A and B, i to ii). Analyses of the protein particle sizes (table S1) and interatomic distance distributions showed that both ESV1 and LESV are composed of structured domains extended by elongated, more dynamic

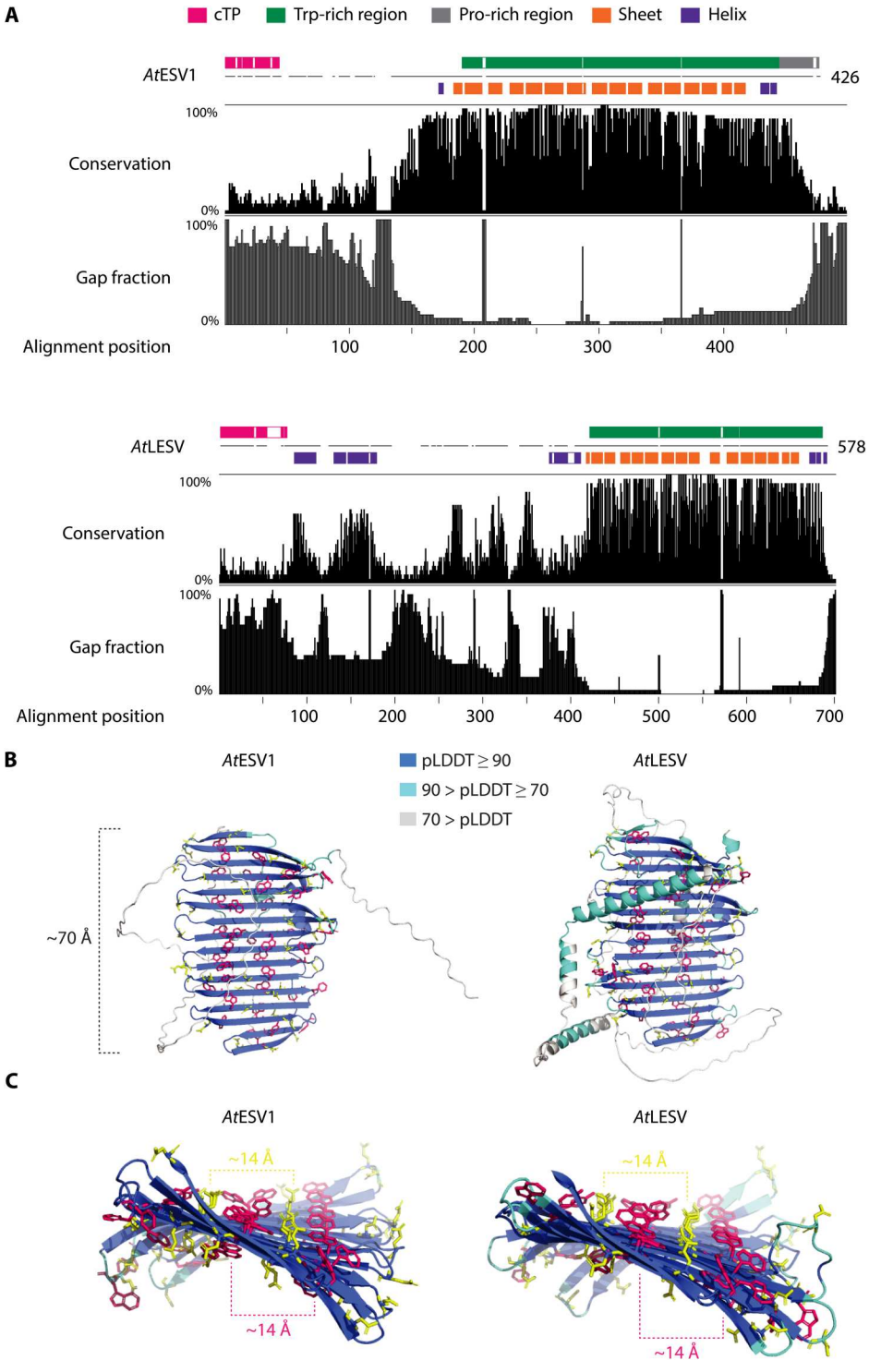


Fig. 1. Structures and conservation of ESV1 and LESV proteins. (**A**) Conservation of homologous ESV1 and LESV protein sequences. Gap frequencies are based on ≥25 orthologous ESV1 and LESV sequences from (*14*) (for alignments, see data S1, A and B). Indicated in colored boxes above the *Arabidopsis* sequences are the chloroplast transit peptide (cTP), Trp-rich region, and Pro-rich region (ESV1 only) of the respective proteins. Colored boxes below represent the AlphaFold secondary structure predictions [see (B)]. Note that the alignment position does not correspond to the *Arabidopsis* protein residue numbering because of gaps. (**B**) AlphaFold predictions of ESV1 and LESV. Cartoon representations of the full-length proteins; the side chains of aromatic (Trp, Tyr, and Phe) and acidic residues (Asp and Glu) are shown in pink and yellow, respectively, and the protein backbone is colored according to confidence of the model (pLDDT value; see color key). (**C**) Isolated β sheet plane regions of ESV1 and LESV, reoriented to reveal the view along the length of the planes.

domains. We calculated 10 ab initio molecular envelopes for both proteins using DAMMIF (*19*), fitting the experimental data with χ^2 of 1.7 (ESV1) and χ^2 of 2.3 (LESV), and averaged and filtered them using DAMAVER/DAMFILT (*20*). The results (fig. S3, A and B, iii) suggest that both proteins contain oblate domains, whose dimensions and shapes are consistent with the shape of the conserved β sheets of the AlphaFold models.

It was previously proposed that the Trp-rich domain may be responsible for the association of ESV1 and LESV with starch granules (*14*). The Trp-rich regions alone were sufficient to mediate starch granule localization when transiently expressed in *Nicotiana benthamiana* (fig. S4). Together, these data strongly suggest that the Trp-rich regions of ESV1 and LESV constitute atypical carbohydrate-binding surfaces.

ESV1 and LESV promote insoluble glucan accumulation in engineered yeast

To investigate the roles of ESV1 and LESV on starch biosynthesis in a simplified biological context, the proteins were expressed in *S. cerevisiae* cells previously engineered to synthesize starch-like glucans. Two yeast stains—designated lines 28 and 29—were used. Line 28 contains the *Arabidopsis* enzymes SS1 to SS4, BE2 and BE3, and accumulates large amounts of soluble glucan, while line 29 additionally contains the heteromultimeric isoamylase composed of ISA1 and ISA2, and a substantial fraction of its glucans are insoluble and semi-crystalline (*13*). Sequences encoding ESV1 and LESV (minus transit peptides), placed under the control of the *CWP2* promoter, were integrated by homologous recombination into the yeast's nuclear genome individually or both together. Expression of the two proteins in the different strains was confirmed by immunoblotting (Fig. 2A).

We assessed the quantities of glucans accumulated by the strains, measuring the soluble and insoluble pools separately (the latter defined as sedimentable by centrifugation—see Materials and Methods). Expression of ESV1 in line 28 background (designated line 559) resulted in the accumulation of small amounts of insoluble glucans in addition to the soluble glucans typical of the parental line (Fig. 2B). This effect was even more pronounced when expressing LESV (line 562) or both proteins simultaneously (line 563). Similar effects were obtained when ESV1 and LESV were expressed in line 29, which already produces insoluble glucans. In this case, the additional expression of ESV1, LESV, or both (lines 569, 572, and 573, respectively) substantially increased the amounts of insoluble glucans to the extent that they exceeded the soluble glucans. Again, LESV expression or the simultaneous expression of both proteins had the greatest effect. In both backgrounds, the increase in insoluble glucans brought about by ESV1/LESV expression was associated with a commensurate decrease in soluble glucans, suggesting that the two proteins may influence the partitioning of glucans between the soluble and insoluble fractions rather than altering the extent of their accumulation per se.

ESV1 and LESV influence insoluble glucan morphology in *S. cerevisiae*

Light microscopy (LM) was used to visualize the patterns of glucan accumulation in iodine-stained cells of the different yeast strains. As previously reported (*13*), cells of line 28 stained uniformly brown, as did ESV1-expressing line 559 (Fig. 2C). However, LESV-expressing lines (562 and 563) stained uneven, possibly reflecting the presence of both soluble and insoluble glucans. LM of line 29 revealed discrete, irregularly shaped particles—presumably the insoluble glucans. Intriguingly, in the ESV1-expressing line 569, the particles appeared larger and more regular in shape. This effect was even clearer in the LESV-expressing line 572 and most prominent in line 573, expressing both proteins. Together with the glucan fractionation and quantification, these data indicate that ESV1 and LESV promote the formation of insoluble glucans in our engineered yeast, acting irrespective of the presence of ISA.

We performed transmission electron microscopy (TEM) of fixed and embedded cells (Fig. 3A) and scanning electron microscopy (SEM) of the purified insoluble glucans (Fig. 3B). As previously reported (*13*), TEMs of line 28 revealed large numbers of small, uniform glycogen-like particles in the cytoplasm. Consistent with this, no larger particles resembling insoluble glucans could be purified for SEM. Upon expression of ESV1, LESV, or both proteins (lines 559, 562, and 563, respectively), TEM revealed irregularly shaped deposits of particulate matter (Fig. 3A, arrowheads) alongside glycogen-like material similar to that in the parental line. Using SEM, the insoluble material from these three lines appeared as small, aggregated particles (Fig. 3B). The insoluble glucans in line 29 were visible by TEM as large, irregular particles alongside glycogen-like material in the cytoplasm and, by SEM, as granular structures with rough surfaces (Fig. 3, A and B) (*13*). The expression of ESV1, LESV, or both proteins (lines 569, 572, and 573, respectively) in line 29 substantially altered the morphology of the particles; many now appeared as aggregates of discrete, larger particles with smooth surfaces. This was evident from both TEMs (where less glycogen-like material was seen) and from SEMs of the extracted glucans and consistent with our LM observations (Fig. 2C). The influence on particle morphology, in terms of size and surface uniformity, was more pronounced upon expression of LESV compared with ESV1 but most obvious when both proteins were expressed together.

Comparing the glucans' primary structure by their chain length distribution (CLD) profiles revealed that the primary glucan structures were similar to that of the parental lines (see fig. S5A and Supplementary Text). This is consistent with the absence of any known enzymatic function for ESV1 and LESV and suggests that their influence on glucan partitioning is not exerted by changing the glucan's chain lengths or branch point frequency. Rather, it suggests that glucans transition more readily into an insoluble state. In yeast line 29, short malto-oligosaccharides (MOSs) accumulate besides soluble and insoluble, starch-like glucans owing to isoamylase action, which liberates chains from the branched glucans made by the SSs and BEs (*9*, *21*). Expression of ESV1 and/or LESV in line 29 substantially reduced the accumulation of MOS (again, this effect was strongest upon expression of LESV compared with ESV1; fig. S6), indicating that, in the presence of these proteins, a greater proportion of the glucan transitions into an insoluble form, with a lower degree of debranching by ISA.

Aberrant starch formation occurs in the *lesv* mutant

Next, we investigated the roles of ESV1 and LESV *in planta*. *esv1* mutants were initially identified via their tendency to degrade starch too quickly, resulting in early nighttime starvation. However, *lesv* mutants had normal starch levels and turnover, with a mutant phenotype only evident upon LESV overexpression (*14*), which partly contrasts with the observed impact of LESV on

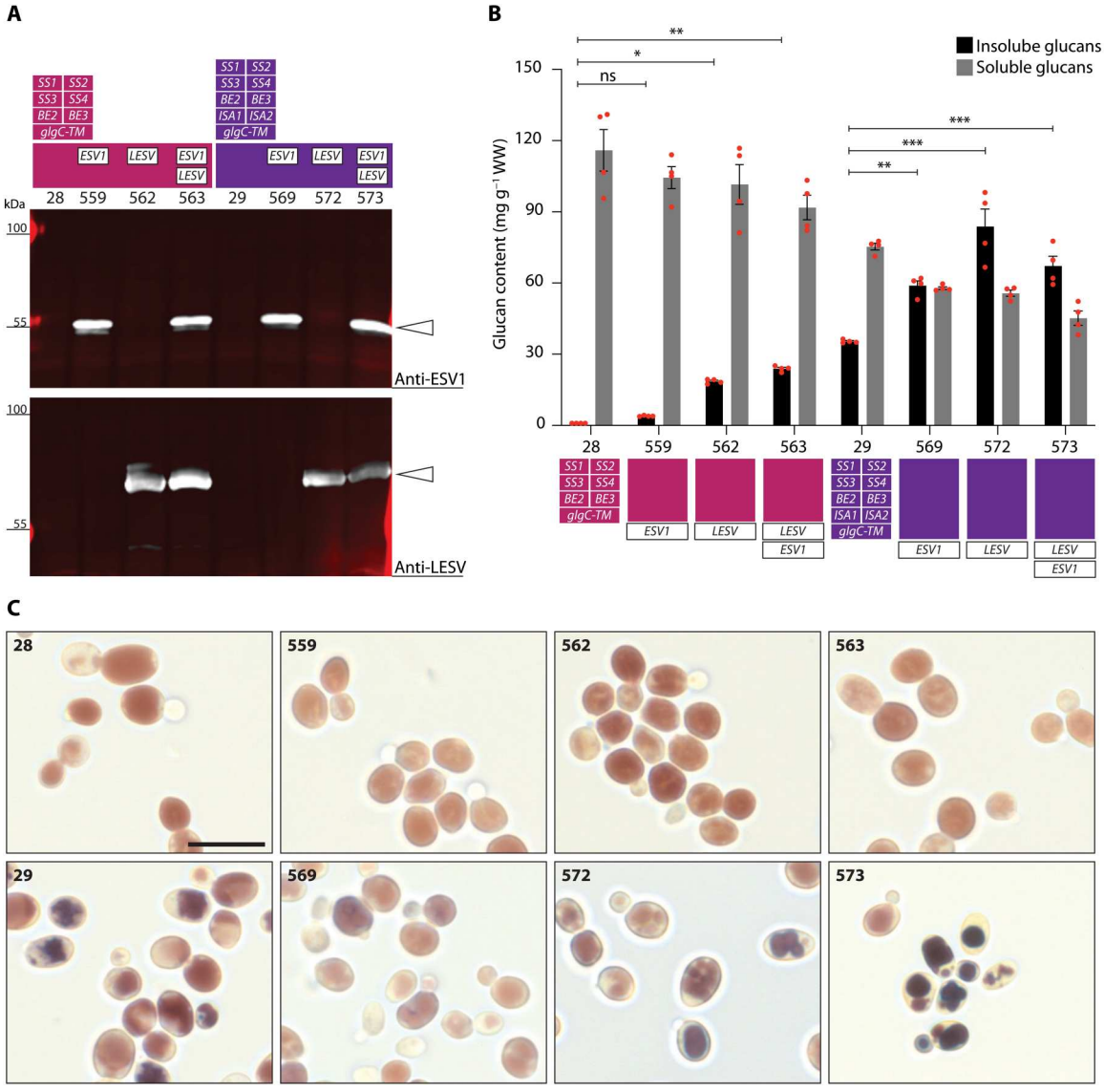


Fig. 2. Expression of ESV1 and LESV in different yeast genetic backgrounds. (**A**) Immunoblots of total protein extracts from yeast strains expressing ESV1 (45 kDa), LESV (72 kDa), or both proteins in the 28 and 29 genetic backgrounds (expression sets indicated above the strain number) (*13*). ESV1 and LESV were detected using protein-specific antibodies in each case (white arrowheads). (**B**) Quantification of insoluble and soluble glucans of the yeast strains shown in (A), grown in liquid culture for 5.75 hours under inducing conditions. Values are means ± SE ($n = 4$); independent replicate cultures arose from different precultures. WW, wet weight. Statistical comparisons were performed using two-way analysis of variance (ANOVA) with Dunnett's multiple comparisons test; see data S2 (A and B). Note that only comparisons for insoluble glucans to the respective parental strain are indicated in the graph. *$P \leq 0.05$, **$P \leq 0.01$, and ***$P \leq 0.001$. ns, not significant. (**C**) Light micrographs (LMs) of iodine-stained yeast cells grown as in (B). For strain specifications, refer to (A) and (B). Scale bar, 10 μm.

glucan partitioning in yeast. While glucan levels in the *lesv-1 Arabidopsis* T-DNA insertion line (hereafter referred to as *lesv*) appeared normal (Fig. 4A), and most *lesv* chloroplasts contained normal-looking starch granules, about 5% of the chloroplasts examined contained aberrant starch. Instead of normal lenticular granules, these plastids contained small granules with irregular surfaces surrounded by very small particles or just the small particles (Fig. 4B). These aberrant starch granules were not seen in the wild type nor in the *esv1-2* (hereafter referred to as *esv1*) T-DNA line.

We challenged the *lesv* mutant by subjecting it to a single prolonged night (16 hours in total) so as to completely destarch it, enforcing the re-creation of starch granules from scratch in the subsequent light period (*22*). After a subsequent 8-hour day, wild-type plants had produced near-normal levels of starch. However, *lesv* had a notable phenotype: It accumulated far less glucans in total, a substantial fraction of which was soluble (Fig. 4C). Visualization of these glucans by TEM showed that most *lesv* chloroplast sections contained a mix of aberrant granules and small, presumably soluble, particles (58%) or exclusively small particles (37%). Normal starch was observed in only 5% of

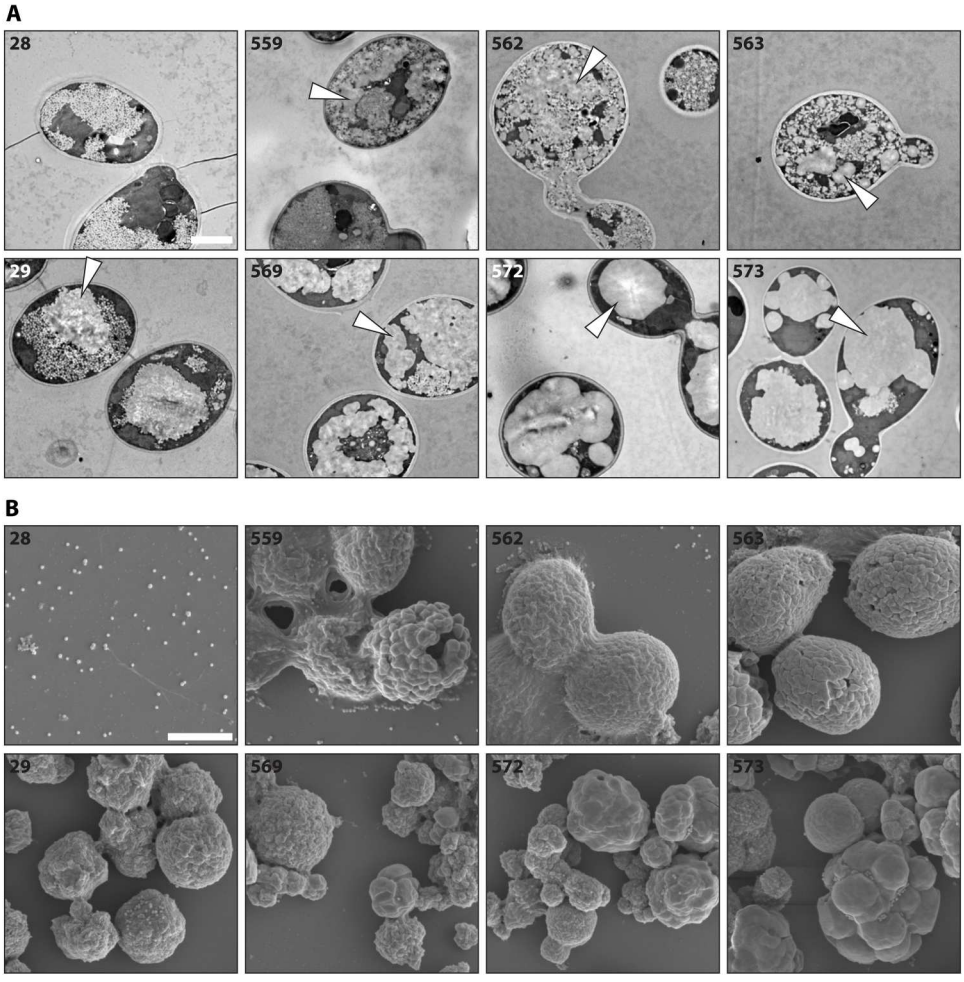


Fig. 3. Appearance of soluble and insoluble glucans produced by yeast strains expressing ESV1 and LESV. (**A**) TEMs of the indicated yeast strains (for strain specifications, refer to Fig. 2). White arrowheads indicate presumably insoluble glucan structures. (**B**) SEMs of insoluble particles purified from the indicated yeast strains. Scale bars, 2 μm (A) and (B).

chloroplast sections (Fig. 4D). The CLD profile of the starch was like wild-type starch, but that of the soluble glucans was enriched in very short chains (length < 6 glucose units; see fig. S5B and Supplementary Text). These short chains are indicative of amylolytic attack (*9*), and, indeed, the daytime level of maltose (the product of β-amylases) was greatly increased in *lesv* following the destarching treatment but not under normal conditions (Fig. 4, E and F). These data suggest that LESV strongly facilitates the process of de novo formation of insoluble starch granules. However, in regular day-night cycles, this deficiency is compensated for by other factors.

