Heinrich Heine
University
Düsseldorf

Stochastic modelling of glycogen: structure, metabolism, and related disorders

Inaugural-Dissertation

zur Erlangung des Doktorgrades
der Mathematisch-Naturwissenschaftlichen Fakultät
der Heinrich-Heine-Universität Düsseldorf

vorgelegt von
Yvan Rousset
geboren in
Nîmes, Frankreich

Düsseldorf, April 2023

aus dem Institut für Quantitative und Theoretish Biologie
der Heinrich-Heine-Universität Düsseldorf

Gedruckt mit der Genehmigung der
Mathematisch-Naturwissenschaftlichen Fakultät der
Heinrich-Heine-Universität Düsseldorf

Berichterstatter:

1. Prof. Oliver Ebenhöh

2. Prof. Martin Lercher

Tag der mündlichen Prüfung: 31/08/2023

Abstract

Organisms must efficiently manage their energy resources to survive as nutrient availability can vary significantly over time, and other stresses may temporarily increase energy demands. Therefore, internal energy stores are necessary to respond to changes in energy supply and demand. These stores are filled when nutrients are abundant and depleted when demand exceeds available supply. Glucose plays a central role in energy metabolism for most organisms as it serves as a direct substrate for catabolic pathways. Animals, fungi, and most bacteria store glucose as glycogen, a macro-polymer made of glucose organized in branched linear chains. Cycles of glycogen degradation and breakdown ensure maintaining glucose homeostasis, as well as fueling other organs in mammals. Four enzymes are directly responsible for glycogen synthesis and degradation: glycogen synthase, glycogen branching enzyme, glycogen phosphorylase, and glycogen debranching enzyme. The interplay between these four enzymes ensures the correct building of the glycogen molecule. Despite being widely investigated since 1950, numerous questions remain unclear. The interplay between the kinetics of these enzymes and the structure of glycogen is not fully characterized. The precise mechanism at work during branching and debranching is not well understood. Moreover, the effects of certain genetic conditions on glycogen metabolism and structure is still to be explored.

This thesis introduces a spatially resolved and stochastic model for the synthesis and degradation of glycogen. By using the Gillespie algorithm to track single reaction events, the model allows for a detailed exploration of glycogen structure. Experimental measurements of structural features as signatures of enzyme activities were used to constrain different branching scenarios. The model can also replicate numerous other experimental data such as the density profile and radius of the glycogen granules. Additionally, the model can be used to investigate other effects such as steric hindrance and enzymatic mechanisms, potentially in polysaccharides other than glycogen.

In the second part of this work, we developed algorithmic methods to couple deterministic chemical systems with stochastic ones. We present the periodic-coupling algorithm, which comprises a stochastic module communicating with a classical ordinary differential equation (ODE) solver at a given frequency, enabling the tracking of single stochastic reactions in a regular ODE model. The algorithm outperforms a full stochastic approach and enables the coupling of our 3D structural model to a kinetic model of glycogen metabolism. With this approach, we can simultaneously track the evolution of a small glycogen metabolic model and the glycogen granule properties, which allows for a characterization of the reciprocal effect of the granule structure on the kinetic model. Additionally, it facilitates the investigation of simplified models for glycogen storage diseases that we discuss.

Finally, we provide a discussion on the fractal view of glycogen, as well as a toy model to establish the basics of further investigation of β and α glycogen granule interactions.

Acknowledgements

I would like to express my sincere gratitude to all the individuals who have contributed to the completion of this PhD thesis.

First and foremost, I would like to thank my supervisor, Oliver, for his support, guidance, and freedom he gave me to explore my research interests. I am grateful for the insightful discussions, constructive feedback, and for providing me with an excellent working environment.

I am also indebted to Adélaïde for her valuable feedback and for occasionally pushing me to aim higher.

My sincere thanks go to Nima for his useful discussions and for helping me refine my ideas. I am also grateful to Marvin for programming questions, and to all the members of QTB for the pleasant atmosphere. Especially, I want to thank Hettie, the best colleague with whom one can share an office.

I am indebted to Janina for her expertise in refactoring some code.

Additionally, I would like to thank the people involved in PoLiMeR, including Prof. Barbara Bakker, Prof. Robert Field, Dr. Maaike Oosterveer, Gaia, and Kishore, for their useful discussions on glycogen and GSDs. I am also grateful to Christoff, Kishore, and Karen for their support during my secondment in Groningen. Moreover, I would like to express my appreciation to the other members of PoLiMeR for the cool atmosphere during progress meetings.

Finally, I also want to thanks all the people I met in Düsseldorf for making me feel at home and for making me want to stay in the city.