Overexpression of ESV1 or LESV promotes starch accumulation in *Arabidopsis* isoamylase mutants

To further reveal the impact of ESV1 and LESV on glucan crystallinity, we modulated their expression in the *isa1isa2* double mutant background. This isoamylase-deficient mutant predominantly makes soluble phytoglycogen in mesophyll cell chloroplasts but still produces some starch in epidermal and bundle sheath cells (*9*). We reasoned that this mutant background may represent an intermediate setting between our previously used yeast system and wild-type plants. Previously characterized *Arabidopsis* lines overexpressing yellow fluorescent protein (YFP)–tagged versions of ESV1 or LESV (*ESV1-OX*#3-2 and *LESV-OX*#4-6, hereafter referred to as *ESV1-OX* and *LESV-OX*, respectively) (*14*) were crossed to *isa1isa2*, and plants homozygous for the mutant and overexpression loci were identified. Immunoblotting confirmed the overexpression of ESV1-YFP and LESV-YFP in *isa1isa2ESV1-OX* and *isa1isa2LESV-OX*, respectively (Fig. 5A). For unknown reasons, ESV1-YFP was less abundant in *isa1isa2ESV1-OX* than in the *ESV1-OX* parental line. We reconfirmed granule association of these YFP-tagged protein versions by confocal microscopy (fig. S7) (*14*).

Next, these lines were characterized with respect to starch and phytoglycogen content at the end of the day (EOD). Consistent with previous reports, *ESV1-OX* had elevated starch contents and rosettes stained darker blue-black with I_2/KI solution, indicative of its high amylose content (Fig. 5, B and C) (*14*). *LESV-OX* plants had similar starch contents as the wild type but stained less intensely, also consistent with its previously reported low amylose level. TEM confirmed that *ESV1-OX* had larger granules than the wild type in all cell types examined, while *LESV-OX* accumulated numerous small starch granules that were variable in appearance

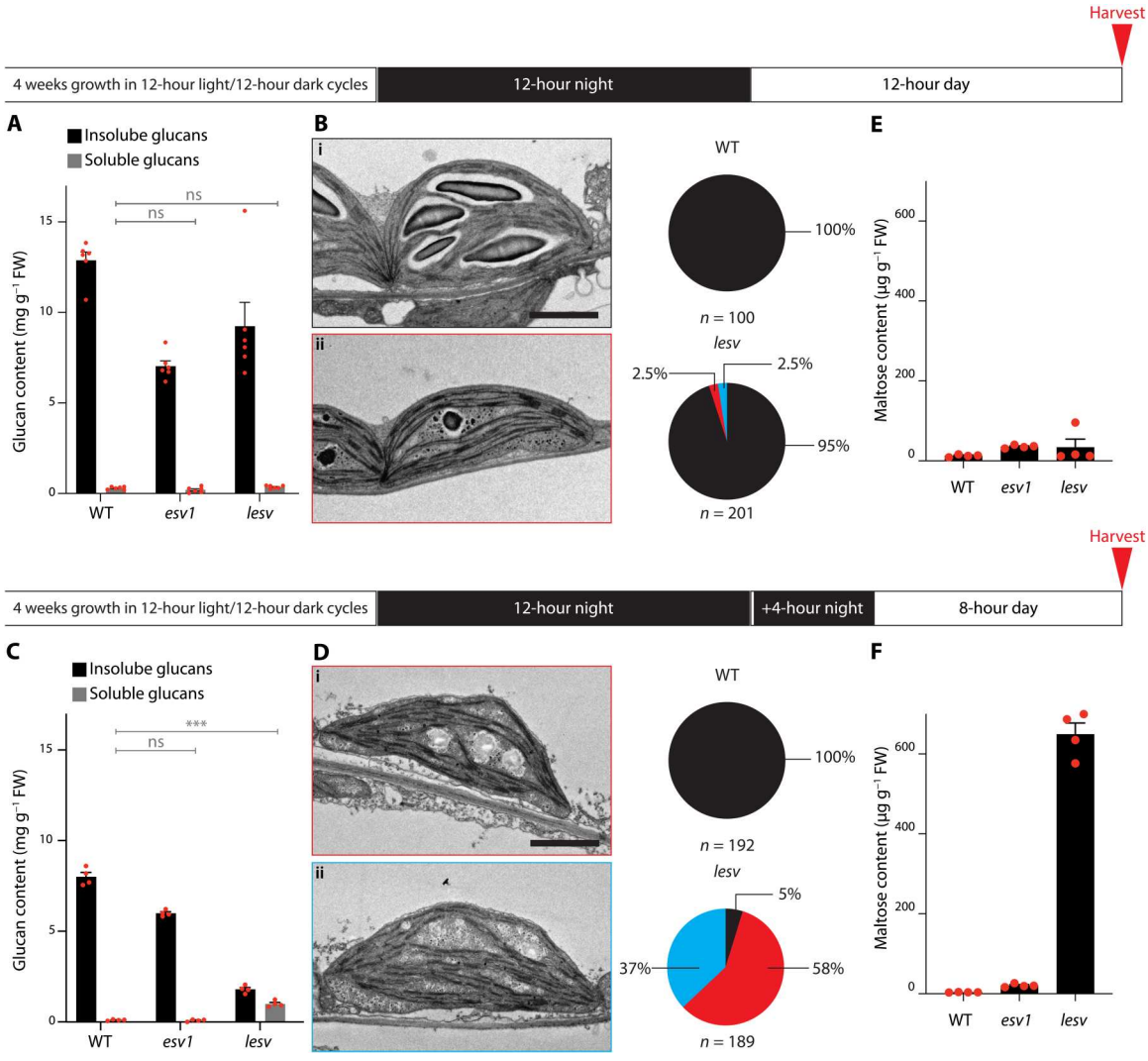


Fig. 4. The *lesv* single mutant has a conditional mutant phenotype. (**A**) Glucan quantification of plants grown in 12-hour light/12-hour dark cycles, harvested as indicated. FW, fresh weight. Values are means ± SE ($n = 6$). Statistical comparisons were performed using two-way ANOVA with Dunnett's multiple comparisons test; see data S3 (A and B). Only comparisons to the wild type for soluble glucans are shown. WT, wild type. (**B**) TEMs of selected *lesv* chloroplast sections, obtained from plants grown and sampled as in (A). Most *lesv* plastids appear wild-type like (i), but a few contain unusual glucans (ii). Scale bar, 2 µm. Pie charts show quantitative data of plastid section classifications; black, sections containing regular starch granules; red, sections containing starch and apparent phytoglycogen; blue, sections containing phytoglycogen-like inclusions only. (**C**) Glucan quantification of plants grown as in (A), subjected to a single prolonged night (16 hours), and harvested as indicated. Values are means ± SE ($n = 4$). Statistical comparisons were performed as in (A); see data S4 (A and B). Only comparisons to the wild type for soluble glucans are shown. $^{***}P \leq 0.001$. (**D**) TEMs of *lesv* plants grown and harvested as in (C). Most chloroplast sections contain either a mixture of starch and phytoglycogen (i) or phytoglycogen-like inclusions only (ii). Scale bar, 2 µm. Pie charts indicate respective quantifications, as in (B). (**E** and **F**) Quantification of maltose in the plants in (A) (E) and (C) (F). Note that maltose is not included in the soluble glucans as measured in (A) and (C). Values are means ± SE ($n = 4$). Statistical comparisons were performed using one-way ANOVA with Dunnett's multiple comparisons test; see data S5 (A and B).

(Fig. 6) (*14*). The *isa1isa2* double mutant, with predominantly soluble phytoglycogen, stained orange-brown, and TEM analysis supported its previously reported cell-type–specific phenotype (Fig. 6 and fig. S8A) (*9*).

Both *isa1isa2ESV1-OX* and *isa1isa2LESV-OX* lines stained a darker color than *isa1isa2* at EOD, but their phenotypes differed (Fig. 5B); *isa1isa2ESV1-OX* not only had similar amounts of phytoglycogen as *isa1isa2* but also had starch levels comparable to the wild type. However, while *isa1isa2LESV-OX* also had wild-type levels of starch, it had very little phytoglycogen. Thus, in both lines, the starch fraction predominated over the phytoglycogen (Fig. 5C). TEMs of leaf tissue yielded results consistent with these quantitative analyses; *isa1isa2ESV1-OX* had both phytoglycogen and large starch granules in its mesophyll cells (Fig. 6) but only starch granules in epidermal and bundle sheath cells (fig. S8A). In contrast, *isa1isa2LESV-OX* contained many small starch granules in the plastids of all cell types examined but very little phytoglycogen (Fig. 6 and fig. S8A). These phenotypes were consistent over the leaf sections, as determined by LM (fig. S8B).

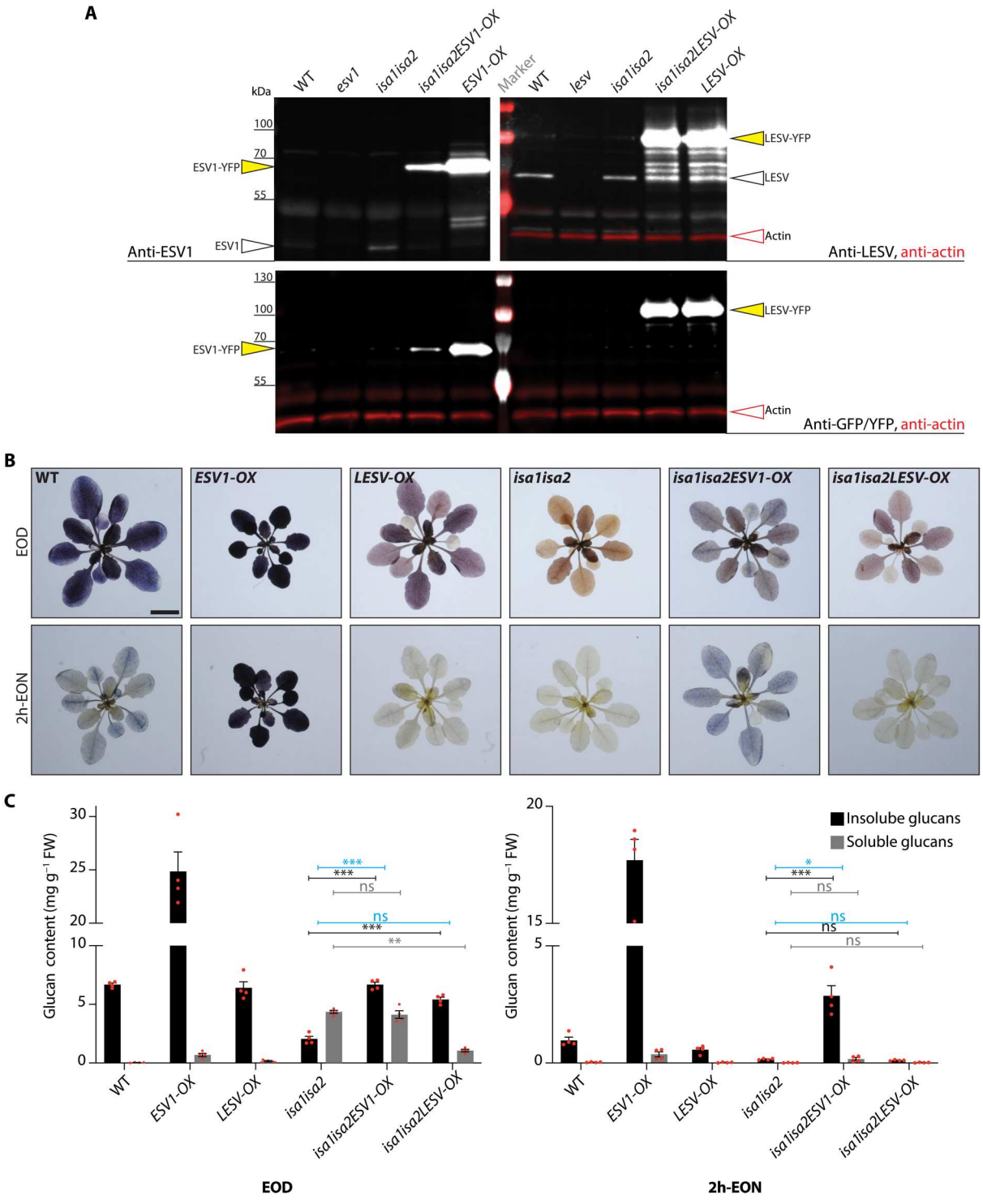


Fig. 5. Glucan accumulation and turnover in plants overexpressing ESV1 and LESV in the *isa1isa2* background. (**A**) Endogenous ESV1 and LESV and overexpressed ESV1-YFP and LESV-YFP, as assessed by immunoblotting of total leaf protein extracts. Actin (in red) served as a loading control (omitted in the anti-ESV1 immunoblot for clarity). (**B**) *Arabidopsis* rosettes harvested at the EOD and 2 h-EON stained with Lugol's solution. Scale bar, 1 cm. (**C**) Glucan quantification of plants harvested at the EOD and 2h-EON. Values are means ± SE ($n = 4$ biological replicates). Statistical comparisons were performed using two-way ANOVAs with Dunnett's multiple comparisons test. Comparisons of total summed glucans are indicated in blue, those of soluble glucans (phytoglycogen) in gray, and those of insoluble glucans (starch) in black. $^{*}P \leq 0.05$, $^{**}P \leq 0.01$, and $^{***}P \leq 0.001$. For clarity, only selected comparisons are shown. See data S6 (A and B).

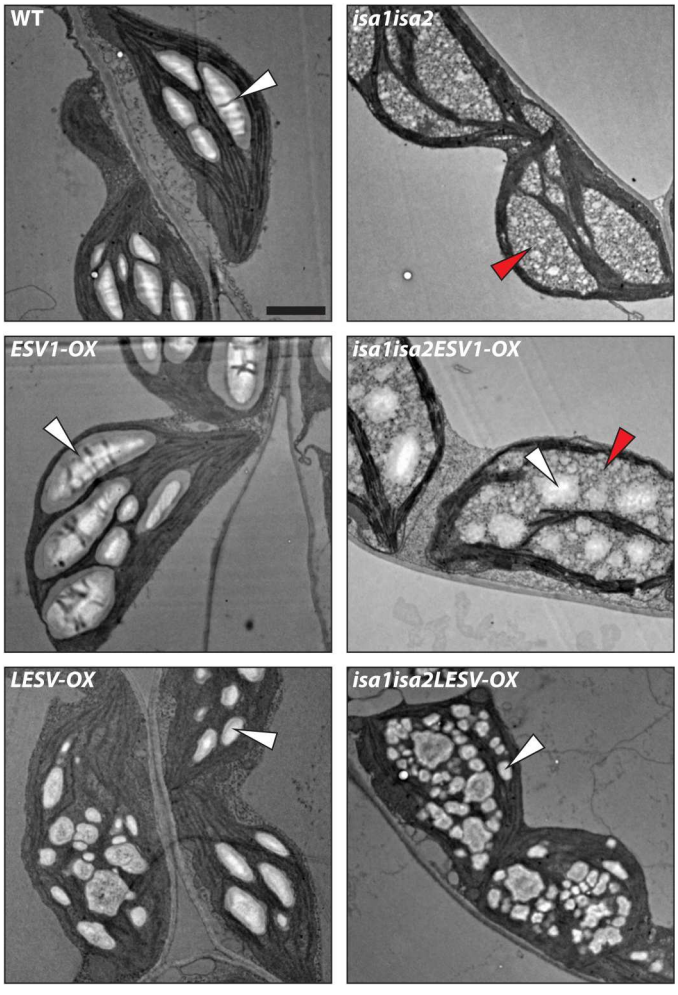


Fig. 6. Appearance of glucans forming in mesophyll chloroplasts of plants overexpressing ESV1 or LESV in the *isa1isa2* background. Leaf tissue was harvested at the EOD, and the chloroplasts and starch granules therein visualized using TEM. White and red arrowheads indicate starch granules and phytoglycogen, respectively. Scale bar, 2 μm.

Similar I_2/KI staining and quantitative analyses were carried out 2 hours before the end of the night (2h-EON), when *esv1* mutant plants have already prematurely degraded their starch (*14*). As expected, wild-type plants still contained some starch and rosettes stained lightly (Fig. 5, B and C). *ESV1-OX* rosettes stained much stronger and had excess starch, consistent with a reduced rate of starch mobilization, while *LESV-OX* did not. In *isa1isa2*, both starch and phytoglycogen had been consumed by this time. *isa1isa2ESV1-OX* had degraded its phytoglycogen but retained much of its starch. In contrast, *isa1isa2LESV-OX* had degraded its starch and what little phytoglycogen it had. These data suggest that ESV1 seems to limit the degradation of starch, while having little impact on the amount pf phytoglycogen that accumulates. LESV appears to promote the formation starch instead of phytoglycogen, without limiting degradation.

Because both overexpression lines were made using YFP-tagged proteins, we used confocal microscopy to examine the proteins′ distribution. *ESV1-OX* and *LESV-OX* lines yielded YFP fluorescent patterns consistent with starch granule localization (fig. S7) (*14*). In both *isa1isa2ESV1-OX* and *isa1isa2LESV-OX*, the fluorescence was primarily in the form of smaller, grouped punctae with a more diffuse background in areas between the chlorophyll autofluorescence, where glucans accumulate.

Glucan turnover and crystallinity are affected by loss of ESV1 or LESV in *isa1isa2*

Next, we investigated the consequence of losing ESV1 or LESV function on the *isa1isa2* phenotype by generating the *isa1isa2esv1* and *isa1isa2lesv* triple mutants. At the EOD, the *esv1* and *lesv* mutants stained similarly to the wild type with I_2/KI. However, *esv1* contained less starch than the other two lines (Fig. 7, A and B) (*14*). As before, *isa1isa2* accumulated predominantly phytoglycogen and less starch. Both the *isa1isa2esv1* and *isa1isa2lesv* triple mutants accumulated less glucan compared with their respective parental lines, with the reduction being particularly pronounced in *isa1isa2esv1*. However, the ratio of phytoglycogen:starch differed between the lines; in *isa1isa2lesv*, it was higher than in *isa1isa2*, while in *isa1isa2esv1*, it was lower (Fig. 7B). At 2h-EON, the wild type and *lesv* mutant still stained with I_2/KI and contained a little starch, but *esv1* did not stain and contained almost none. The rosettes of *isa1isa2*, *isa1isa2esv1*, and *isa1isa2lesv* were also devoid of glucans at this time (Fig. 7, A and B), indicating that these lines had prematurely used their glucan reserves.

To see how the cell-type–specific starch and phytoglycogen phenotype of *isa1isa2* was affected in the triple mutants, we investigated leaf mesophyll, epidermal, and bundle sheath cell plastids by TEM (Fig. 8 and fig. S9). As expected, mesophyll chloroplasts of *isa1isa2* contained phytoglycogen with few starch-like particles, while epidermal and bundle sheath cell plastids contained only starch granules. In *isa1isa2esv1*, the mesophyll cells contained less glucan, but a mix of starch and phytoglycogen was still evident; however, in epidermal cells, neither starch nor phytoglycogen was observed, consistent with the overall very low glucan content of this line (Fig. 7B). A contrasting result was obtained with *isa1isa2lesv*: The plastids in the mesophyll, epidermal, and bundle sheath cells all contained only phytoglycogen (Fig. 8 and fig. S9).

The accumulation of phytoglycogen, as opposed to starch, in *isa1isa2* is ultimately due to the aberrant structure of the glucan produced, yet overexpression or mutation of ESV1 or LESV shifted the starch:phytoglycogen ratio in a similar overall way as in engineered yeast cells (Figs. 2 to 6). We again tested whether this was due to a modified structure of the glucan by obtaining the CLDs for the glucans from each of the genetic backgrounds. As seen for the yeast, the absence or overabundance of ESV1 or LESV had only minor impacts on the glucan′s primary structure despite causing major changes in their turnover and/or partitioning between the soluble and insoluble phases (see fig. S5B and Supplementary Text).

Last, we used small-angle x-ray scattering (SAXS) to investigate whether the presence of ESV1 or LESV influenced the lamellar structure of insoluble glucans. The starch-like glucans synthesized in yeast were shown previously to be enriched in long chains and to have a relatively weak lamellar repeat of over 13 nm compared to *Arabidopsis* amylopectin with its ca. 10-nm repeat (*13*). These observations were replicated here, with insoluble glucans from yeast line 29 giving a SAXS pattern consistent with a 13.1-nm repeat. The expression of ESV1, LESV, or both proteins in line 29 resulted in an altered SAXS pattern indicating a shorter repeat of 10.6 to 10.8

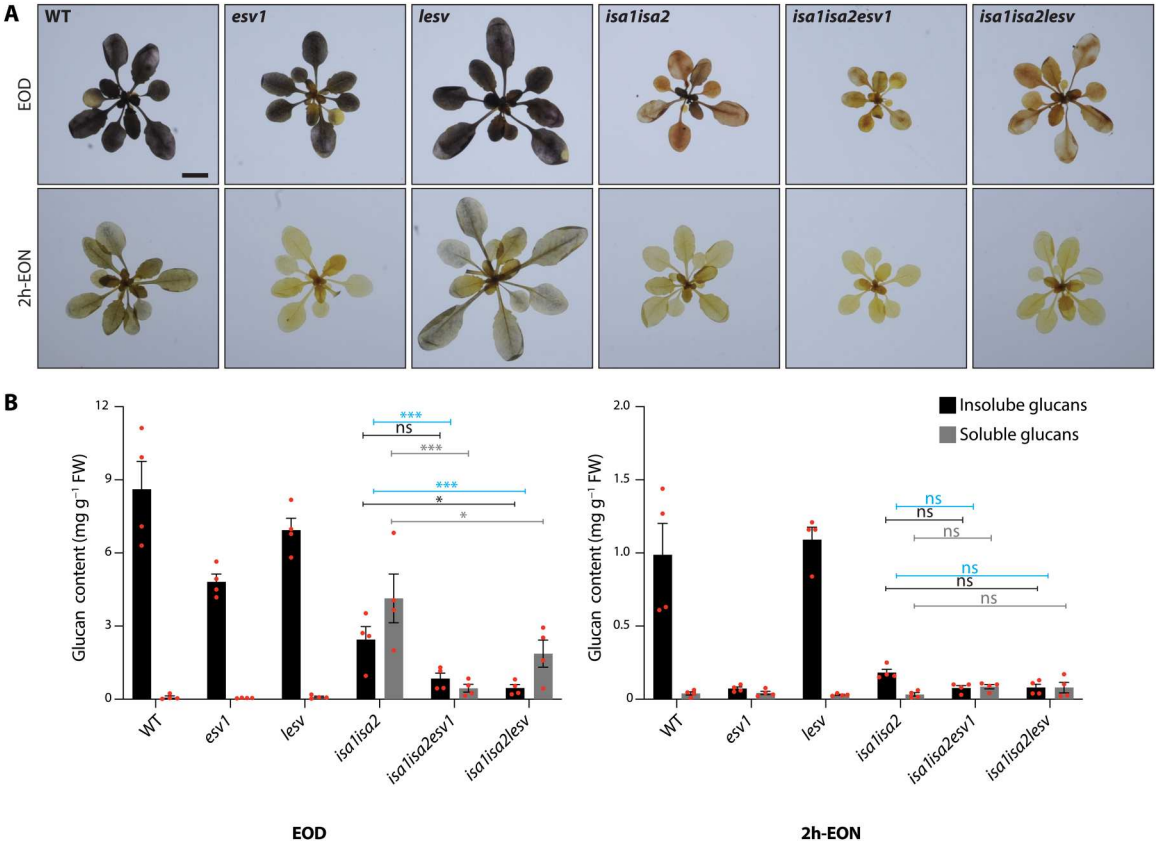


Fig. 7. Glucan content of *isa1isa2*, *esv1*, *lesv*, and higher-order mutants thereof. (**A**) Entire rosettes were harvested at the EOD and 2h-EON and stained with Lugol's solution. Scale bar, 1 cm. (**B**) Glucan quantification in plants, harvested at the EOD and 2h-EON. Values are means ± SE ($n = 4$ biological replicates). Statistical comparisons were performed using two-way ANOVA with Dunnett's multiple comparisons test. Comparisons of total summed glucan contents are indicated in blue, soluble glucans (phytoglycogen) in gray, and insoluble glucans (starch) in black. For clarity, only selected comparisons are shown. *$P \leq 0.05$ and ***$P \leq 0.001$. See data S7 (A and B).

nm (Fig. 9A and table S2). Similarly, the insoluble glucans produced in line 28 expressing ESV1, LESV, or both proteins had short repeats (9.7 to 10.6 nm). Similar SAXS analysis applied to the starches extracted from plants showed that the loss or overexpression of either ESV1 or LESV did not change the lamellar repeat substantially, whether in the wild type or *isa1isa2* background: In all cases, the SAXS patterns were consistent with lamellar repeats in the region of 10 nm (Fig. 9B).

DISCUSSION

The insight obtained from this work, using structural modeling and a set of complementary experimental approaches, provides a strong case for protein-mediated phase transition of amylopectin during starch biosynthesis. This is a paradigm-changing idea, since it has previously been assumed that phase transition is a biophysical process, occurring spontaneously when the amylopectin biosynthetic enzymes (SSs, branching, and debranching enzymes) generate a crystallization-competent structure. We suggest that, while these spontaneous crystallization can occur, it is promoted by the action of LESV. Furthermore, once a crystalline state is attained, we suggest that this state is stabilized by ESV1.

ESV1 and LESV have unique starch binding domains

The ESV1 and LESV proteins were originally identified as starch-binding proteins (*14*). This property was proposed to be due to the unusual Trp-rich C-terminal domain that both proteins have, which we confirmed experimentally (fig. S4). Surface-exposed Trp residues, as well as the other aromatic amino acids (phenylalanine and tyrosine), are well known to form interaction sites for glucans, which bind via nonpolar CH-π staking interactions (*23*). The high-confidence structural models for the C-terminal domains of both ESV1 and LESV (Fig. 1), supported by our biophysical analysis of the recombinant proteins (fig. S3), are suggestive of a unique and remarkable protein-glucan interaction surface. The predictions for these domains are almost identical for both *Arabidopsis* proteins (fig. S2) and for several orthologs tested , consistent with the high degree of conservation between species (Fig. 1).

The regular spacing of the side chains of ca. 30 of the 40 aromatic amino acids results in two strips adorning both sides of the extended antiparallel β sheet. Thus, extensive interactions with several glucan chains could occur. The β sheet domain has a length of approximately 7 nm, and the spacing of the aromatic amino acid strips is around 1.4 nm (Fig. 1, B and C). These dimensions are highly conspicuous, being similar to the predicted lengths and packing distances, respectively, of the double helices within the crystalline lamellae (*3*). We therefore propose that the strips of aromatic amino

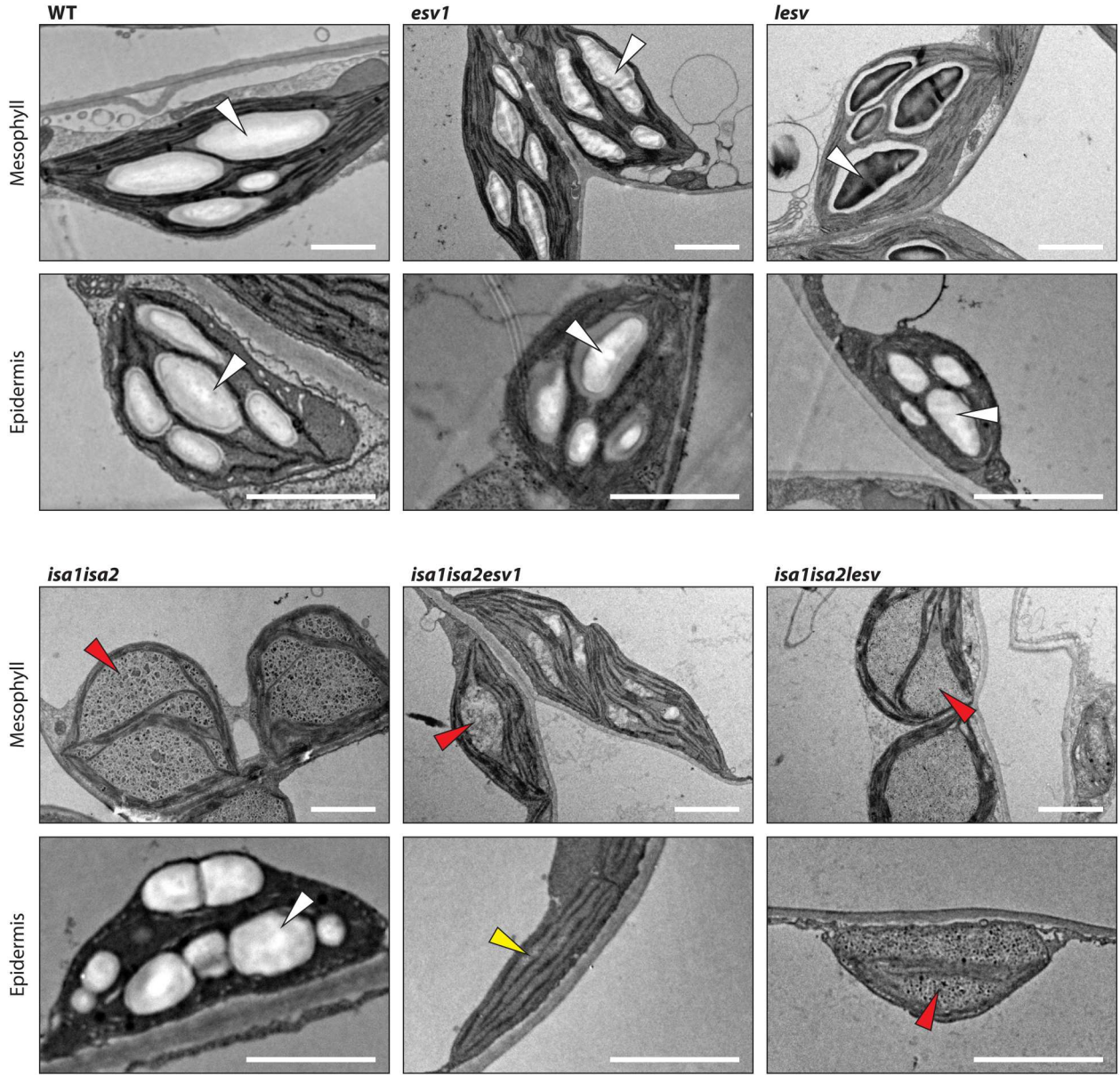


Fig. 8. Glucans in mesophyll and epidermal cell plastids of *isa1isa2*, *esv1*, *lesv*, and higher-order mutants. Leaf tissue was harvested at the EOD, and plastids and starch granules observed using TEM. Arrowheads indicate starch granules (white), phytoglycogen (red), and plastid sections lacking glucans entirely (yellow). Scale bars, 2 μm.

acids may coordinate the binding of neighboring double helices within an amylopectin molecule (Fig. 10). The frequent surface-exposed acidic residues (aspartates and glutamates) located between the aromatic residues could also participate in glucan binding by hydrogen bonding with the hydroxyl groups of the glucosyl residues (*24*).

Despite the similarity of their C termini, ESV1 and LESV differ substantially at their N termini, the structural predictions for which are less confident. No ordered structure is predicted by AlphaFold for the N terminus of ESV1, and it is notable that this part of the protein is poorly conserved between species. In contrast, parts of the N terminus of LESV are predicted to fold into α helices, consistent with our CD spectroscopy of the recombinant protein (Fig. 1 and fig. S2). Furthermore, there is overlap between this prediction and areas of high sequence conservation between species. That said, AlphaFold failed to predict ordered structure for a substantial part of the N terminus, and the predicted aligned errors for most of the N-terminal residues are high. Therefore, it seems likely that the exact organization of the helices, relative to each other and to the C-terminal domain, may differ to that shown in Fig. 1B. Two areas of conservation within the N terminus also feature prominent aromatic amino acids (fig. S1B), suggesting that it might also participate in glucan binding. Further studies will be needed to determine the structure of the N terminus of LESV, assess its binding capabilities, and see how these interrelate.

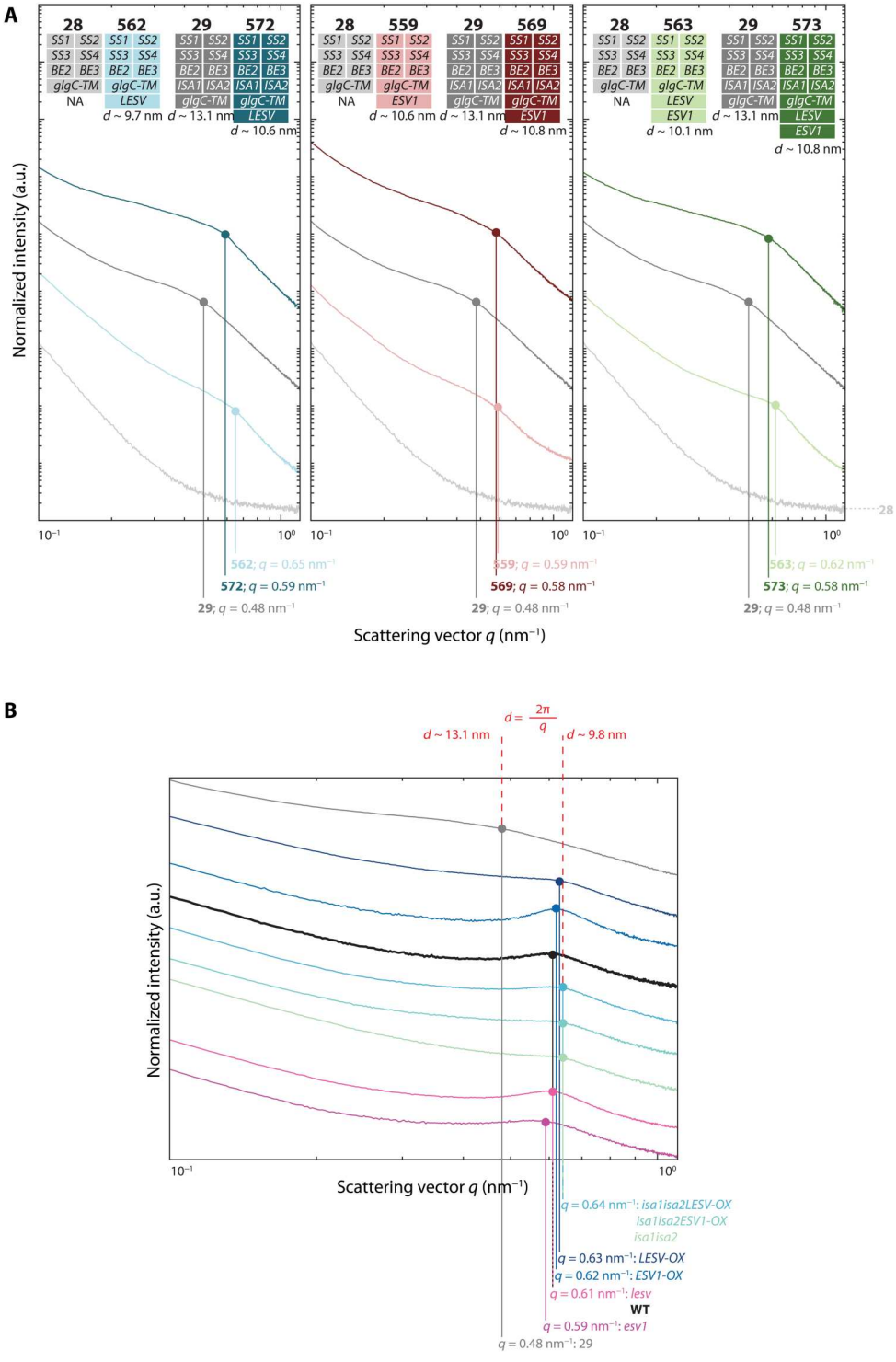


Fig. 9. SAXS of purified yeast and plant glucans. (**A**) Stacked SAXS plots obtained from glucans purified from yeast strains expressing ESV1 and LESV. The local intensity maxima of individual samples were manually selected and are highlighted by colored dots. Vertical lines indicate the respective maxima's *q* values. The respective calculated repeat distances (*d*) are indicated below the strains' genetic descriptions. As expected, no maximum could be detected, and thus, no *d* was calculated, for yeast strain 28, due to the absence of insoluble glucans. NA, not applicable. (**B**) Stacked SAXS plots obtained from insoluble glucans purified from different plant genotypes. Displayed data are as in (A). In both (A) and (B), only one replicate measurement is shown; refer to table S2 for a summary of replicate analyses. a.u., arbitrary unit.

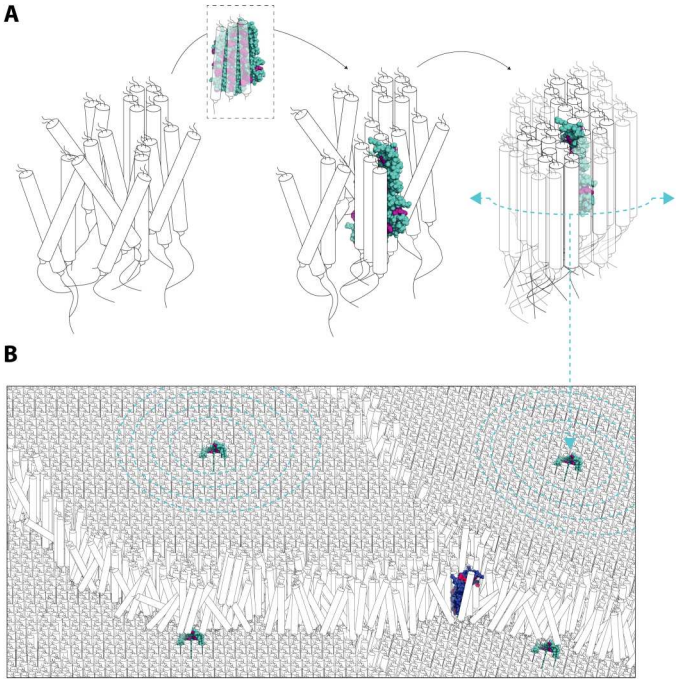


Fig. 10. Model of the proposed ESV1 and LESV functions. (**A**) LESV interacts with glucans that have adopted a helical secondary structure via its Trp-rich domain and possibly aided by its N-terminal domain (not shown) and thereby facilitates their arrangement into compact, ordered tertiary structures. Once seeded, regular glucan arrangements can self-propagate and spread. (**B**) Regions of ordered glucan helices, seeded by LESV, form adjacent blocks of crystalline lamellae. ESV1 binds to and stabilizes exposed helices at the margins of these regions, restricting access to hydrolytic activities. In both (A) and (B), ESV1 and LESV proteins are represented as simplified versions of their predicted AlphaFold structures (only Trp-rich region is shown; aromatic residues are highlighted by color; pLDDT values are disregarded). See Fig. 1B as comparison.

What is also clear, both from previous studies and from data presented here, is that the functions of the two proteins differ. The phenotypic effects of mutations abolishing expression of each gene, or overexpressing of each, are distinct (*14*). Furthermore, while aspects of the *esv1* mutant phenotype could be complemented by expression of the *Marchantia polymorpha* and *Oryza sativa* ESV1 orthologs, they could not by ectopic overexpression of LESV (*16*).

A role for LESV promoting amylopectin phase transition

Collectively, our data suggest that LESV fulfills a previously unrecognized role in starch biosynthesis by promoting the phase transition of glucans. The evidence comes from several directions. First, in yeast cells engineered to make starch-like glucans, LESV is able to increase the fraction that transitions into an insoluble state, over that which crystalizes spontaneously. The resultant granules were larger and more uniform in appearance. The ability of LESV to promote phase transition was even seen in yeast cells where the complement of starch biosynthetic enzymes was incomplete, i.e., in line 28 where isoamylase is missing and where the resultant polymer otherwise remained exclusively in the soluble phase (Figs. 2 and 3). Second, this influence of LESV was also seen in isoamylase-deficient plants, where its overexpression substantially decreased the amount of soluble phytoglycogen and increased the amount of starch (Figs. 5 and 6). Consistent with this, the residual starch made in the *isa1isa2* double mutant was largely absent in the *isa1isa2lesv* triple mutant, suggesting that the endogenous LESV already promotes starch formation despite the suboptimal amylopectin branching pattern (Fig. 7). Third, starch formed after an extended night is highly abnormal in the *lesv* single mutant, with a substantial proportion remaining soluble as phytoglycogen despite the full complement of starch biosynthetic enzymes (Fig. 4). This result contrasts with the relatively mild phenotype of the *lesv* mutant that was reported previously (*14*) and was revealed only after destarching the plants.