Contents

Chapter 1

Introduction and Motivation

1.1 Metabolism

Metabolism refers to the set of chemical reactions that insure the sustain of life within organisms. These reactions serve three functions: converting the energy found in nutrients into a usable form for fuelling cellular processes, transforming these nutrients into the building blocks necessary for proteins, lipids, nucleic acids, and carbohydrates. It enables organisms to grow, reproduce, maintain their structure, and respond to their surrounding environments. Metabolism is often used as a way to define what is a living organism [Taylor et al., 2021]. From a metabolic perspective, viruses are not classified as living entities due to their inability to transform energy within themselves.

Metabolic reactions can be classified into two types: catabolic and anabolic reactions. Catabolic reactions involve the breakdown of complex compounds, such as the conversion of glucose to pyruvate during cellular respiration. In contrast, anabolic reactions involve the synthesis or building up of more complex compounds, including proteins, carbohydrates, lipids, and nucleic acids [Berg et al., 2018]. Typically, catabolic reactions release energy while anabolic reactions consume energy. These reactions are catalyzed by enzymes that help them to overcome their activation energy. In anabolism reaction, an energy source is usually required to go against the standard free energy potential, typically provided by ATP, a molecule often described as the energy currency of cells.

These reactions are organized in metabolic pathways which consist in a series of reactions that step by steps transforms one chemical into another. These pathways are highly complex and involve numerous enzymes and proteins. Important ones, among others, are glycolysis, TCA cycle, fatty acid oxidation, gluconeogenesis and glycogenolysis. Glycolysis consists in the conversion of glucose into pyruvate, producing ATP and NADH. The TCA cycle is a sequence of enzymatic reactions that break down acetyl-CoA to produce ATP and NADH, while releasing CO2. Fatty acid oxidation breaks down fatty acids also producing ATP and NADH. Gluconeogenesis consists in producing glucose from non-carbohydrate precursors such as lactate, amino acids, and glycerol. Glycogenolysis, on the opposite produce glucose from carbohydrate sources, such as glycogen. The reverse pathway is called glycogenesis and consists in the building up of the glycogen molecule from glucose.

Between organisms, metabolism vary greatly and depends on numerous factors. In human it is typically influenced by age, activity level, temperature, hormones, diet composition and genetics [Lehninger et al., 2017]. Hormones, such as insulin and glucagon, play a central role, by regulating blood glucose levels or the storage of lipids and their utilization [Cryer, 1993; Gerich, 1993].

One objective of this work is to investigate glycogenolysis and glycogenesis and the effect of hormones and genetic variances on these pathway.

1.2 About Glycogen

The glycogen macro-molecule

Glycogen is a polysaccharide molecule, first observed by Claude Bernard in 1857 [Young, 1957]. It is a very large molecule, composed of thousands of glucose molecules linked together by covalent glycosidic bonds [Drochmans, 1962; Ggunja-Smith et al., 1971]. There are two kinds of glycosidic bonds $\alpha - 1,4$ and $\alpha - 1,6$, where the numbering refers to the positions of the carbon atoms involved in the bond, and α to the stereochemical configuration. These $\alpha - 1,4$ linkage forms linear chains of glucose units constituting glycogen. On these chains, some glucose molecule are linked with $\alpha - 1,6$ linkages, which serves as branching points to another linear chain. These arrangement between chains and branches makes glycogen a highly branched polymer [Lodish et al., 2007].

There is 5 main enzymes involve in its synthesis and breakdown. The synthesis is initiated by Glycogenin (GN). Its role is to form a primary linear glucose chains by catalyzing the transfer of a glucose molecule from UDP-glucose, until the chains reach a degree of polymerisation (DP) around 6 [Roche, 2002]. From these initiated chains, another enzyme will pursue the synthesis by elongating the chains. This enzyme is Glycogen Synthase (GS) and is responsible for the elongation of the chains. It adds a glucose unit at the end of a linear chain (called non-reducing end), using UDP-glucose as substrate. Once the chains become long enough, Glycogen Branching Enzyme (GBE) will be able to cleave one $\alpha - 1,4$ bond and to branch the cleaved part of the chain onto another glucose through an $\alpha - 1,6$ linkage, creating a new chain. The combined action of these two enzymes will ensure glycogen synthesis and gives glycogen some of its structural features, as for example its average chain length, or the frequency in the branching points.

In order for an organism to make use of its stored glycogen, it must first be broken down into glucose unit (glycogenolysis). This process begins with a Glycogen Phosphorylase (GP) breaking down the bonds between individual sugar molecules within each chain, releasing glucose in form of glucose-1-phosphate. When chains have been shorten enough, Glycogen Debranching Enzyme (GDE) will unbranched the short chain, releasing directly one glucose and transferring the remaining part of the chain onto another. The resulting sugar molecules from both reactions are then released into circulation where they can be used for energy

production or used in other metabolic pathways. The mechanisms of these reactions are not elaborated on here. In this work, one of the main focus will be to address this question. We will provide a more comprehensive explanation of these reactions, along with the relevant considerations for their modeling in chapter 2 and 3.