Clearly, starch granules can form in the absence of LESV, but our results suggest that LESV promotes this by seeding glucan phase transition. We propose that the Trp-rich domain serves as a template to align double helices that form spontaneously between neighboring chains of amylopectin (Fig. 10A). We envisage that each strip of aromatic amino acids on each side of the β sheet domain could bind one double helix, and, once properly aligned on LESV, these double helices could then align others, propagating a wave of phase transition, resulting in a crystalline lamella. Given that LESV is found encapsulated inside starch granules—something also true for LESV orthologs from different species (*14*, *15*)—we assume that at least a fraction of it then remains associated with the crystalline lamellae. The seeding function may also require other parts of the LESV protein, including the conserved α-helical domains, or the regions with additional conserved aromatic residues but for which no structure is predicted.

This model could explain many of our observations. On one hand, when the glucan branching pattern is suboptimal (e.g., in the absence of the trimming isoamylase), the chance that double helices spontaneously self-align to seed a crystalline structure will be reduced, but the presence of a LESV template could allow it to occur. On the other hand, when the branching pattern is optimal, crystallization may occur spontaneously and self-propagate even in the absence of a LESV template. This is important, because it could explain the near-normal starch phenotype of the *lesv* mutant when growing under standard diel conditions (Fig. 4, A and B). During a normal night, most but not all starch is degraded (*2*). Any residual starch could itself serve as the template for the crystallization of freshly synthesized amylopectin the following day. Only when completely destarched, therefore, does the absence of both LESV and residual starch result in highly aberrant glucan formation (Fig. 4, C and D). It is also notable that, upon LESV overexpression, the number of starch granules is greatly increased. This could also be interpreted as accelerated crystallization, such that freshly synthesized soluble glucans assemble into more numerous distinct granules, rather than being added to the growing surface of existing granules.

Our SAXS analysis of the insoluble glucans adds a further angle of support for our proposed function for LESV (Fig. 9). The full suite of starch synthesizing enzymes in the yeast line 29 creates a crystallization-competent glucan, which spontaneously adopts a tertiary structure with a 13.1-nm repeat. However, the presence of LESV not only increased the amount that crystallizes but also shortened the lamellar repeat close to the 10-nm values seen for *Arabidopsis* starch. The 9- to 10-nm repeat in plant starch are formed by a crystalline (~6 nm) and amorphous (~3 nm) layer (*3*, *4*). The 7-nm length of the Trp-rich β sheet of LESV could thus guide the formation of double helices in this range, such that the next amorphous

layer can be initiated on top of it. It would be interesting to see if engineering LESV to alter the length of its β sheet could influence the lamellar repeat length upon overexpression in yeast or in plants.

A role for ESV1 in stabilizing semi-crystalline amylopectin

We propose a slightly different role for ESV1, which is to stabilize the amylopectin within the starch granule once it has formed, rather than to promote its crystallization in the first place. The evidence for this again comes from several directions. First, in yeast cells, ESV1 has a much smaller effect on promoting glucan phase transition compared with LESV (Fig. 2). Thus, while the Trp-rich domain of ESV1 seemingly can trigger phase transition, it is either less efficient than the equivalent domain of LESV or other parts of the LESV protein that ESV1 lacks also assist in the phase transition process (as discussed above). Second, when overexpressed in isoamylase-deficient plants, ESV1 increased the amount of starch but did not do so by decreasing the amount of soluble phytoglycogen (Fig. 5). Rather, the increase in starch is probably due to a block in its degradation at night; while the phytoglycogen is degraded, excess starch remains at dawn, leading to a higher overall glucan content. This excess starch accumulation is also seen when ESV1 is overexpressed in the wild type (*14*). Third, in the *esv1* mutant, starch is degraded too fast during the night and even concurrently with its accumulation during the day (*14*). However, no evidence for aberrantly formed granules or for soluble glucan accumulation comparable to the destarched *lesv* mutant was seen (Figs. 4, 7, and 8).

Given their semi-crystalline nature, starch granules are considered inherently resistant to enzymatic degradation. This view has been reinforced by the discovery that transient glucan phosphorylation is critical for normal degradation, with the reasoning that it disrupts the semi-crystalline structure (*1*, *25*–*29*). However, it is unlikely that the starch granule surface is uniform, and there may be areas where the structure is susceptible to degradation, regardless of phosphorylation. For example, at locations where one crystalline lamella ends and the next begins with a change in orientation, the double helices may be less stable. We envision that ESV1 could play a role at such sites, recognizing and stabilizing these weak points. Thus, in the absence of ESV1, these sites would be exposed, explaining the uncontrolled degradation in the *esv1* mutant. Conversely, ESV1 overexpression could render the granule more stable and even interfere with the process of phosphorylation, as proposed from in vitro work with recombinant ESV1 (*30*). Given this proposed role, we speculate that ESV1 may be primarily associated with the granule surface rather than becoming encapsulated like LESV. This is consistent with fewer ESV1 peptides being present inside starch granules compared to LESV (*14*). However, further experimental work will be required to test this hypothesis.

We suggest that soluble glucans, i.e., phytoglycogen, are less effectively protected by ESV1 than starch; all the phytoglycogen in *isa1isa2* is degraded at night, even when ESV1 is overexpressed (Fig. 5C). That said, phytoglycogen levels were very low in the *isa1isa2esv1* triple mutant (Fig. 7B), suggesting that in *isa1isa2*, ESV1 limits phytoglycogen turnover to some extent.

In conclusion, our work points toward a previously unrecognized biochemical process at play during starch biosynthesis. The amylopectin biosynthetic enzymes, working together on a common substrate, generate a glucan product with a propensity to crystallize. We propose that crystallization is facilitated by LESV and subsequently stabilized by ESV1. Clearly, this model needs further rigorous testing. The individual roles of each protein need to be defined, as do their potential combined effects when both are present and interacting with nascent glucans during starch biosynthesis or the mature granules themselves. The advantage of protein-mediated phase transition also needs to be established; it may help plants to make starch robustly in a wide variety of tissues and under different conditions. Variable factors such as temperature and plastid stroma solute content, i.e., during water stress, could differentially influence the starch biosynthetic enzymes, leading to alterations in amylopectin structure. Such an effect was reported for the starch in barley seeds developing at different temperatures (*31*). Furthermore, these variable factors may act as control parameters influencing the likelihood of spontaneous crystallization of the glucan. Thus, the presence of proteins that provide a template to nucleate crystallization and stabilize it could ensure efficient starch granule production and carbon storage even in suboptimal conditions. Last, the widespread conservation of both *ESV1* and *LESV* genes and the presence of the proteins in starches obtained from multiple sources (*14*) suggest that these results from *Arabidopsis* will be more broadly applicable in plants, including many of the world's most important staple crops.

MATERIALS AND METHODS

Experimental design

The goal of this study was to better understand the molecular roles of the *Arabidopsis thaliana* ESV1 and LESV proteins in starch metabolism. This was approached from three angles: first, by studying the proteins' sequences and 3D structures using bioinformatic methods and validating in vitro techniques using recombinantly expressed purified proteins; second, by studying the influence of ESV1 and LESV expression on glucan production and properties in an engineered yeast system; and, last, by investigating the effects of ESV1 and LESV overexpression or absence *in planta* using the *isa1isa2* double mutant as a sensitive genetic background.

Bioinformatic analyses

For conservation analyses, orthologous streptophyte ESV1 and LESV sequences were extracted on the basis of the phylogenetic tree depicted previously (*14*) and aligned using Clustal Omega version 1.2.4. Alignments were annotated, and the conservation was depicted using CLC Genomics Workbench 12 (QIAGEN). Secondary structure predictions shown above the conservation plots are annotated on the basis of the respective *Arabidopsis* AlphaFold structures, irrespective of the respective regions' pLDDT values.

WebLogos were constructed using WebLogo 3 (*32*) using compositional adjustment assuming a typical amino acid usage pattern. Predicted protein structures were obtained from AlphaFold [F4I9G2 (ESV1) and Q5EAH9 (LESV)] (*17*) and visualized using PyMOL (*33*).

Cloning, expression, and purification of recombinant proteins

The coding sequence (CDS) of *LESV* (minus the 56 N-terminal residues predicted to encode the transit peptide), codon optimized for *Escherichia coli*, was cloned into the vector pET28a+ (Novagen), in-frame with an N-terminal 6XHis tag followed by a thrombin cleavage site, giving rise to the construct *pET28a-LESV*. Similarly, the codon-optimized CDS of *ESV1* was inserted into pET24a+ in-

frame with a C-terminal 6XHis tag, giving rise to *pET24a-ESV1*. Because commonly used algorithms do not reliably predict a transit peptide for ESV1, we omitted 95 N-terminal residues, corresponding to a region that is poorly conserved in alignments of orthologous sequences.

pET28a-LESV was transformed into *E. coli* strain Rosetta (DE3) pLysS (Novagen). Precultures were grown overnight in LB medium supplemented with antibiotics at 37°C with agitation. For the main culture, ZY liquid broth supplemented with antibiotics was inoculated and incubated with agitation for 3 hours at 37°C, and bacterial cultures were then incubated overnight at 20°C. Cells were collected; resuspended in 50 mM tris buffer (pH 8), 300 mM NaCl, 10 mM imidazole, bovine pancreas deoxyribonuclease I (10 μg/ml), 20 mM MgSO4, and complete protease inhibitor cocktail EDTA-free tablets (one tablet per liter of culture; Roche) at 4°C; and then disrupted using an Avestin Emulsiflex disruptor. The bacterial extract containing soluble proteins was collected by centrifugation (10,000*g*) for 1 hour at 4°C. The extract was subjected to a first step of purification by immobilized metal affinity chromatography (IMAC) using His-Trap FP 5-ml column (Cytiva) equilibrated with 50 mM tris buffer (pH 8), 300 mM NaCl, and 10 mM imidazole. Recombinant protein was eluted using one step of 250 mM imidazole in equilibration buffer. This was followed by a second purification step by size exclusion chromatography using a HiLoad 16/60 Superdex 200 column (Cytiva) pre-equilibrated with 50 mM tris (pH 8), 150 mM NaCl, 10% (v/v) glycerol, and 2 mM dithiothreitol (DTT).

pET24a-ESV1 was transformed into the bacterial strain BL21 codon+. Precultures were grown overnight in LB medium supplemented with antibiotics at 25°C with agitation. For the main culture, LB liquid broth supplemented with antibiotics was inoculated and incubated with agitation at 37°C to an optical density of 0.5 to 0.6 (A_{600}), whereupon protein expression was induced by addition of isopropyl-β-D-thiogalactopyranoside (final concentration of 1 mM) for 3 hours at 37°C. Cultures were then incubated overnight at 20°C. Harvesting of bacteria and protein extract preparation were as described above. The extract was subjected to a first step of purification by IMAC using His-Trap FP 5-ml column (Cytiva) equilibrated in 50 mM tris (pH 7.5), 300 mM NaCl, and 40 mM imidazole. Bound protein was then eluted in one step in 50 mM tris (pH 7.5), 300 mM NaCl, and 250 mM imidazole. The protein solution was dialyzed against 50 mM tris (pH 7.5), 100 mM NaCl, 10% (v/v) glycerol, and 2 mM DTT.

Synchrotron radiation CD

CD spectroscopy data were measured by synchrotron radiation CD at the DISCO beamline of the SOLEIL Synchrotron (Gif-sur-Yvette, France). Five microlitres of ESV1 at 6.3 mg/ml and 2 μl of LESV at 13.4 mg/ml were deposited between two CaF_2 coverslips with a guaranteed pathlength of 20 and 10 μm, respectively (*34*). The beam size of 4 mm by 4 mm and the photon-flux per nanometer step of 2×10^{10} photons s^{-1} in the spectral band from 270 to 170 nm prevented radiation-induced damage (*35*). Spectra were collected consecutively over time and are the mean of 3 accumulations. The buffer baseline was recorded sequentially and subtracted from the spectra before taking into account the concentration in residues. Data were processed using CDToolX (*36*).

SAXS of recombinant proteins

For SAXSs, recombinant protein samples were concentrated using Vivaspin centrifugal concentrators with a 10-kDa cutoff (Sartorius). Protein concentrations were determined using a NanoDrop Spectrophotometer (ND1000; Thermo Fisher Scientific). The maximum concentrations obtained for ESV1 and LESV for SAXS experiment are 4 and 21 mg ml^{-1} respectively.

Recombinant protein solutions were subject to centrifugation (10 min, 4°C, 10,000g) before x-ray analysis to eliminate potential aggregates. SAXS experiments were conducted on the SWING beamline at Synchrotron SOLEIL (λ = 1.033 Å). All solutions were mixed in a fixed-temperature (15°C) quartz capillary. The monodispersed samples of proteins were injected into a size exclusion column (SEC-3, 150 Å; Agilent) using an Agilent HPLC system and eluted directly into the SAXS flowthrough capillary cell at a flow rate of 0.2 ml min^{-1} (*37*). Protein samples (50 μl each) were then injected for SAXS measurements. During the first minutes of the elution, 150 frames were collected and averaged to account for buffer scattering, which was subtracted from selected frames corresponding to the main elution peak. Data reduction to absolute units, frame averaging, and subtraction were done using FOXTROT (*37*). All subsequent data processing, analysis, and modeling steps were carried out with programs of the ATSAS suite (*38*). The radius of gyration (*Rg*) was derived by the Guinier approximation using PRIMUS (*20*). The program GNOM (*39*) was used to compute the pair-distance distribution functions [*P*(*r*)] and feature the maximum dimension of the macromolecule (*D*max).

Cloning of yeast constructs and generation and growth of strains

Plasmids and primers used for cloning are provided in data S8. Constructs for heterologous expression in *S. cerevisiae* (yeast) were cloned using the protocols and toolkit for modular cloning in yeast (*40*). The CDSs of *ESV1* and *LESV*, less their putative chloroplast transit peptides, were cloned into pYTK001 (*40*) as detailed in data S8. The CDSs were introduced between the strong yeast *CWP2* promoter (pBP124) and *CYC1* terminator (pBP224) and assembled into yeast integration vectors designed for single transcription units (pBP386 and pBP387 in case of *ESV1* and *LESV*, respectively), resulting in the final expression vectors pBP389 and pBP392 for the expression of ESV1 and LESV, respectively, from the yeast locus XII-5 of the previously described yeast expression platform (*13*, *41*). For dual expression of ESV1 and LESV, both transcription units were assembled into the vector pBP249 (remaining positions were filled using primers mimicking the connector overhangs, added directly to the Bsm BI–mediated golden gate reaction; see data S8), which also targets the XII-5 locus, resulting in the final expression vector pBP393. Plasmids containing the CDS in pYTK001 were verified by sequencing. The correctness of the integration and final expression vectors was confirmed by diagnostic restriction digests using Pst I and Sac II.

S. cerevisiae strains derive from haploid CEN.PK113-11C. Strains expressing ESV1 and/or LESV were generated by transforming Not I–linearized expression vectors pBP389, pBP392, and pBP393 into yeast strains 28 or 29 (*13*), essentially using the transformation protocol described previously (*42*). Integration at the expected locus was confirmed by polymerase chain reaction (PCR) and amplicon sequencing as described previously (*13*). Genotypes of yeasts are given in table S3.

Media were prepared as described in (*43*), and yeasts were grown in shake flasks as described in (*13*). In brief, yeast cells from overnight cultures in YPD medium [1% (w/v) Bacto yeast extract (BD), 1% (w/v) Bacto peptone (BD), and supplemented with 2% (w/v) glucose] were inoculated in main cultures containing YPGal medium and harvested after 5.75 hours with shaking at 30°C. YPGal [which contains 2% (w/v) galactose] represents the inducing condition for all transgenes, except *ESV1* and *LESV*, which are driven by the *CWP2* promoter and thus constitutively expressed. Replicate main cultures arose from independent precultures.

Plant materials and growth conditions

The mutants *esv1-2* and *lesv-1* (*14*), referred to as *esv1* and *lesv*, respectively, were crossed to *isa1-1isa2-1* (*9*), here referred to as *isa1isa2*, to generate the two triple mutants *isa1isa2esv1* and *isa1isa2lesv*. *isa1isa2* was also crossed to *ESV1-OX #3-2* and *LESV-OX #4-6* (*14*), referred to as *ESV1-OX* and *LESV-OX*, respectively, to generate *isa1isa2ESV1-OX* and *isa1isa2LESV-OX*. Plants homozygous for the respective mutations were identified by PCR-based genotyping (see table S4 for oligonucleotide primers). For the *LESV-OX #4-6* line, homozygous plants were identified by PCR-based genotyping, while for *ESV1-OX #3-2*, homozygous plants were identified by screening for BASTA resistance on plates. Wild-type (Col-0 ecotype) and mutant plants were grown as previously described (*21*). All mutant lines used in this study are in the Col-0 ecotype background.

Protein extraction from yeast and plants and immunoblotting

Total protein extracts from yeast were prepared as previously described (*13*). Protein concentration was determined with a Bradford-based protein assay (Bio-Rad) using bovine serum albumin as standard. A total of 20 μg protein was loaded per lane for SDS–polyacrylamide gel electrophoresis (SDS-PAGE).

For total protein extracts from plants, leaves from 4-week-old rosettes were harvested and snap frozen in liquid nitrogen. Glass homogenizers were used to homogenize the plant material in extraction medium [1 ml/100 mg of fresh weight; 40 mM tris-HCl (pH 6.8), 5 mM $MgCl_2$, 2% (w/v) SDS, and complete protease inhibitor (Roche)]. Insoluble debris was pelleted by centrifugation (5 min, 4°C, 20,000*g*) before loading equal sample volumes for SDS-PAGE.

For immunoblotting, proteins were transferred onto low-fluorescence polyvinylidene difluoride membranes following SDS-PAGE and probed with antibodies specifically recognizing ESV1, LESV, or green fluorescent protein (GFP). Antisera against *Arabidopsis* ESV1 and LESV (*14*) were used at concentrations of 1:200 for affinity purified anti-ESV1 and 1:3000 for anti-LESV crude serum. YFP-tagged proteins and plant actin (used as a loading control) were detected using commercial antibodies (anti-GFP: Torrey Pines Biolabs, TP401, at a concentration of 1:5000; anti-actin: Sigma-Aldrich, clone 10-B3, at a concentration of 1:10,000).

Glucan extraction, quantification, and CLD analysis

For yeast, extraction protocols used were as described previously (*13*). Soluble and insoluble glucans were quantified using enzymatic assays after digestion to glucose, as described (*44*). For *Arabidopsis*, starch and phytoglycogen were extracted and quantified as described previously (*21*). CLDs of soluble and insoluble glucans extracted from yeast were obtained using protocols described previously (*13*). Similarly, CLDs of *Arabidopsis* starch and phytoglycogen were obtained using protocols described earlier (*21*).

Iodine staining

Yeast cells from liquid cultures were stained with Lugol's solution and imaged as described in (*13*). To stain *Arabidopsis*, rosettes were harvested at specified time points, decolorized in hot 80% (v/v) ethanol, rinsed in water, stained in Lugol's solution for 3 min, and rinsed in water again before imaging.

Scanning electron microscopy

For SEM imaging of insoluble glucans from yeast, glucans were purified using Percoll cushions as described earlier (*13*), coated with platinum, and imaged using an FEI Magellan 400.

TEM and LM of embedded materials

For imaging yeast cells by TEM, protocols used were as described previously (*13*), with minor adjustments to the chemical fixation and gelling of cells before osmium staining. For the chemical fixation, cells grown in liquid culture in complex medium were prefixed with glutaraldehyde [50 mM sodium cacodylate (pH 6.8), 1 mM $MgCl_2$, 1 mM $CaCl_2$, and 2% (v/v) glutaraldehyde] for 1 hour at 20°C, pelleted by centrifugation (5 min, 20°C, 1500 *g*), and resuspended in the same fixative solution. After fixation using a BioWave (TedPella), the cells were kept at 4°C overnight. Cells were then pelleted and washed four times with water and once with 100 mM sodium cacodylate buffer (pH 6.8). Before post-staining with osmium, yeast cells were resuspended in water and mixed with one volume of 6% (w/v) low melting point agarose solution. After solidification, yeast cells were further processed, starting with osmium post-fixation as described earlier (*13*). Ultrathin (70 nm) sections were cut using a diamond knife, placed on formvar/carbon-coated copper grids, and stained with 2% (w/v) uranyl acetate and Reynold's lead citrate. Images were acquired using a JEM-1400 Plus JEOL electron microscope.

To image *Arabidopsis* chloroplasts by TEM, young leaves of 4-week-old plants were harvested at EOD and cut into small pieces. These were fixed in 100 mM sodium cacodylate (pH 7.4), 2.5% (v/v) glutaraldehyde, and 2% (w/v) formaldehyde, followed by post-staining in 1% (w/v) osmium tetroxide, dehydration, and embedding into Spurr resin as described in (*22*) and sectioned and imaged as for yeast cells above.

Spurr blocks containing leaf tissue were also used for overview imaging by LM. Semi-thin (500 nm) sections were cut using a diamond knife, stained using toluidine blue O, and imaged using an AxioImager Z2 microscope (Zeiss).

Cloning of ESV1 and LESV constructs for expression in tobacco

Sequences encoding full-length and truncated (Trp-rich regions only) ESV1 and LESV protein versions were amplified using constructs from (*14*) as templates. The resulting PCR products were first cloned into the pDONR221 vector and subsequently recombined into pB7WGY2 (for the truncated versions; in-frame with N-terminal sequence encoding the *Arabidopsis* Rubisco small subunit chloroplast transit peptide followed by sequence encoding YFP) or pB7YWG2 (for the full-length versions; in-frame with a C-terminal sequence encoding YFP) via gateway recombination

cloning technology (Invitrogen), as described previously (*45*). The resulting constructs were transformed into *Agrobacterium tumefaciens*, infiltrated into tobacco (*N. benthamiana*), and the fluorescence was imaged as described previously (*14*).

Confocal microscopy

Fluorescence from *ESV1-OX*, *LESV-OX*, *isa1isa2ESV1-OX*, and *isa1isa2LESV-OX* plants was imaged at end of the day as described previously (*46*) using a Zeiss LSM780 confocal imaging system, using either 514-nm (YFP) and 458-nm (CFP) argon or 633-nm (chlorophyll) helium-neon lasers. Image acquisition was done sequentially using filters ranging from 526 to 624 nm (YFP), 463 to 509 nm (CFP), and 647 to 721 nm (chlorophyll). At least two independent biological replicates were imaged per genotype, and imaging was repeated two or three times using different plant batches.

SAXS of yeast and plant glucans

Yeast glucans for SAXS measurements were purified as described for SEM. For plant glucan purification, whole rosettes of individual plants were harvested and the material crushed with a pestle in extraction buffer [50 mM tris-HCl (pH 8), 2 mM EDTA, and 0.5% (v/v) Triton X-100] The slurry was sequentially filtered through wet nylon meshes with 150-, 30-, and 6-μm pore sizes. Glucans were then pelleted by centrifugation (10 min, 20°C, 6000 *g*). The resulting pellet was first washed with 0.5% (w/v) SDS until the supernatant was clear and then with water.

Approximately 20 μl of dense glucan suspension was injected into glass mark tubes (Hilgenberg) with an outer diameter of 0.7 mm and a wall thickness of 0.01 mm and sealed using a two-component epoxy resin. Glucans were allowed to settle before the capillaries were inserted into a laboratory SAXS system (Xenocs Xeuss 3.0). The scattering signal was recorded for the starch-enriched regions under vacuum using Cu Kα x-ray radiation (λ Cu Kα = 1.5419 Å) and a 2D detector (DECTRIS EIGER2 1M) positioned at a sample detector distance of 1300 mm. The recorded 2D scattering signal was azimuthally integrated to obtain the 1D scattering signal (scattering intensity as a function of the scattering vector). The background scattering from the capillary and water was scaled and then subtracted from the recorded signals to obtain the scattering signal solely from the starch. All data handling was performed using Xenocs XSACT software.

Accession numbers

The *Arabidopsis* genome initiative gene codes for the *Arabidopsis* genes used in this study are the following: *ESV1*, *At1g42430*; *LESV*, *At3g55760*; *ISA1*, *At2g39930*; and *ISA2*, *At1g03310*.

Supplementary Materials

This PDF file includes:
Supplementary Text
Figs. S1 to S9
Tables S1 to S4
Legends for data S1 to S8

Other Supplementary Material for this manuscript includes the following:
Data S1 to S8

View/request a protocol for this paper from *Bio-protocol*.

REFERENCES AND NOTES

1. D. Santelia, S. C. Zeeman, Progress in Arabidopsis starch research and potential biotechnological applications. *Curr. Opin. Biotechnol.* **22**, 271–280 (2011).
2. M. Stitt, S. C. Zeeman, Starch turnover: Pathways, regulation and role in growth. *Curr. Opin. Plant Biol.* **15**, 282–292 (2012).
3. S. Pérez, E. Bertoft, The molecular structures of starch components and their contribution to the architecture of starch granules: A comprehensive review. *Starch* **62**, 389–420 (2010).
4. P. J. Jenkins, R. E. Cameron, A. M. Donald, A universal feature in the structure of starch granules from different botanical sources. *Starch* **45**, 417–420 (1993).
5. T. Sawada, Y. Nakamura, T. Ohdan, A. Saitoh, P. B. Francisco Jr., E. Suzuki, N. Fujita, T. Shimonaga, S. Fujiwara, M. Tsuzuki, Diversity of reaction characteristics of glucan branching enzymes and the fine structure of α-glucan from various sources. *Arch. Biochem. Biophys.* **562**, 9–21 (2014).
6. B. Pfister, S. C. Zeeman, Formation of starch in plant cells. *Cell. Mol. Life Sci.* **73**, 2781–2807 (2016).
7. S. Ball, H.-P. Guan, M. James, A. Myers, P. Keeling, G. Mouille, A. Buléon, P. Colonna, J. Preiss, From glycogen to amylopectin: A model for the biogenesis of the plant starch granule. *Cell* **86**, 349–352 (1996).
8. A. M. Myers, M. K. Morell, M. G. James, S. G. Ball, Recent progress toward understanding biosynthesis of the amylopectin crystal. *Plant Physiol.* **122**, 989–998 (2000).
9. T. Delatte, M. Trevisan, M. L. Parker, S. C. Zeeman, 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 (2005).
10. S. C. Zeeman, T. Umemoto, W.-L. Lue, P. Au-Yeung, C. Martin, A. M. Smith, J. Chen, A mutant of Arabidopsis lacking a chloroplastic isoamylase accumulates both starch and phytoglycogen. *Plant Cell* **10**, 1699–1711 (1998).
11. F. Wattebled, Y. Dong, S. Dumez, D. Delvallé, V. Planchot, P. Berbezy, D. Vyas, P. Colonna, M. Chatterjee, S. Ball, Mutants of Arabidopsis lacking a chloroplastic isoamylase accumulate phytoglycogen and an abnormal form of amylopectin. *Plant Physiol.* **138**, 184–195 (2005).
12. S. Streb, T. Delatte, M. Umhang, S. Eicke, M. Schorderet, D. Reinhardt, S. C. Zeeman, Starch granule biosynthesis in Arabidopsis is abolished by removal of all debranching enzymes but restored by the subsequent removal of an endoamylase. *Plant Cell* **20**, 3448–3466 (2008).
13. B. Pfister, A. Sánchez-Ferrer, A. Diaz, K. Lu, C. Otto, M. Holler, F. R. Shaik, F. Meier, R. Mezzenga, S. C. Zeeman, Recreating the synthesis of starch granules in yeast. *eLife* **5**, e15552 (2016).
14. D. Feike, D. Seung, A. Graf, S. Bischof, T. Ellick, M. Coiro, S. Soyk, S. Eicke, T. Mettler-Altmann, K. J. Lu, The starch granule-associated protein EARLY STARVATION1 is required for the control of starch degradation in *Arabidopsis thaliana* leaves. *Plant Cell* **28**, 1472–1489 (2016).
15. S. Helle, F. Bray, J. Verbeke, S. Devassine, A. Courseaux, M. Facon, C. Tokarski, C. Rolando, N. Szydlowski, Proteome analysis of potato starch reveals the presence of new starch metabolic proteins as well as multiple protease inhibitors. *Front. Plant Sci.* **9**, 746 (2018).
16. K. Song, D.-W. Lee, J. Kim, J. Kim, H. Guim, K. Kim, J.-S. Jeon, G. Choi, EARLY STARVATION 1 is a functionally conserved protein promoting gravitropic responses in plants by forming starch granules. *Front. Plant Sci.* **12**, 628948 (2021).
17. J. Jumper, R. Evans, A. Pritzel, T. Green, M. Figurnov, O. Ronneberger, K. Tunyasuvunakool, R. Bates, A. Žídek, A. Potapenko, Highly accurate protein structure prediction with AlphaFold. *Nature* **596**, 583–589 (2021).
18. M. Varadi, S. Anyango, M. Deshpande, S. Nair, C. Natassia, G. Yordanova, D. Yuan, O. Stroe, G. Wood, A. Laydon, AlphaFold Protein Structure Database: Massively expanding the structural coverage of protein-sequence space with high-accuracy models. *Nucleic Acids Res.* **50**, D439–D444 (2022).
19. D. Franke, D. I. Svergun, DAMMIF, a program for rapid ab-initio shape determination in small-angle scattering. *J. Appl. Cryst.* **42**, 342–346 (2009).
20. P. V. Konarev, V. V. Volkov, A. V. Sokolova, M. H. J. Koch, D. I. Svergun, PRIMUS: A Windows PC-based system for small-angle scattering data analysis. *J. Appl. Cryst.* **36**, 1277-1282 (2003).
21. B. Pfister, K. J. Lu, S. Eicke, R. Feil, J. E. Lunn, S. Streb, S. C. Zeeman, Genetic evidence that chain length and branch point distributions are linked determinants of starch granule formation in Arabidopsis. *Plant Physiol.* **165**, 1457–1474 (2014).
22. L. Bürgy, S. Eicke, C. Kopp, C. Jenny, K. J. Lu, S. Escrig, A. Meibom, S. C. Zeeman, Coalescence and directed anisotropic growth of starch granule initials in subdomains of *Arabidopsis thaliana* chloroplasts. *Nat. Commun.* **12**, 6944 (2021).

23. K. Ramirez-Gualito, R. Alonso-Rios, B. Quiroz-Garcia, A. Rojas-Aguilar, D. Diaz, J. Jimenez-Barbero, G. Cuevas, Enthalpic nature of the CH/π interaction involved in the recognition of carbohydrates by aromatic compounds, confirmed by a novel interplay of NMR, calorimetry, and theoretical calculations. *J. Am. Chem. Soc.* **131**, 18129–18138 (2009).
24. S. Zhang, K. Y. Chen, X. Zou, Carbohydrate-protein interactions: Advances and challenges. *Commun. Inf. Syst.* **21**, 147–163 (2021).
25. R. Lorberth, G. Ritte, L. Willmitzer, J. Kossmann, Inhibition of a starch-granule–bound protein leads to modified starch and repression of cold sweetening. *Nat. Biotechnol.* **16**, 473–477 (1998).
26. T.-S. Yu, H. Kofler, R. E. Hausler, D. Hille, U.-I. Flugge, S. C. Zeeman, A. M. Smith, J. Kossmann, J. Lloyd, G. Ritte, 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 (2001).
27. G. Ritte, J. R. Lloyd, N. Eckermann, A. Rottmann, J. Kossmann, M. Steup, The starch-related R1 protein is an α-glucan, water dikinase. *Proc. Natl. Acad. Sci. U.S.A.* **99**, 7166–7171 (2002).
28. O. Kötting, D. Santelia, C. Edner, S. Eicke, T. Marthaler, M. S. Gentry, S. Comparot-Moss, J. Chen, A. M. Smith, M. Steup, STARCH-EXCESS4 is a laforin-like phosphoglucan phosphatase required for starch degradation in *Arabidopsis thaliana*. *Plant Cell* **21**, 334–346 (2009).
29. O. Kötting, K. Pusch, A. Tiessen, P. Geigenberger, M. Steup, G. Ritte, Identification of a novel enzyme required for starch metabolism in Arabidopsis leaves. The phosphoglucan, water dikinase. *Plant Physiol.* **137**, 242–252 (2005).
30. I. Malinova, H. Mahto, F. Brandt, S. AL-Rawi, H. Qasim, H. Brust, M. Hejazi, J. Fettke, EARLY STARVATION 1 specifically affects the phosphorylation action of starch-related dikinases. *Plant J.* **95**, 126–137 (2018).
31. J. A. Cuesta-Seijo, A. J. De Porcellinis, A. H. Valente, A. Striebeck, C. Voss, L. Marri, A. Hansson, A. M. Jansson, M. H. Dinesen, J. U. Fangel, Amylopectin chain length dynamics and activity signatures of key carbon metabolic enzymes highlight early maturation as culprit for yield reduction of barley endosperm starch after heat stress. *Plant Cell Physiol.* **60**, 2692–2706 (2019).
32. G. E. Crooks, G. Hon, J.-M. Chandonia, S. E. Brenner, WebLogo: A sequence logo generator. *Genome Res.* **14**, 1188–1190 (2004).
33. L. Schrödinger, W. DeLano, The PyMOL molecular graphics system, version 2.0 (Schrödinger LLC, 2017).
34. M. Réfrégiers, F. Wien, H.-P. Ta, L. Premvardhan, S. Bac, F. Jamme, V. Rouam, B. Lagarde, F. Polack, J.-L. Giorgetta, DISCO synchrotron-radiation circular-dichroism endstation at SOLEIL. *J. Synchrotron Radiat.* **19**, 831–835 (2012).
35. A. J. Miles, R. W. Janes, A. Brown, D. Clarke, J. Sutherland, Y. Tao, B. A. Wallace, S. Hoffmann, Light flux density threshold at which protein denaturation is induced by synchrotron radiation circular dichroism beamlines. *J. Synchrotron Radiat.* **15**, 420–422 (2008).
36. A. J. Miles, B. A. Wallace, CDtoolX, a downloadable software package for processing and analyses of circular dichroism spectroscopic data. *Protein Sci.* **27**, 1717–1722 (2018).
37. G. David, J. Perez, Combined sampler robot and high-performance liquid chromatography: A fully automated system for biological small-angle x-ray scattering experiments at the Synchrotron SOLEIL SWING beamline. *J. Appl. Cryst.* **42**, 892–900 (2009).
38. D. Franke, M. V. Petoukhov, P. V. Konarev, A. Panjkovich, A. Tuukkanen, H. D. T. Mertens, A. G. Kikhney, N. R. Hajizadeh, J. M. Franklin, C. M. Jeffries, D. I. Svergun, ATSAS 2.8: A comprehensive data analysis suite for small-angle scattering from macromolecular solutions. *J. Appl. Cryst.* **50**, 1212–1225 (2017).
39. D. I. Svergun, Determination of the regularization parameter in indirect-transform methods using perceptual criteria. *J. Appl. Cryst.* **25**, 495–503 (1992).
40. M. E. Lee, W. C. DeLoache, B. Cervantes, J. E. Dueber, A highly characterized yeast toolkit for modular, multipart assembly. *ACS Synth. Biol.* **4**, 975–986 (2015).
41. M. D. Mikkelsen, L. D. Buron, B. Salomonsen, C. E. Olsen, B. G. Hansen, U. H. Mortensen, B. A. Halkier, Microbial production of indolylglucosinolate through engineering of a multigene pathway in a versatile yeast expression platform. *Metab. Eng.* **14**, 104–111 (2012).
42. R. D. Gietz, R. H. Schiestl, High-efficiency yeast transformation using the LiAc/SS carrier DNA/PEG method. *Nat. Protoc.* **2**, 31–34 (2007).
43. B. Pfister, J. M. Shields, T. Kockmann, J. Grossmann, M. R. Abt, M. Stadler, S. C. Zeeman, Tuning heterologous glucan biosynthesis in yeast to understand and exploit plant starch diversity. *BMC Biol.* **20**, 207 (2022).
44. C. Hostettler, K. Kölling, D. Santelia, S. Streb, O. Kötting, S. C. Zeeman, Analysis of starch metabolism in chloroplasts, in *Chloroplast Research in Arabidopsis* (Springer, 2011), pp. 387–410.
45. K.-J. Lu, B. Pfister, C. Jenny, S. Eicke, S. C. Zeeman, Distinct functions of STARCH SYNTHASE 4 domains in starch granule formation. *Plant Physiol.* **176**, 566–581 (2018).
46. M. R. Abt, B. Pfister, M. Sharma, S. Eicke, L. Burgy, I. Neale, D. Seung, S. C. Zeeman, STARCH SYNTHASE5, a noncanonical starch synthase-like protein, promotes starch granule initiation in Arabidopsis. *Plant Cell* **32**, 2543–2565 (2020).

Acknowledgments: We thank A. Ruckle for help with plant cultivation, ScopeM for providing access to microscopy facilities, and T. Weber from the ETH x-ray service platform for technical support with the SAXS equipment, and I. Burgert for access. We further acknowledge DISCO beamline and the regular access to the SAXS beamline SWING at synchrotron SOLEIL (St. Aubin, France) through BAG MX-20181002 and MX-20201190 and are thankful for the expert technical support provided by the beamline staff. **Funding:** This work was supported by the Swiss National Science Foundation (grant 31003A_182570 to S.C.Z.), a Chinese Scholarship Council PhD fellowship (NSCIS no. 201706310142 to C.L.), and ETH Zurich (core funding to S.C.Z.).
Author contributions: S.C.Z. and C.L. conceived and designed the experiment with input from B.P., C.B., and M.R.A. C.L., M.R.A., R.O., M.R., D.S., A.H., M.S., M.F.-S., and S.E. performed the experiments. C.L., R.O., M.R., C.B., and M.R.A. analyzed the data. S.C.Z., C.L. and M.R.A. wrote the manuscript. S.C.Z., B.P., C.L., C.B., and M.R.A. revised the manuscript. **Competing interests:** The authors declare that they have no competing interests. **Data and materials availability:** All data needed to evaluate the conclusions in the paper are present in the paper and/or the Supplementary Materials. SAXS data have also been deposited in the Small Angle Scattering Biological Data Bank (SASBDB) with the following entry codes: ESV1, SASDRR5; LESV, SASDRS5.

Submitted 18 January 2023
Accepted 19 April 2023
Published 26 May 2023
10.1126/sciadv.adg7448

ScienceAdvances

LIKE EARLY STARVATION 1 and EARLY STARVATION 1 promote and stabilize amylopectin phase transition in starch biosynthesis

Chun Liu, Barbara Pfister, Rayan Osman, Maximilian Ritter, Arvid Heutinck, Mayank Sharma, Simona Eicke, Michaela Fischer-Stettler, David Seung, Coralie Bompard, Melanie R. Abt, and Samuel C. Zeeman

Sci. Adv., **9** (21), eadg7448.
DOI: 10.1126/sciadv.adg7448

Science Advances
AAAS

Supplementary Materials for

LIKE EARLY STARVATION 1 and EARLY STARVATION 1 promote and stabilize amylopectin phase transition in starch biosynthesis

Chun Liu *et al.*

Corresponding author: Samuel C. Zeeman, szeeman@ethz.ch

Sci. Adv. **9**, eadg7448 (2023)
DOI: 10.1126/sciadv.adg7448

The PDF file includes:

Other Supplementary Material for this manuscript includes the following:

Supplementary Text

Given the marked increase in insoluble glucans and the altered appearance of the particles upon expression of ESV1 and/or LESV, either in yeast cells or in the *isa1isa2* mutant background, we investigated whether there were changes to the glucans' primary structure, as revealed by their chain length distribution (CLD) profiles. Insoluble and soluble glucans from each line were debranched enzymatically and the resultant linear chains analyzed by HPAEC-PAD (Figure S5).

CLD analysis of glucans in the yeast lines:

The yeast line 28 synthesized soluble glucan that, compared to Arabidopsis amylopectin, was rich in short (DP4-7) and medium length (DP16-30) chains, and had fewer short (DP10-15) and long (DP>36) chains. Its structure was more similar – though not identical – to soluble phytoglycogen from the Arabidopsis *isa1isa2* double mutant. The CLDs of the insoluble glucans resulting from expression of ESV1, LESV, or both proteins in line 28, i.e. lines 559, 562 and 563, were remarkably similar to that of line 28, except for decreased numbers of very short chains (particularly DP5-7) and slightly increased numbers of longer chains. The soluble glucans, which still dominated in all four lines, were essentially identical.

The CLD of the glucans from line 29 differed from those of line 28, having fewer short chains (DP6-18) and more chains over DP25 (Figure S5A; (*13*)), presumably as a result of selective debranching by ISA. Line 29 has both soluble and insoluble glucans with the former having slightly fewer short chains (DP5-9) than the latter (Figure S5A; (*13*)). Upon ESV1 expression (line 569), the CLD of the insoluble glucan was unchanged. Upon LESV expression (line 572), or expression of both proteins together (line 573), there was a further slight decrease in the number of the short chains (particularly DP3-8) and slight increases in the number of longer chains in the insoluble glucans compared with line 29. There were also slight changes in the CLDs of the soluble glucans in these lines.