The role of Glycogen

Glycogen is the main energy store in animals, most bacteria and fungi [Berg et al., 2018]. It can be found in both cytoplasm and the nucleus of cells [Stick, 2008]. In mammals It is primarily stored in the liver, muscles and brain where it can be quickly broken down into glucose when needed for energy. The synthesis of glycogen is one of the ways in which organisms store excess energy, in order to use it later. It can be seen as a capacitor in electricity. One of the uses of these is to reduce the fluctuations of an electrical signal, by accumulating excess electrons when there is a high signal and releasing them when the signal is low. This analogy can be extended in a simplified way to the level of glucose in an organism, particularly in humans in terms of glucose levels in the blood. In particular, hepatic glycogen will act as an enormous reserve of glucose and help to regulate the level of glucose in the blood, preventing glucose from accumulating in cytoplasm and increase in osmolarity Meléndez-Hevia et al., 1993. A molecule very similar to glycogen, in terms of its chemical composition, is starch. Both molecules are composed of chains of glucose linked together in $\alpha - 1,4$ and $\alpha - 1,6$ bonds. However, the two molecules differ in their branching frequency, with glycogen having a higher degree of branching than both forms of starch (amylose or amylopectin). Despite their seemingly simple chemical notation, these structural differences have important implications for the characteristics of the two molecules.

The way in which unbranched chains are arranged in starch is crucial for the formation of double helix structures, which are responsible for starch's crystallinity and insolubility. This solid form of storage also affects the kinetics of degradation and glucose release. Conversely, glycogen's highly branched structure makes it soluble and more readily degraded. Therefore, it is essential to comprehend the different branching patterns during the synthesis of these molecules.

β and α glycogen granules

Electron microscopy has identified two types of glycogen structures, known as α and β granules Drochmans, 1962. α granules are predominantly found in the liver, and are aggregate of several β granules. These β granules are single glycogen molecules as we described above. They have a size of about 20-30 nm in diameter and a molecular weight of 10^6-10^7 while α granules are bigger with a diameter of up to 300 nm and a molecular weight of 10^8. It is believed that these aggregates provide a slower energy source [Sullivan et al., 2014], by minimizing their surface to volume ratio. It is a reasonable assumption if we assume that most of the breakdown activity take place at the surface, and that enzymes can not access inner part of glycogen. In the liver, when glycogen concentration is at maximum, glycogen is found in form of β granule, which

later aggregate into α granules via protein bindings [Sullivan et al., 2010]. In section 5 we will discuss in more details these two form of glycogen.

Models for glycogen

The geometry of glycogen and how it occupies 3D space is critical for understanding enzyme activities at the surface and the impact of steric hindrance. Constructing models of glycogen can provide valuable insights into these processes and help researchers gain a better understanding of the underlying mechanisms.

In the 1960s, researchers tried to established a general pattern for glycogen that could explain experimental data. One particularly notable model was the Meyer-Bernfeld model Meyer and Bernfeld, 1940, which can be seen in Figure 1.1-left. The model respects a specific ratio in which the number of A chains divided by the number of B chains is approximately unity [first sources]+ Marshall and Whelan, 1974. A chain is defined Gunja-Smith et al., 1970a as a chain that is only connected to the rest of the molecule through its reducing chain end and does not carry any "daughter" chains. In contrast, B chains carry other A and/or B chains. The Meyer-Bernfeld pattern matches this ratio and has the characteristic that all of its chains terminate at the surface of the molecule, meaning that no chains are buried.

The discovery of a glycogen debranching enzyme from Cytophaga Gunja-Smith et al., 1970c, an isoamylase, allowed for further investigation into the detailed structure of glycogen. This enzyme can hydrolyze $\alpha - 1,6$ bonds only when the branch contains at least three glucose units, leaving chains linked in $\alpha - 1,6$ with a maltosyl chain uncut. When applied to $\phi - \beta$ limit dextrin glycogen (glycogen reduced with phosphorylase and beta-amylase, transforming the accessible A chains into maltosyl groups), the enzyme breaks all $\alpha - 1,6$ bonds between B chains. The resulting product is a mixture of B chains linked to untouched A chains in form of a maltosyl group, and B chains branched to potentially untouched A chain because they are located inside the granule. Further application of β-amylase to the mixture would reduce the unbranched B chains as well as the untouched A chains, yielding maltose units.

The Meyer-Bernfeld pattern is characterized by the absence of buried A chains or B chains carrying only B chains. If we assume such a structure and subject it to an enzymatic process involving (i) creating a phi-beta limit dextrin, (ii) using cytophaga isoamylase, and (iii) reducing the remaining mixture with beta-amylase, no maltose should be released from step (iii). However, results from a study by Gunja-Smith in 1970 show that the degree of beta-amylolysis is approximately 44%, which is roughly the number of maltose molecules that could potentially be released from the total glucose units in the mixture (0.5 x total number of glucose units). The authors concluded that there must be B chains without A chains and/or with buried A chains (not available to phosphorylase in step (i)). Therefore, the proposed Meyer-Bernfeld model cannot be used to accurately describe glycogen molecules.