Overall, these data show that presence of ESV1 and/or LESV resulted in either no or minor changes in the structure of the glucans, with the CLDs most closely resembling that of the parental line in each case. We interpret these data to mean that the proteins do not directly alter the glucan structure. The small changes could be explained if ESV1 and/or LESV promote the transition of the glucan into the insoluble phase. On the one hand this might enrich for certain glucan structures, while on the other hand it might reduce the susceptibly of the now insoluble glucan to further enzymatic modification. It is also possible that the binding of ESV1 and/or LESV to a still soluble glucan could mask it and prevent other enzymes from using it as substrate.

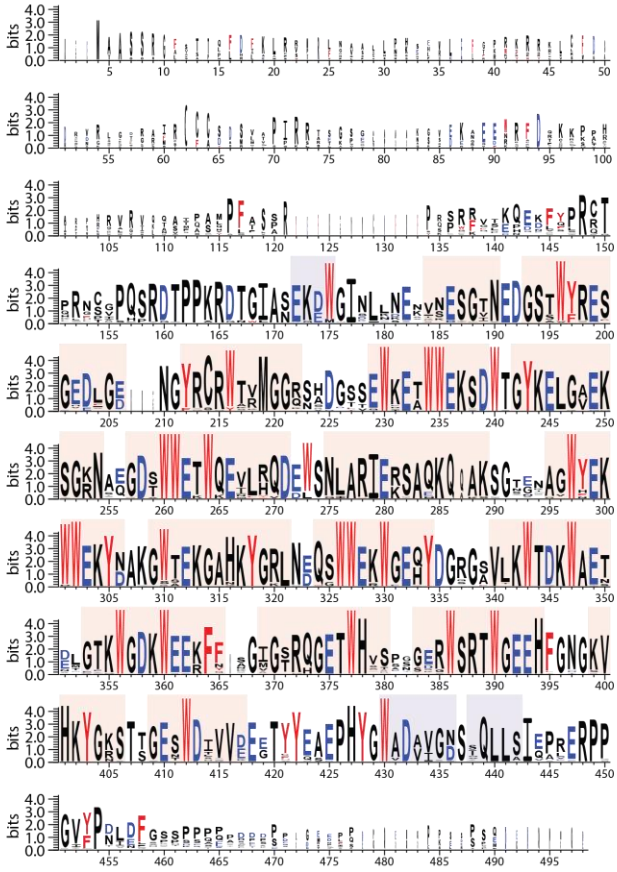
CLD analysis of starch and soluble glucans accumulating in the Arabidopsis lesv mutant after de-starching with an extended night:

After an extended night, the *lesv* mutant accumulates a mixture of starch and soluble glucan the following day (Figure 4C). The starch CLD profile from this *lesv* mutant starch was essentially indistinguishable from that of the wild type (Supplemental Figure 5B). However, the soluble glucan had a different structure, with increased numbers of chains of DP3-8 and a decrease in most chains of DP11 and longer. These changes are reminiscent of the soluble phytoglycogen of *isa1isa2* (*9*). The marked increase in chains, especially those shorter that DP6, are a strong indicator of amylolytic degradation, since the shortest chain transferred by branching enzymes during amylopectin biosynthesis is DP6 (*5*). This is supported by the very high levels of the β-amylase degradation product, maltose, in the *lesv* mutant under these conditions (Figure 4E).

CLD analysis of glucans in the Arabidopsis lines modulated for ESV1, LESV and the ISA1/ISA2 debranching enzyme:

Starch extracted at the end of a normal day from the wild type, and from the *esv1* and *lesv* mutants had CLD profiles that were indistinguishable from each other (Figure S5C), consistent with earlier findings (*14*). In *isa1isa2* starch, however, there was a slight increase in chains of DP5-8 and a decrease in chains of DP12-16. These changes were more pronounced in the soluble phytoglycogen of *isa1isa2*, also as previously described (*9*).
The overexpression of ESV1 and LESV, in *isa1isa2ESV1-OX* and *isa1isa2LESV-OX* respectively, hardly altered the CLD profile of starch from *isa1isa2* - only chains of DP5-8 were slightly decreased in abundance. The phytoglycogen CLDs from *isa1isa2ESV1-OX* and *isa1isa2LESV-OX* were identical to that of *isa1isa2* (Figure S5C). The loss of ESV1 in the *isa1isa2esv1* triple mutant caused a minor increase in short chains (DP5-9) in the starch CLD, relative to *isa1isa2*. The loss of LESV in the *isa1isa2lesv* triple further enriched these short chains in the starch and resulted in fewer chains of DP12-16. This CLD was conspicuously similar to that of the phytoglycogen, and, since *isa1isa2lesv* accumulated very little starch (Figure 6A and B), may represent some phytoglycogen contamination of the insoluble material. The CLDs of the phytoglycogen, were more similar between the three genotypes (Figure S5C).
As with the yeast glucans, the CLDs of starch and phytoglycogen in these plants is unaltered or altered only to a minor degree depending on whether ESV1 and LESV are present.

A



B

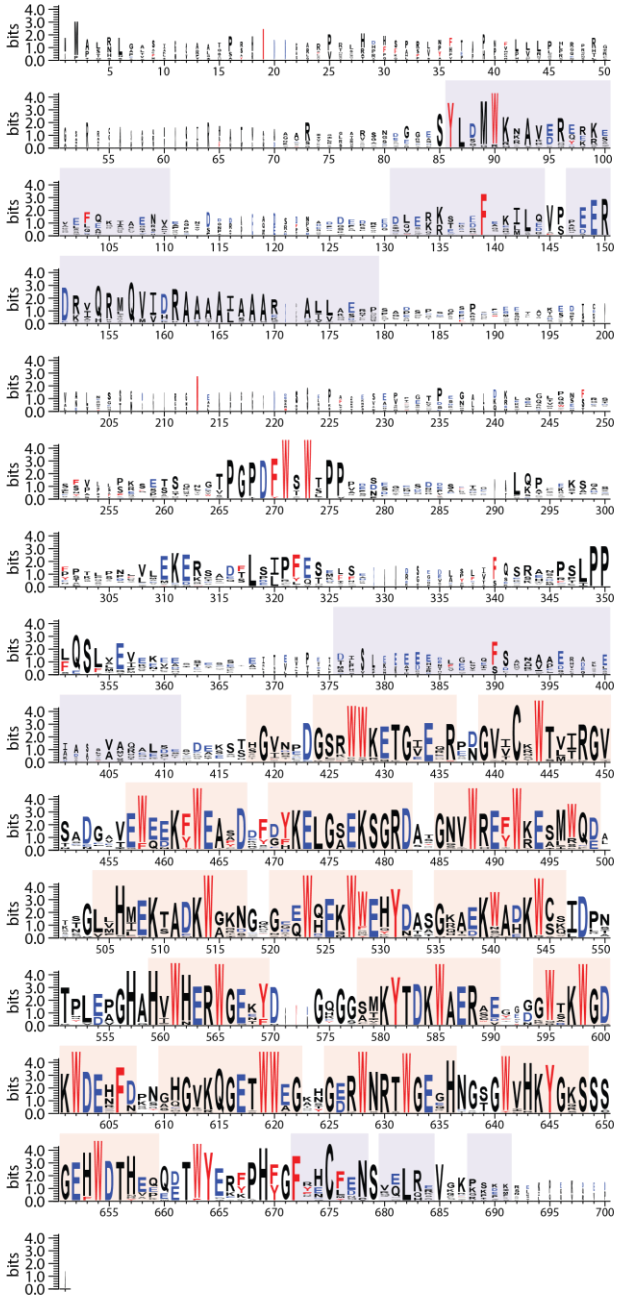


Fig. S1. Conservation hotspots in ESV1-like and LESV-like sequences. Displayed are WebLogos of the multiple sequence alignments of orthologous ESV1 (**A**) and LESV (**B**) sequences (used to create main Figure 1A). Aromatic amino acids are colored red, acidic ones blue. Underlaid in transparent boxes are the AlphaFold secondary structure predictions for the respective Arabidopsis protein, mapped to the alignment (coloring as in Figure 1A).

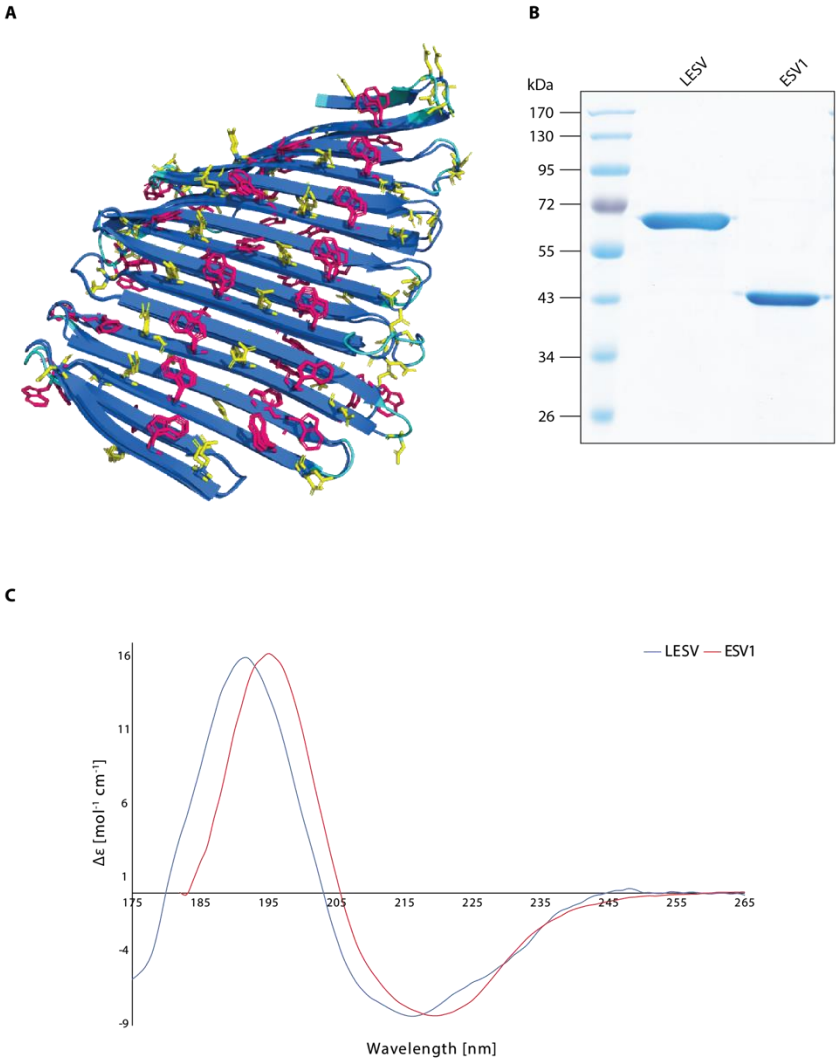


Fig. S2. The ESV1 and LESV beta-sheet planes. (**A**) Shown is a closeup of the overlaid ESV1 and LESV beta-sheet planes (as displayed in main Figure 1C). (**B**) Coomassie blue stained 10% SDS-PAGE. For both LESV and ESV1, 2 µg of the purified protein samples used for SAXS and CD experiments were loaded. (**C**) Far UV SRCD spectra of recombinant ESV1 and LESV proteins. Spectra represent the mean of three independent acquisitions. Only data with a HT (High Tension) of the photomultiplier tube > 400 V are represented.

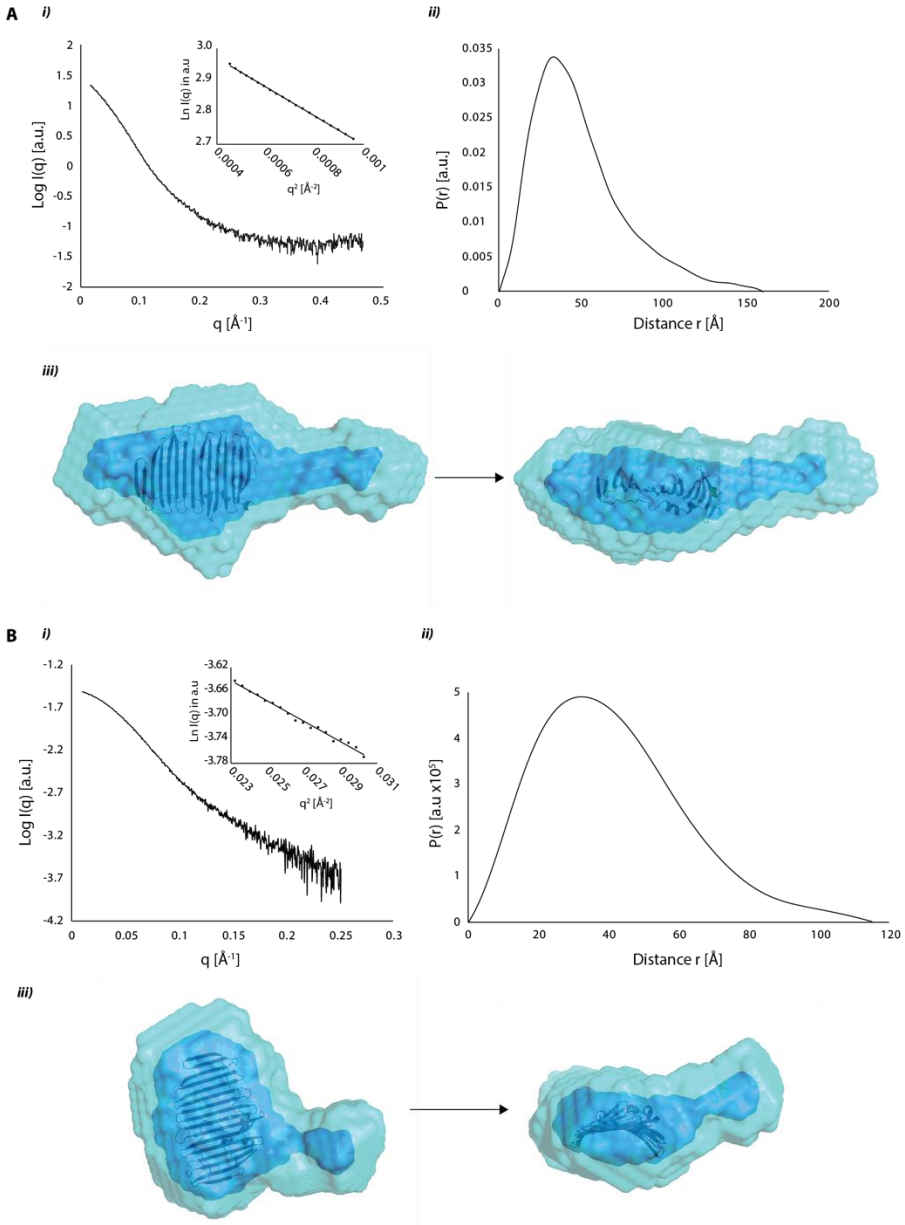


Fig. S3. SAXS data and *ab initio* models for recombinant LESV (A) and ESV1 (B) proteins. *i*) Experimental data plotted as a function of the scattering vector q (the Guinier plot is represented as an inset), *ii*) Distance distribution function and *iii*) Average *ab initio* envelope predicted by DAMMIF/DAMAVER (from 10 *ab initio* protein models computed by DAMMIF, in light cyan) superposed with the most typical shape filtered by DAMFILT (navy blue). The proteins' structural parts modeled with a confidence score (pLDDT) ≥ 90 (cartoon) are manually fitted into the molecular envelope. a.u., arbitrary unit.

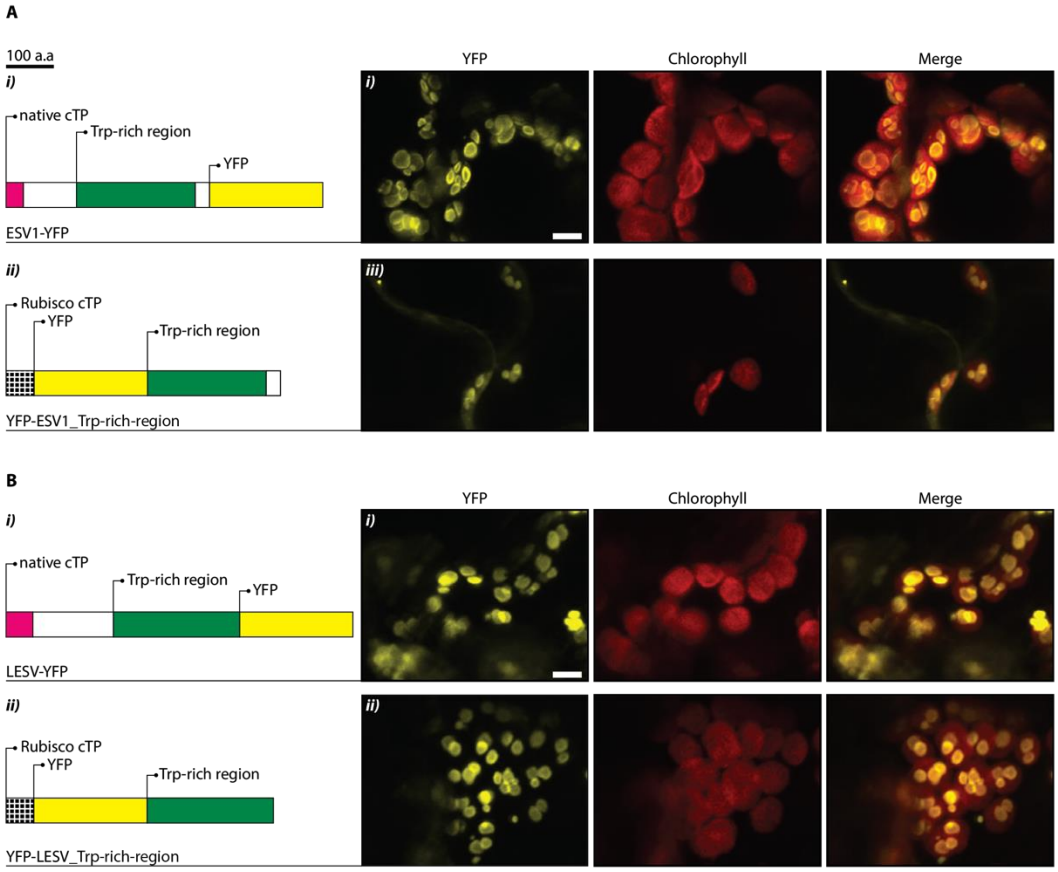
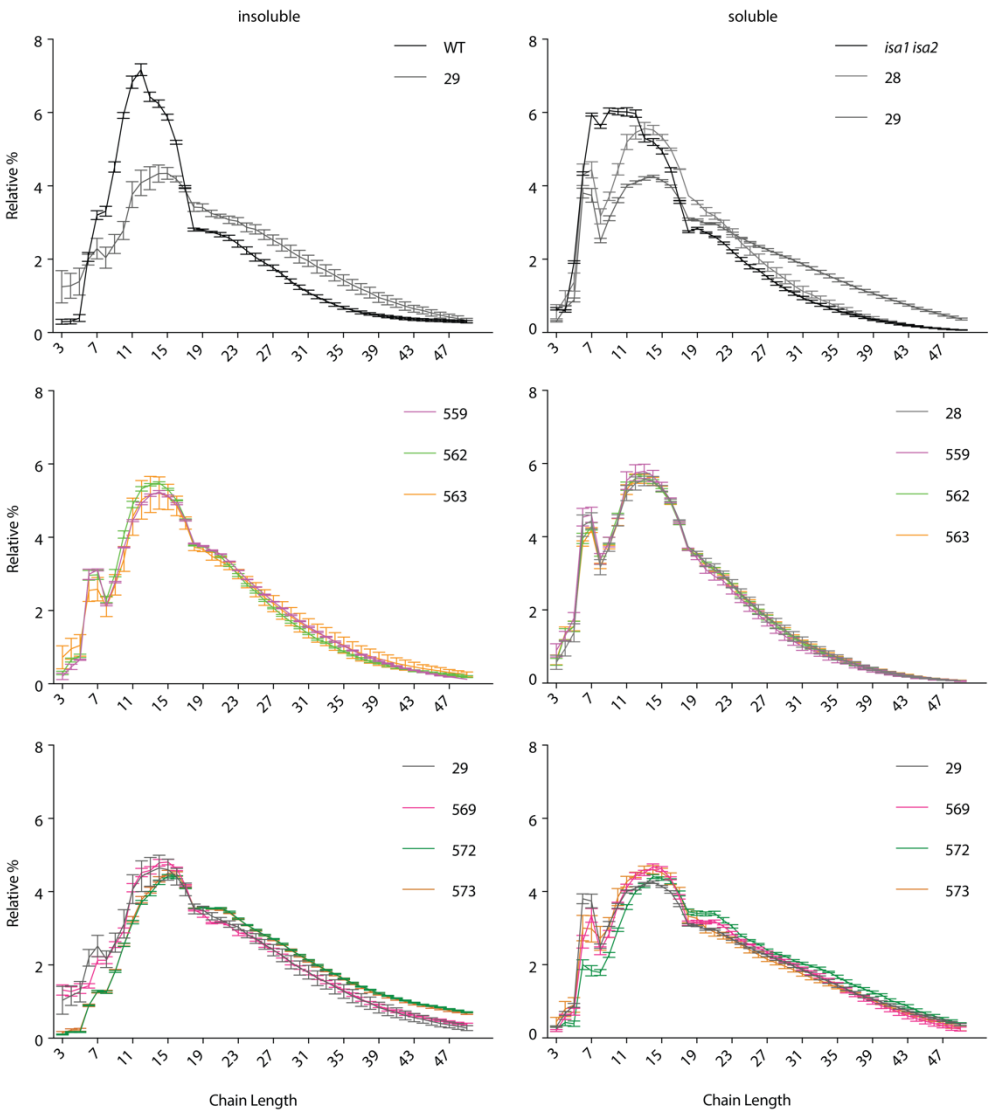


Fig. S4. Sub-plastidial localization of full-length and truncated ESV1- and LESV variants. Full-length and N-terminally truncated, YFP-tagged variants of ESV1 (**A**) and LESV (**B**) were transiently expressed in *Nicotiana benthamiana* leaves and their localizations determined by confocal microscopy. The respective protein variants' domain structures (*i-ii* for both ESV1 and LESV) are schematically depicted on the left of the confocal images (scale indicated in A is valid for B, too; a.a., amino acid residues). The Arabidopsis Rubisco small subunit cTP was used for the N-terminal YFP fusions. Presence of the Trp-rich regions alone is sufficient to mediate starch granule binding in either case. Scale bar, 5 μm.

A


insoluble
soluble
WT
29
isa1 isa2
28
29
559
562
563
28
559
562
563
29
569
572
573
29
569
572
573
Relative %
Chain Length

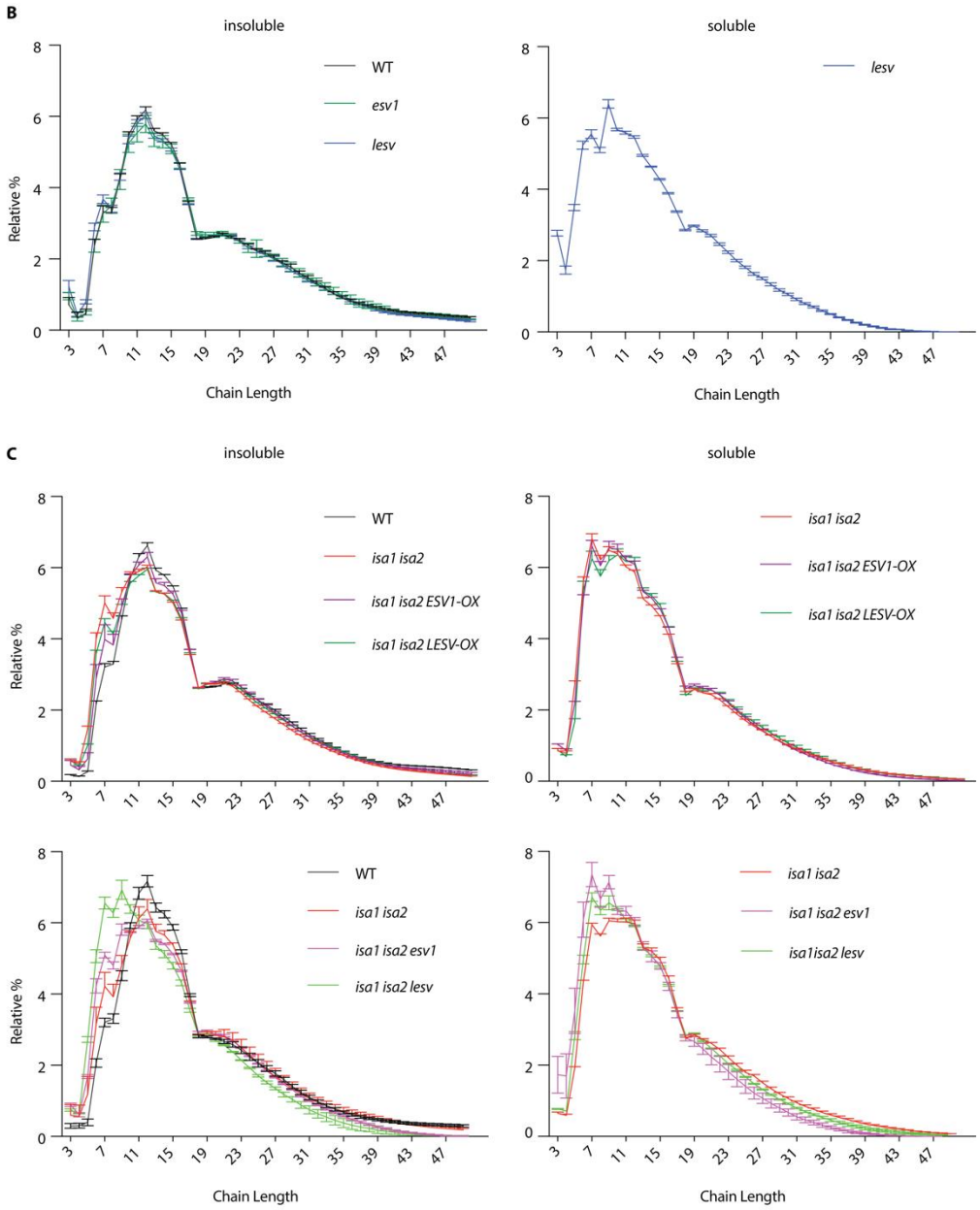


Fig. S5. Fine structure of yeast and plant glucans. Chain length distributions (CLDs) of (**A**) debranched insoluble and soluble glucans purified from the indicated yeast strains, (**B**) *esv1* and *lesv* glucans compared to wild-type Arabidopsis starch (de-starched for 16 h and then harvested after 8 h light), and (**C**) glucans from plant lines harvested after a regular day, compared to wild type (Col-0) Arabidopsis starch and *isa1isa2* phytoglycogen. Values are means ± SE (*n*=4 independent replicate cultures or biological replicate plants, respectively); relative percentages of values were obtained by dividing peak areas for each chain length by the sum of all analyzed peaks. See Supplementary Text for a discussion of the results.

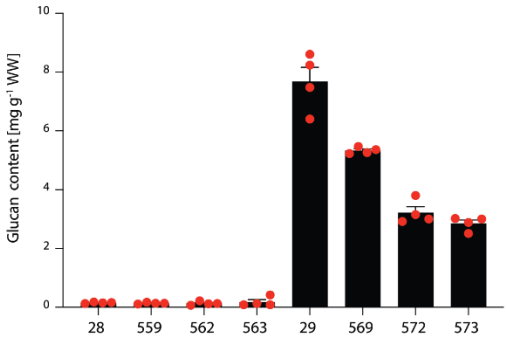


Fig. S6. MOS content of yeast strains grown for 5.75 h in liquid culture under inducing conditions. MOS were measured in the supernatant remaining after methanol precipitation of soluble polyglucans. MOS were enzymatically digested to glucose, which was then quantified using a spectrophotometric assay. Values are means ± SE (*n*=4 independent replicate cultures). See Supplementary Data S2A.

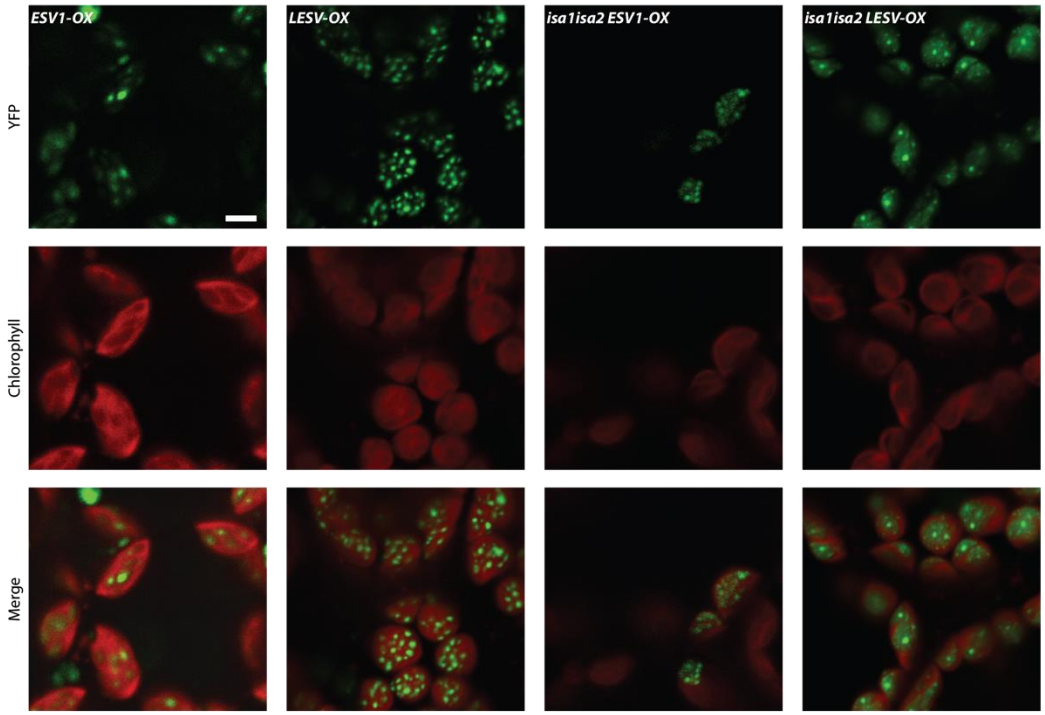


Fig. S7. ESV1 and LESV localization *in planta*. Fluorescent signal in Arabidopsis leaves of lines overexpressing YFP-tagged ESV1 or LESV, imaged at the EOD using confocal microscopy. Scale bar, 5 μm.

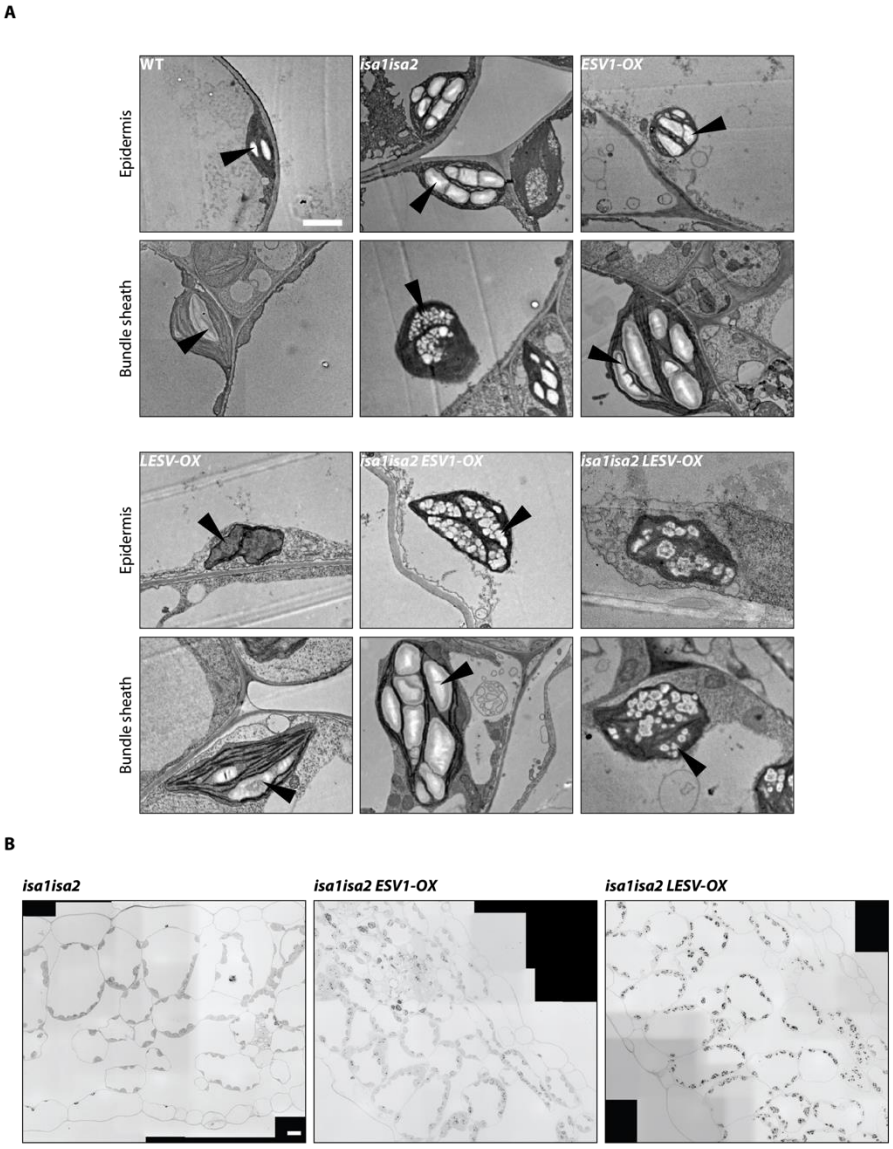


Fig. S8. TEM and overview LM images of ESV1-YFP and LESV-YFP overexpression plants. (**A**) TEM images of epidermal and bundle sheath cells. Black arrowheads indicate starch granules. Scale bar, 2 μm. (**B**) LM overviews, showing starch granules in the chloroplasts of leaf sections prepared from *isa1isa2, isa1isa2 ESV1-OX* and *isa1isa2 LESV-OX*. Individual images were stitched using Fiji to create overviews. Scale bar, 10 μm.

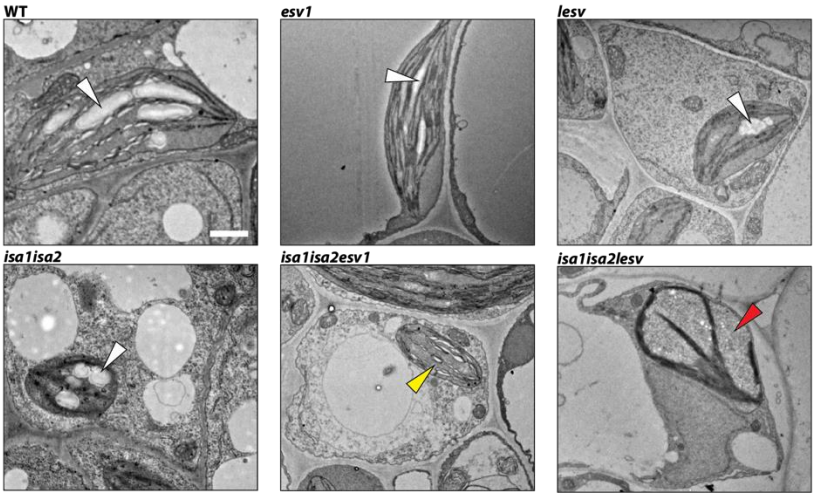


Fig. S9. TEM images of plastids of the bundle sheath. Arrowheads indicate starch granules (white), phytoglycogen (red), and empty plastid sections (yellow). Scale bar, 1 μm.

Table S1. SAXS Structural parameters.

Beamline SWING, SOLEIL	**LESV**	**ESV1**
I(0) [cm^{-1}]	23.31 +/-0.033	0.031 +/- 5.4e-05
R_g [Å]	36.79 +/-0.08	32.03 +/-0.14
q_{min} [$Å^{-1}$]	0.014	0.010
qR_g max	1.18	1.10
Coefficient of correlation, R^2	0.999	0.999
P(r) analysis		
I(0) [cm^{-1}]	23.82 +/-0.034	0.031 +/-5.8e-05
R_g [Å]	39.71 +/-0.102	32.26 +/-0.03
q range $_{min}$ [$Å^{-1}$]	0.014 to 0.4	0.010 to 0.25
Porod volume [$Å^3$]	121099	83295
Dmax	160	115

Table S2. SAXS measurements on purified yeast and plant glucans. Summary of all SAXS measurements for yeast strains and plant genotypes, with the location of the maxima in reciprocal space q_{max} and respective repeat distance d. Replicates in bold are the replicates shown in Figure 9.

Yeast strain #	***n***	**Replicate**	**q_{max} [nm^{-1}]**	**d [nm]**
562	*2*	**1**	**0.65**	**9.7**
		2	0.64	9.8
29	*2*	**1**	**0.48**	**13.1**
		2	0.48	13.1
572	*2*	**1**	**0.59**	**10.6**
		2	0.57	11.0
559	*3*	1	0.58	10.8
		2	0.6	10.5
		3	**0.59**	**10.6**
573	*1*	**1**	**0.58**	**10.8**
563	*2*	1	0.62	10.1
		2	**0.62**	**10.1**
569	*1*	**1**	**0.58**	**10.8**

Plant genotype		**Replicate**	**q_{max} [nm^{-1}]**	**d [nm]**
isa1isa2	*3*	1	0.64	9.8
		2	**0.64**	**9.8**
		3	0.63	10.0
LESV-OX	*3*	**1**	**0.64**	**9.8**
		2	0.63	10.0
		3	0.63	10
WT	*1*	**1**	**0.61**	**10.3**
esv1	*2*	1	0.59	10.6
		2	**0.58**	**10.8**
lesv	*1*	**1**	**0.61**	**10.3**
isa1isa2 ESV1-OX	*2*	**1**	**0.61**	**10.3**
		2	0.61	10.3
ESV1-OX	*1*	**1**	**0.62**	**10.1**
isa1isa2 LESV-OX	*2*	1	0.61	10.3
		2	**0.62**	**10.1**

Table S3. Description of yeast strains used in this study.

Strain	Genotype	Source
28	*MATa MAL2-8C SUC2 his3Δ gdb1Δ gph1Δ glg1Δ glg2Δ glc3::BE3 XII-2::BE2 gsy1::GlgC-TM X-2::SS1 XI-2::SS2 gsy2::SS3 XII-1::SS4*	(*13*)
559	*MATa MAL2-8C SUC2 his3Δ gdb1Δ gph1Δ glg1Δ glg2Δ glc3::BE3 XII-2::BE2 gsy1::GlgC-TM X-2::SS1 XI-2::SS2 gsy2::SS3 XII-1::SS4 XII-5::ESV1 HygR*	This study
562	*MATa MAL2-8C SUC2 his3Δ gdb1Δ gph1Δ glg1Δ glg2Δ glc3::BE3 XII-2::BE2 gsy1::GlgC-TM X-2::SS1 XI-2::SS2 gsy2::SS3 XII-1::SS4 XII-5::LESV HygR*	This study
563	*MATa MAL2-8C SUC2 his3Δ gdb1Δ gph1Δ glg1Δ glg2Δ glc3::BE3 XII-2::BE2 gsy1::GlgC-TM X-2::SS1 XI-2::SS2 gsy2::SS3 XII-1::SS4 XII-5::ESV1-LESV HygR*	This study
29	*MATa MAL2-8C SUC2 his3Δ gdb1Δ gph1Δ glg1Δ glg2Δ glc3::BE3 XII-2::BE2 gsy1::GlgC-TM X-2::SS1 X-4::SS2 gsy2::SS3 XII-1::SS4 XI-2::ISA1-ISA2*	(*13*)
569	*MATa MAL2-8C SUC2 his3Δ gdb1Δ gph1Δ glg1Δ glg2Δ glc3::BE3 XII-2::BE2 gsy1::GlgC-TM X-2::SS1 X-4::SS2 gsy2::SS3 XII-1::SS4 XI-2::ISA1-ISA2 XII-5::ESV1 HygR*	This study
572	*MATa MAL2-8C SUC2 his3Δ gdb1Δ gph1Δ glg1Δ glg2Δ glc3::BE3 XII-2::BE2 gsy1::GlgC-TM X-2::SS1 X-4::SS2 gsy2::SS3 XII-1::SS4 XI-2::ISA1-ISA2 XII-5::LESV1 HygR*	This study
573	*MATa MAL2-8C SUC2 his3Δ gdb1Δ gph1Δ glg1Δ glg2Δ glc3::BE3 XII-2::BE2 gsy1::GlgC-TM X-2::SS1 X-4::SS2 gsy2::SS3 XII-1::SS4 XI-2::ISA1-ISA2 XII-5::ESV1-LESV HygR*	This study

Table S4. Oligonucleotide primer sequences used for ESV1 and LESV cloning for expression in *N. benthamiana*, and mutant Arabidopsis alleles and primer sequences used for their genotyping.