Whelan [Ggunja-Smith et al., 1971; Gunja-Smith et al., 1970b] proposed a revision of the Meyer-Bernfeld structure drawn in figure 1.1-right. It explains the above observations (A:B=1

and degree of beta-amylolysis = 44%). In this model, which will become known as the Whelan model, 50% of the B chains (the outermost ones) carry an average of two A chains, and the other B chains carry 2 others B chains.

From a schematic view to a widespread idea of a fractal model

The Whelan model has been interpreted more literally over time. In their study of the structure of Maltoheptaose, Goldsmith et al. [Goldsmith et al., 1982] provide valuable insights into the organization of glycogen branches at a molecular level. They discovered the average angles in the alpha1,4 glycosidic bonds responsible for a left-handed helical configuration of glucose residues among the chains. They found that the helix has 6.5 residues per turn, with a rise per residue of 0.24 nm and a cross-sectional area of 1.3 nm^2.

Assuming that glycogen follows the pattern described in the model perfectly, with n regular tiers (layers), one can reasonably state that the number of branches in tier n is 2^{n-1}, and therefore the total number of chains is given by

$$\sum_{i=1}^{t} 2^{i-1} = \frac{1-2^t}{1-2}.$$

From this, one should expect the glycogen content in the two outermost tiers to be around 75%. The authors claim that this is supported by in-vivo observations [Gunja-Smith et al., 1970a]. With enzymatic arguments, the authors further assume that two branches are separated by 4

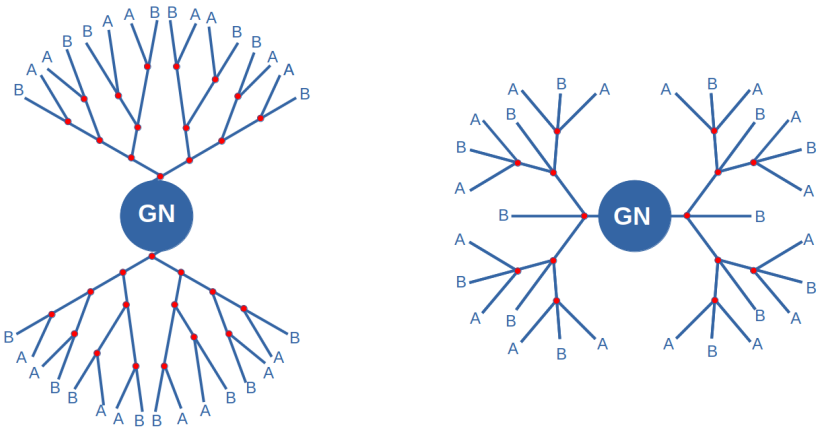

Figure 1.1: **Schematic view of the Meyer-Bernfeld and Whelan models. Left**: Illustration of the Meyer-Bernfeld model. All the A and B chains' non reducing ends are located at the surface. **Right**: The schematic model proposed by Gunja-Smith et al, known as the Whelan model. Figure adapted from Gunja-Smith et al., 1970b

free glucose units. This assumption leads to another: on average, a branch starts halfway up a chain. Therefore, one can calculate the contribution in length of each tier:

$$l = \frac{1}{2} \cdot n \cdot \ell + 0.35$$

where $n = 13$ is the number of glucose units per chain according to the model, $\ell = 0.24\ nm$ is the contribution in length of one glucose unit (rise per residue), and finally, the value of 0.35 comes from the contribution of the $\alpha - 1,6$ bond.

Assuming a molecular weight of $M = 10^7$ for glycogen, we can distribute the 55,500 glucose units in a single molecule along a tiered pattern with chains of 13 units. The total number of chains is given by 2^t where t is the number of tiers. With $t = \ln(\frac{55000}{13})/\ln(2)$, it is found that $t \approx 12$. Therefore, the radius of such a molecule, which is $R = t \cdot l$, is roughly $R \approx 21$ nm.

In a study by Goldsmith et al. [Goldsmith et al., 1982], the authors claimed that an hypothetical 13th tier of glycogen would not be possible due to volume occupancy reaching 1 between the 12th and 13th tiers. The article claim that this mechanism is responsible for the uniformity in size of beta granules, as reported in [Drochmans, 1962].

The concept of glycogen as a fractal structure has been extended by Melendez-Hevia and colleagues [Meléndez et al., 1999; Meléndez-Hevia et al., 1993], who proposed that it arises from an optimality principle. However, more recent research has presented arguments and observations that challenge this perspective [Kim and Duhamel, 2023].

What may have supported this model is a result from polymer physics [Fréchet, 2003] in which certain dendrimer exhibit a density that increase exponentially with the radius, with their external part becoming to dense and providing any further growth.

In Chapter 5 (section 5.1), we explore this issue in greater depth and addressed arguments that oppose this viewpoint.

Zhang et al., 2018 have proposed a Monte Carlo approach to numerically simulate glycogen biosynthesis, in which glucose units are placed on a three-dimensional grid. The biosynthesis of the granule is simulated by adding glucose units on 26 neighbouring positions around the end of the growing chain, leading to the emergence of limited growth as a result. This supports the common idea that steric hindrance restricts granule growth. In order to account for the effects of steric hindrance, the authors set thresholds for the number of grid points that can be occupied locally (equivalent to put a local limit density). If this threshold is reached, some enzymatic reactions cannot take place. This allows for the consideration of an environment which is too dense for an enzyme to access the chains. However, these model parameters are refined empirically in order to reproduce certain observations without predicting any new ones. A second default is the way linear glucose chains are modeled that disregards structural considerations as suggested in [Goldsmith et al., 1982]. Some critical aspects of their models, however, seem to have been corrected after the publication of two erratum.

1.3 Glycogen related diseases

Glycogen storage diseases

Glycogen storage diseases (GSDs) are a group of inherited metabolic disorders that affect the body's ability to store and break down glycogen. In each GSD, one enzyme involved in glycogen synthesis or degradation is either absent or defective, leading to various symptoms in different parts of the body, such as the liver, skeletal muscle, and heart [Hicks et al., 2011]. GSDs affect approximately 1 in 20,000 to 1 in 40,000 individuals worldwide [Ozen, 2007a].

A schematic of the principal reactions involved in glycogen synthesis and degradation is shown in Figure 1.2. Each enzyme and its corresponding GSD are marked in orange.

GSD type 1, also known as Von Gierke disease, is the most common type of GSD affecting approximately 1 in 100,000 individuals worldwide [Hicks et al., 2011]. This disorder is caused by an alteration of glucose-6-phosphatase (G6Pase) or G6P transporter (G6PT), which catalyzes the conversion of glucose-6-phosphate to glucose for export into the blood. In GSD1, glycogen accumulates in liver and muscle cells leading to hepatomegaly, hypoglycemia, hyperlipidemia,

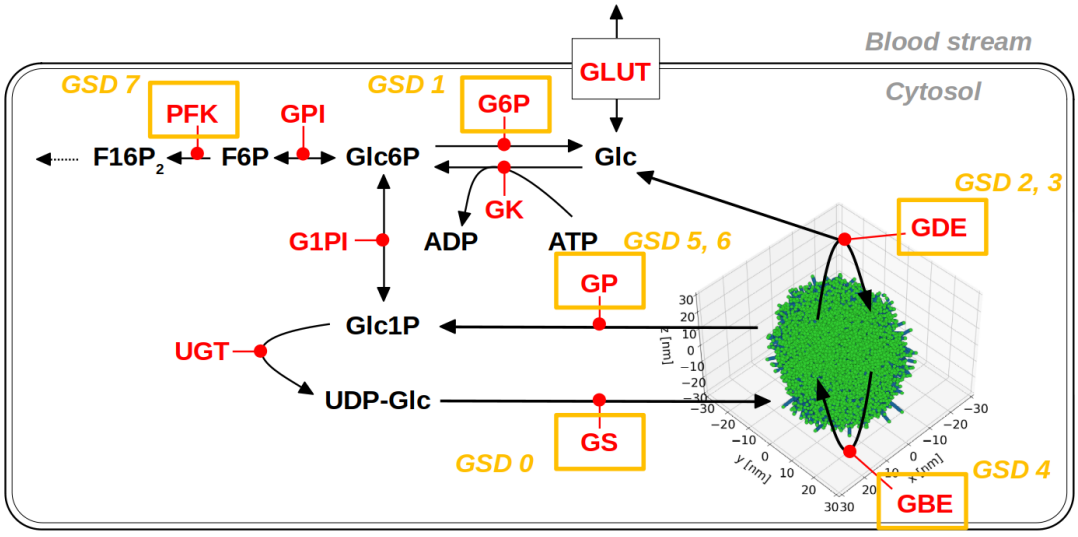

Figure 1.2: **Schematic view of the principal reactions involved in GSDs.** This scheme show the mains metabolites (black) and enzyme (red) involved in some of the GSDs. The corresponding GSDs are highlited in yellow. When two GSDs correspond to one enzyme it means that the same reaction is affected in a different tissues. The Glycogen molecule in shown in green (representing the non-reducing ends). It is a simulated one from our model.

and lactic acidosis [Bali et al., 1993]. The inability to convert Glucose-6-Phosphate into Glucose enhances the two reactions using it as a substrate. The reaction catalyzed by G1PI leads to an increase in Glucose-1-Phosphate, ultimately leading to enhanced glycogen synthesis through GS.