Construct to clone	**Primers used to select mutant alleles (shown 5'to 3')**	**Reference**
ESV1-full length CDS (expression of ESV1-YFP)	Fw: GGGGACAAGTTTGTACAAAAAAGCAGGCTTC ACCATGAGCGAAATGGCGG Rv: GGGGACCACTTTGTACAAGAAAGCTGGGTCTT GTGGTTGGTCAGGG	This study
Truncated *ESV1* CDS (expression of YFP-ESV1_Trp-rich-region)	Fw: GGGGACAAGTTTGTACAAAAAAGCAGGCTTC ACCGAAGATGGTAGTAGTTGGTTTAGAG Rv: GGGGACCACTTTGTACAAGAAAGCTGGGTCTC ATTGGATCGAAAGCAACT	This study
LESV-full length CDS (expression of LESV-YFP)	Fw: GGGGACAAGTTTGTACAAAAAAGCAGGCTTC ACCATGGCTTTGCGTTTAGGTGTTTCTATAGGG Rv: GGGGACCACTTTGTACAAGAAAGCTGGGTCGG ACATATCAGAAGGCTTCTTAACGGCT	This study
Truncated *LESV* CDS (expression of YFP-LESV_Trp-rich-region)	Fw: GGGGACAAGTTTGTACAAAAAAGCAGGCTTC ACCGAAGATGGATTGAAGTGGTGGAAGCAAA CG Rv: GGGGACCACTTTGTACAAGAAAGCTGGGTCCT AGGACATATCAGAAGGCT	This study

Gene/AGI code	**Mutation type, position**	**Mutant allele**	**Primers used to select mutant alleles (shown 5'to 3')**	**Ecotype**	**Reference**
ESV1 AT1G42430	T-DNA Insertion in intron 1	*esv1-2*, GABI_031C11	Fw: CTCCAAGGCTTACTGGTCCTC Rv: ATATTGACCATCATACTCATTGC T-DNA: CTACAAATTGCCTTTTCTTATCGAC	Col-0	(*14*)
LESV AT3G55760	T-DNA Insertion in intron 1	*lesv-1*, SALK_006713	Fw: CTTTGAGAAGACAGTGGGTGG Rev: ATAAGTGGTGCAGCATTGACC T-DNA: TGGTTCACGTAGTGGGCCATCG	Col-0	(*14*)
LESV-OX	Single copy Insertion in	#4-6	Fw: ACGATCGAATGGTGAGCACTCAATTC Rv: GCTATGTCCAATCCCATCAATCACAGC T-DNA: CGCAATTATACATTTAATACGCG	Col-0	(*14*)
ISA1 AT2G39930	T-DNA Insertion in exon 13	*isa1-1*, SALK_042704	Fw: GGGACAGCCTATGTGATCTGCC Rv: TGGGAAACCATGAGGGAAACA T-DNA: GCGTGGACCGCTTGCTGCAACT	Col-0	(*21*)
ISA2 AT1G03310	X-ray Single base pair deletion in exon 1	*isa2-1*	Fw: GGTGACGTATTTACCGATGGA Rev: TGACACTTTGAGCAGCAACC The *isa2-1* amplicon is cut by NlaIV	Col-0	(*21*)

Data S1A. (separate file)
Alignment of ESV1 protein sequences used to create Figure 1A and S1A.

Data S1B. (separate file)
Alignment of LESV protein sequences used to create Figure 1A and S1B.

Data S2. (separate file)
Yeast glucan quantification. Raw data and statistical analysis shown in Figure 2B and S6.

Data S3. (separate file)
Glucan quantification of plants grown under regular 12-h light/12-h dark cycles. Raw data and statistical analysis shown in Figure 4A.

Data S4. (separate file)
Glucan quantification of plants initially grown under regular 12-h light/12-h dark cycles and then subjected to a single prolonged night (16 h dark). Raw data and statistical analysis shown in Figure 4C.

Data S5. (separate file)
Maltose quantification of plants initially grown under different diel cycle regimes. Raw data and statistical analysis shown in Figure 4E and F.

Data S6. (separate file)
Glucan quantification of *ESV1-OX* and *LESV-OX* plants. Raw data and statistical analysis shown in Figure 5C.

Data S7. (separate file)
Glucan quantification of triple mutant plants. Raw data and statistical analysis shown in Figure 7B.

Data S8. (separate file)
Plasmids and primers used for heterologous gene expression in yeast.

Annex

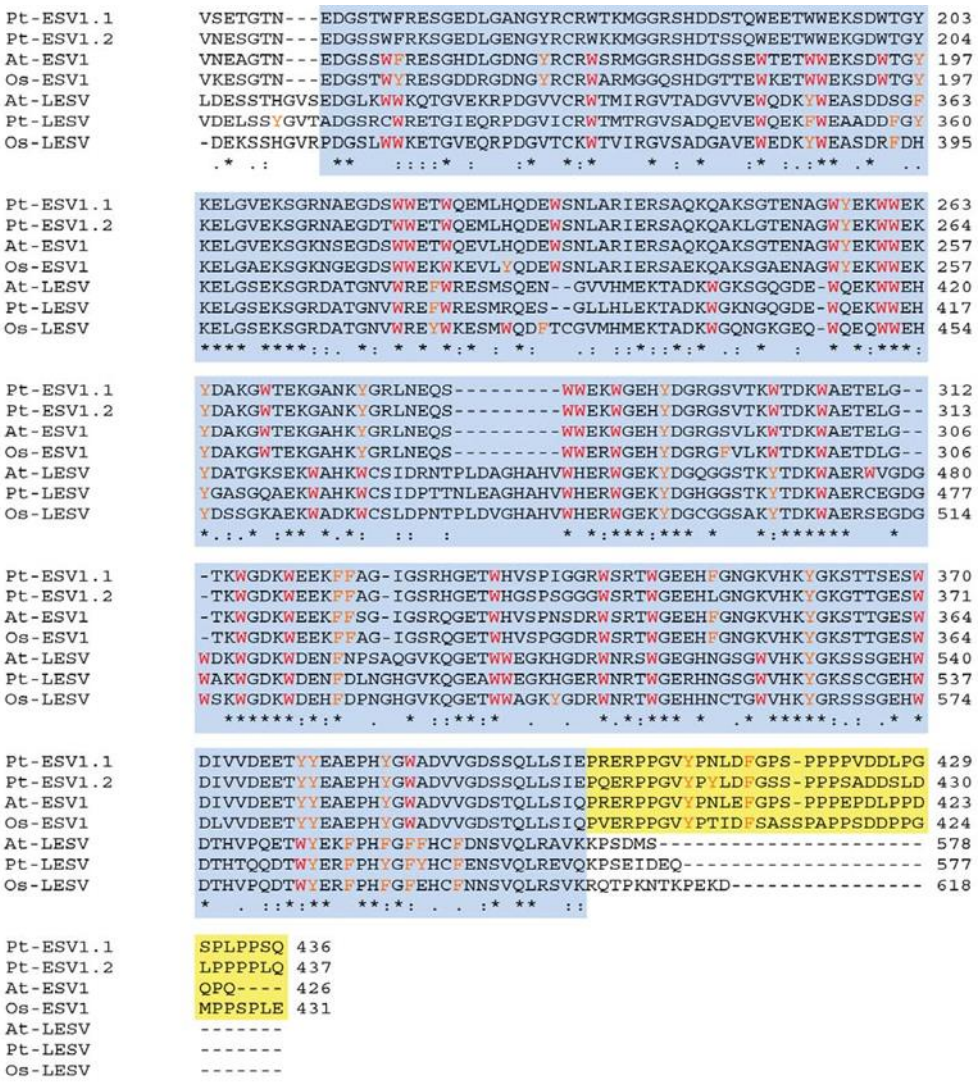


Figure S1: Sequence alignment produced by ClustalO of ESV1 and LESV protein sequences from different organisms, specifically, the tryptophan-rich regions of these proteins from *Arabidopsis*, poplar (*Populus trichocarpa*), and rice (*Oryza sativa*). The tryptophan-rich region is shown in blue, while the proline-rich region of ESV1 is shown in yellow. Tryptophan residues are shown in red, other aromatic amino acids are shown in orange. Below the protein sequences is a key denoting conserved sequence (*), conservative mutations (:), semi-conservative mutations (.), and non-conservative mutations ( ) (Feike et al. 2016).

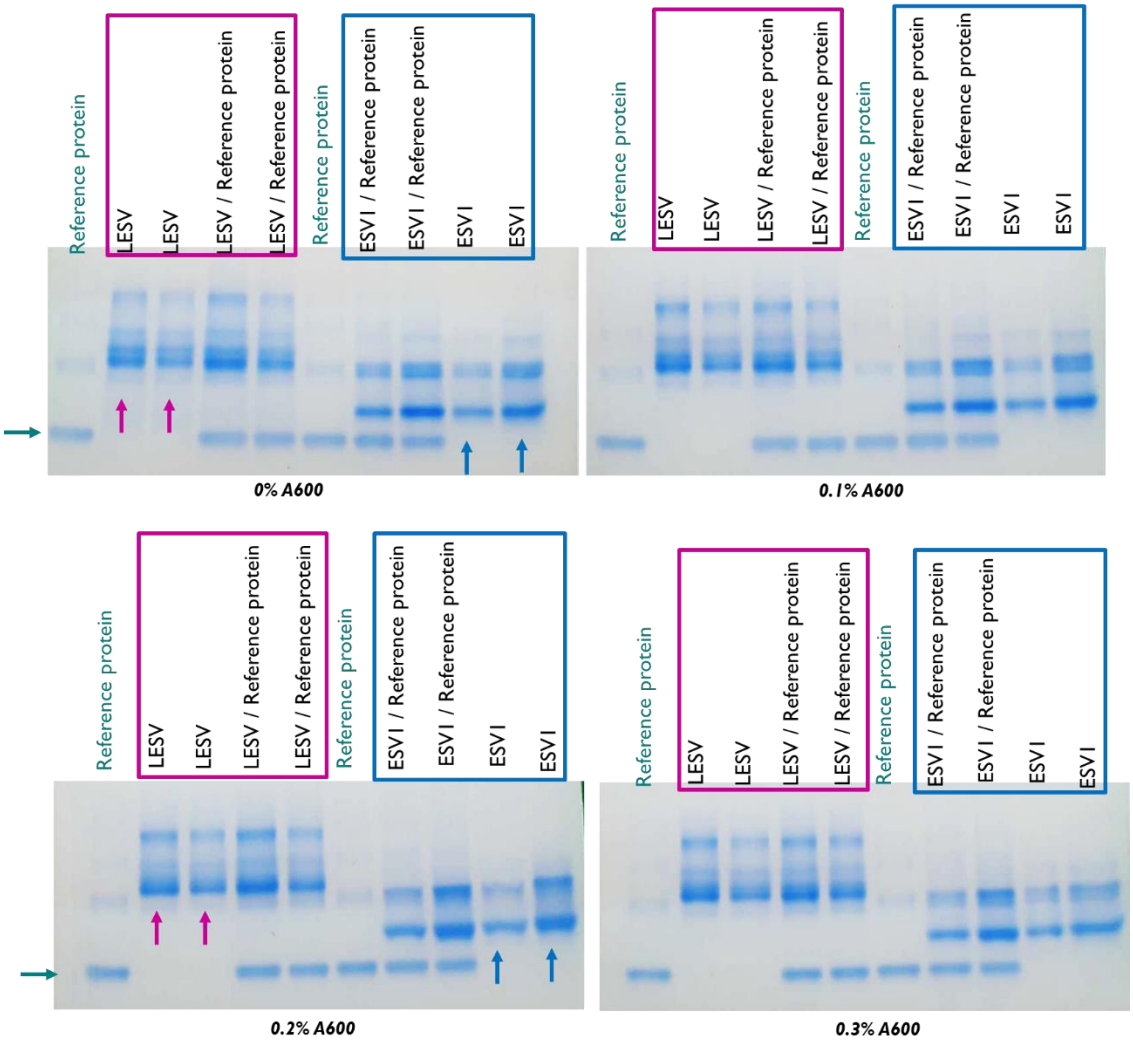


Figure S2: ***Analysis of ESV1 and LESV binding to A600 by EMSA gel with a concentration of 0%, 0,1%, 0.2% and 0.3% of potential ligands.*** Separation of both proteins using native gels, separation gels were supplemented with 0,1%, 0.2%, 0.3% of A600. Here, 1µg of ESV1 and LESV were loaded per lane. As a further control 1µg of the reference protein was separated on native gels. The gels were stained with bromophenol blue. For each gel, the green arrow indicates the migration of the reference protein. The blue and the pink arrows indicate the migration of ESV1 and LESV proteins respectively.

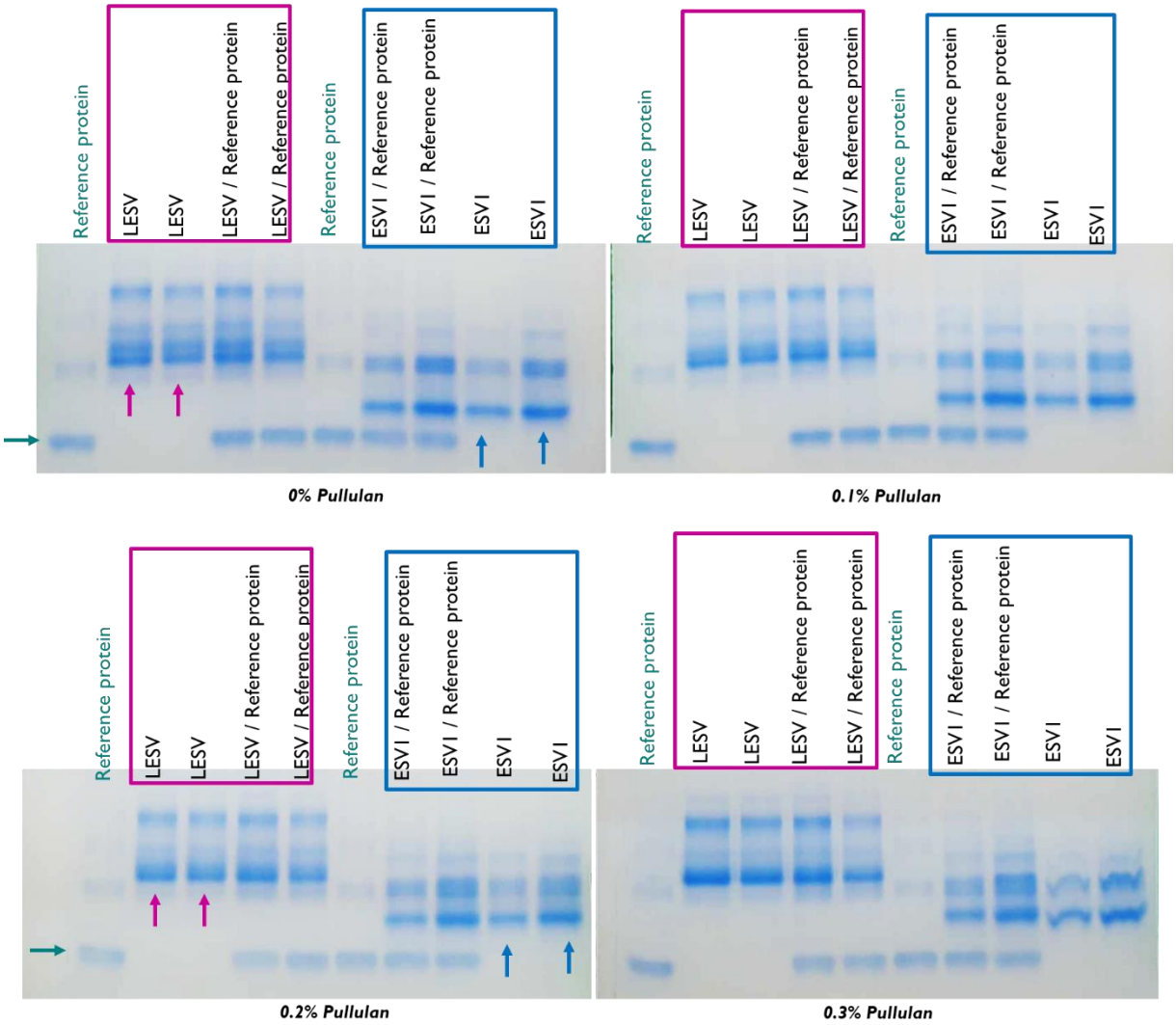


Figure S3: ***Analysis of ESV1 and LESV binding to pullulan by EMSA gel with a concentration of 0%, 0,1%, 0.2% and 0.3% of potential ligands.*** Separation of both proteins using native gels, separation gels were supplemented with 0,1%, 0.2%, 0.3% of A600. Here, 1µg of ESV1 and LESV were loaded per lane. As a further control 1µg of the reference protein was separated on native gels. The gels were stained with bromophenol blue. For each gel, the green arrow indicates the migration of the reference protein. The blue and the pink arrows indicate the migration of ESV1 and LESV proteins respectively.

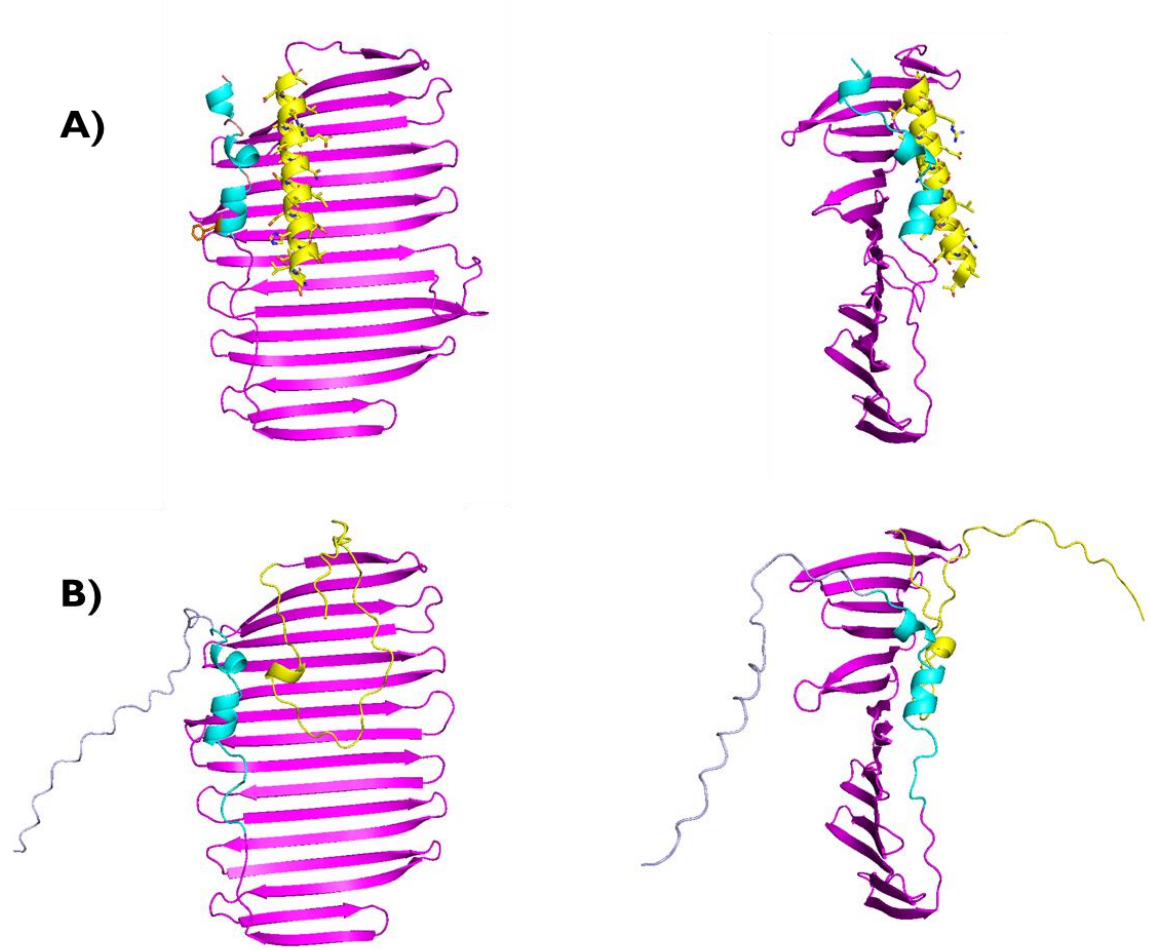


Figure S4: ***Structural conserved motifs on the Face A of LESV (top) and ESV1 (bottom).*** The structures are represented as cartoon, the common β-sheet is colored magenta, the common C-terminal helix is colored cyan and the lon helix of LESV and the long loop of ESV1 are colored yellow. The right panel represents the left panel after a rotation of 90° along y.