GSD type 0, also known as Glycogen Synthase deficiency, is characterized by hypoglycemia as one of its main symptoms Orho et al., 1998. This is because the impaired function of Glycogen Synthase leads to a reduced synthesis of glycogen, which in turn prevents efficient glucose release during fasting. To use again the analogy with electric capacitor, the blood glucose level is way more subject to variation (here feeding and fasting periods).

GSD type 2 and type 3, also known as Pompe and Cori's disease respectively, are caused by a deficiency in one Glycogen Debranching Enzyme (GDE), either in the cytosol or in the lysosome. It lead to glycogen accumulation and consequently hepatomegaly. The prevalence of Pompe's disease is up to 1 in 13,000 in certain populations [Dasouki et al., 2014; Hicks et al., 2011].

GSD type 4, also known as Andersen disease, is caused by Glycogen Branching Enzyme (GBE) deficiency and also leads to hepatomegaly. It typically form abnormal glycogen structure that resemble amylopectin [Hicks et al., 2011].

GSD type 5 and 6 are associated with Glycogen Phosphorylase (GP) in muscle and liver respectively. GP is the enzyme responsible for breaking down the linear chain. Hypoglycemia and hepatomegaly are observed in GSD6 only.

GSD type 7 affects the conversion of fructosee-1 phosphate into fructose-1,6-biphosphate, one of the essential metabolites in Glycolysis.

It is important to note that this enumeration of GSDs is not exhaustive. In Chapter 4, we will address some of these diseases in the context of our model.

Other glycogen related diseases

One other disease directly link to glycogen is Lafora. Lafora disease is a rare, inherited disorder that affects the nervous system. It is a progressive, neurodegenerative disorder characterized by the accumulation of abnormal glycogen-like structures in cells of the central and peripheral nervous systems, resulting in epilepsy, cognitive decline and other neurological symptoms Tagliabracci et al., 2008.

Adult Polyglucosan Body Disease (APBD) is an inherited, progressive neurological disorder that affects the peripheral and central nervous systems [Suzuki et al., 1971]. It is caused by mutations in the GBE1 gene, resulting in an accumulation of abnormal glycogen in neurons, called polyglucosan bodies. APBD is a rare condition that affects fewer than one in every million people worldwide. In our study, we will associate APBD with GSD type 4, since our model will not allow to distinguish between them.

Finally, the most known disease related to glycogen is Diabetes. It is metabolic disorder characterized by high levels of blood glucose. It is caused by genetic and environmental factors, and is often associated with a various of health complications Federation, 2011. In type 1 diabetes there is an impaired ability to regulate blood glucose levels due to either inadequate production or utilization of insulin. This results in hyperglycemia. To characterize the mechanisms at work in diabetes, it is important, among other things, to be able to understand how different metabolic conditions affects blood glucose levels. Interestingly, diabetic patients exhibit lower formation of α granule in the liver [Li and Hu, 2020], which will be discussed in 5.

1.4 Research Questions

Glycogen possesses certain structural characteristics that can only be fully understood through complex models. Variations in the molecule's branching patterns, chain length distribution (CLD), and chain A to chain B proportions differ among organisms. While some features of glycogen structure, such as CLD, can be approximated through simplified kinetic models, others cannot. Wu and Gilbert, 2010 have proposed a reductionist approach that utilizes a set of equations to model starch, which is closely related to glycogen. This approach yields promising results and allows for extensive comparison to experimental data. Ordinary Differential Equation-based kinetic models are computationally less demanding than stochastic simulations and enable the simulation of chain length distribution dynamics for simple polymeric systems.

However, a detailed 3D description of glycogen granule is required to access complex branching patterns, distinguishing between A and B chain types as well as considering volume effect. The precise location of branches and the connectivity information is essential to access macroscopic structural quantities such as the granule radius and density. It further allows to consider steric hindrance effects, where enzymes may be prevented from reacting due to local high glycogen density. Chapter 3 presents a model of glycogen that accounts for the 3-dimensional structure. This model is dynamic in the sense that a glycogen granule evolves according to the enzymes in presence, their mechanisms and kinetic activities, exhibiting either a synthesis phase or a degradation phase. We show how this detailed 3D description can provide insight into complex branching mechanisms that are not captured by simple kinetic models.

If having a model of glycogen synthesis and degradation allows us to access the structural characteristics of glycogen as explained above, we can also approach the problem in the opposite way. Starting from structural data and trying to understand what mechanisms are at work in the enzymes to have allowed such a structure. We will see how to use the model in this direction (chapter 3), and in particular how to constrain different possible scenarios on the branching enzyme, and discuss them in the context of the current literature.

One of the common threads of this work has been to consider the reciprocal effects between the structure of glycogen and the kinetics of chemical reactions associated with its synthesis and degradation. Most of the diseases related to glycogen caused, or are caused by structural

deficiencies of glycogen. To study them requires a model coupling reaction kinetics and glycogen structures.

A traditional approach to studying glycogen metabolism would involves using ordinary differential equations (ODE) to model the reactions. However, this approach oversimplifies glycogen as a simple concentration of glucose fixed into glycogen. In reality, glycogen can have different structures that affect its ability to release glucose. For example, a branched glycogen structure may be more efficient in releasing glucose than a linear structure. Certain genetic conditions can also affect glycogen structure and metabolism. To better understand these effects, our detailed structural description of glycogen is required.

In chapter 4 we developed a reduced glycogen metabolism model that accounts for hormonal regulation. It leads us to develop an hybrid algorithm in which stochastic reactions are incorporated in a regular kinetic model. By combining this approach with our structural glycogen model, we have established a framework to investigate the impact of various genetic conditions on both glycogen metabolism and its structure. Our approach has the unique capability of quantifying the role of debranching in glucose homeostasis, as debranching bypass the regular pathway from glucose-1-phosphate to glucose by releasing glucose directly.

In chapter 5, a computational model has been proposed to investigate the aggregation behavior of glycogen granules. Specifically, the model explores the impact of the surface to volume ratio of glycogen on the size and kinetics of granule aggregation.

Chapter 2

Methods

The aim of this chapter is to provide an overview of the key elements that are applicable to all following chapters. In cases where a chapter presents methods that are specific to that particular chapter, the necessary information will be provided within that chapter. Certain elements are shared with the method section of the article featured in chapter 3. In this chapter, a more comprehensive description of the methods is provided, including potential variations.

2.1 A complete description of the model

2.1.1 From self avoiding random-walk to a coarse-grain model for glycogen

One of the main objectives of this study is to explore the structural properties of glycogen. In order to accomplish this, a structural model is required that accurately describes the arrangement of glycogen granules in three-dimensional space. While the models described in the introduction have been proposed to illustrate the branching patterns of glycogen, very few have focused on the organization of glycogen in 3D. In polymer physics, the most simple way to describe a linear polymer is a random walk. The steps ℓ of the random-walk accounting for the distance between two monomers. The linear polymer is therefore modeled by generating the position of the next monomer from the position of the previous one. This is done by generating randomly, with equal probability, a direction in space $\overrightarrow{u}$ such that

$$\overrightarrow{r_{n+1}} = \overrightarrow{r_n} + \overrightarrow{u}$$

with $|\overrightarrow{u}| = \ell$. In order to make such a model more realistic and account for steric-hindrance, one can model the monomer as a sphere of radius ρ. When generating the next position $\overrightarrow{r_{n+1}}$, it has to be ensured that no other monomer occupies the space. This is done by rejecting the new position if one of the distances

$$\forall i \in 1, ..., N, \quad ||\overrightarrow{r_{n+1}} - \overrightarrow{r_i}|| < 2 \cdot \rho,$$

with $||\cdot||$ being the Euclidean distance defined as :

$$||\overrightarrow{r_{n+1}} - \overrightarrow{r_i}|| = \sqrt{(x_{n+1} - x_i)^2 + (y_{n+1} - y_i)^2 + (y_{n+1} - y_i)^2} \tag{2.1}$$

Figure 2.1 depicted when a position is rejected. The green monomer is the generated monomers indexed by $n+1$. On the left side, the monomer indexed by i_0 is overlapping with the proposed position because the distance $D = ||\overrightarrow{r_{n+1}} - \overrightarrow{r_{i_0}}||$ in red is lower that 2ρ, therefore the location is rejected. On the right side, the distance D is greater than 2ρ implying no overlapping of the monomers.

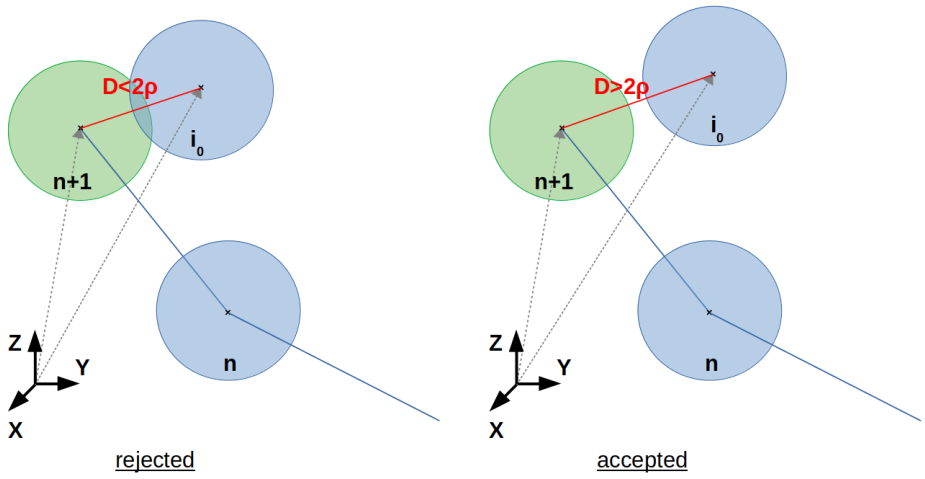

Figure 2.1: **Schematic of the distance criteria in a self-avoiding random walk.** The blue disks represent the previously generated monomers. A new potential monomer (green disk) is generated at a distance of ℓ from the previous monomer n. **Left:** The generated monomer (green) overlap with another monomer indexed by i_0, therefore the generated location is rejected. **Right:** The generated monomer (green) does not overlap with another monomer any of the monomer indexed by $i \in 1, ..., n$, therefore the generated location is accepted

It is inspired by this approach that we designed our numerical procedure to generate a 3D glycogen structure. As previously introduced, glycogen is not a linear polymer but is instead highly branched. This type of polymer is called a dendrimer. Firstly, we will explain how we model linear chains, and secondly, we will describe how we model branching points.

We assume that glycogen is made only of glucose units linked together through α-1,4 and α-1,6 glycosidic bonds. The former will be the linkage inside linear chains, while the latter will form branching points. We assume that these linear chains are always in the form of a single helix, for which many structural parameters have been measured experimentally [Goldsmith et al., 1982]. This article reports that the cross-section of the helix is $1.3nm^2$. The spatial period of the helix, also called a "turn," has a length of 6 to 7 residues. They also determined the rise per residue, which we will refer to as ℓ, to be $0.24nm$. It corresponds to the contribution of one glucose unit to the length of the chains (see Fig 2.2). Finally, the van der Waals radius is found to be 0.65 nm, and we will refer to it as ρ throughout this manuscript. To account for these structural specificities, we will describe the glucose units as spheres with radii equal to

the van der Waals radius. This ensures that we spatially account for the width of the helices. Each of these spheres will be spaced by $\ell = 0.24\ nm$ to account for the contribution, in length, of a monomer to the chain. Unlike in a random walk, we will place these spheres on a straight line representing the direction of the helix.

It can be noticed that the radius of our spherical monomers is greater than the distance between two consecutive monomers. In order to accommodate this within the self-avoiding approach, we allow monomers on the same branch to overlap (refer to Figure 2.2).

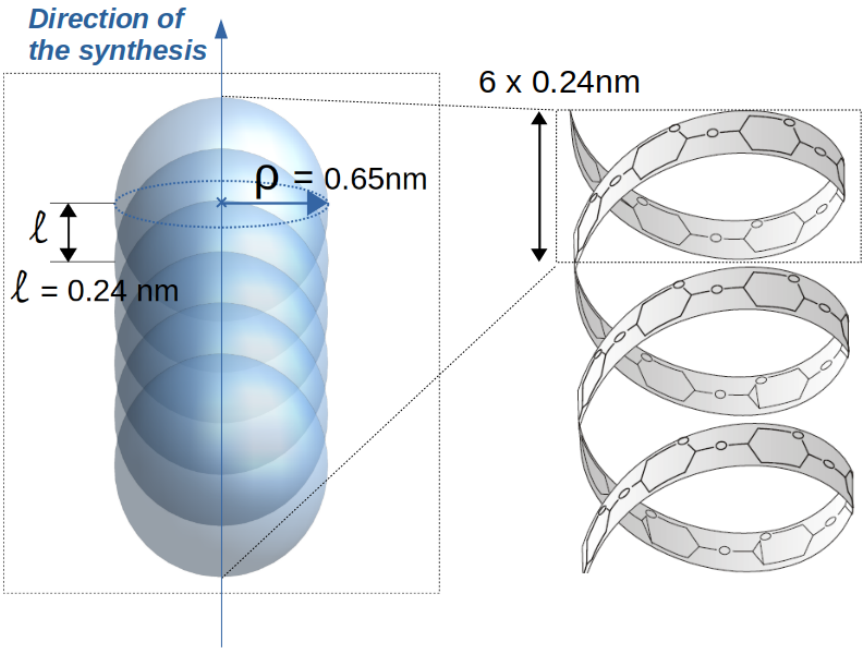

Figure 2.2: **Schematic of the monomer description in linear chains.** Each glucose in an helical linear chains contributes to 0.24 nm. The helix has a radius of 0.65 nm. Inside a linear chains, the models stacks overlapping spheres with radius 0.65 nm, where each sphere is distant from its neighbors by 0.24 nm

With this approach, we create a representation that is simple to computationally implement yet still accurately reflects the spatial properties of helical conformation of linear chains. To properly model branches formed by alpha-1,6 linkages, additional geometric considerations must be taken into account. As shown in Figure 2.3 (left panel, branching A), a new branch is generated from the monomer on the main chain to which it is attached and by two angles (ϕ and ψ) indicating the direction of the new chain. Additionally, the monomer anchoring the branch to the main chain and the first monomer of the branch chain are separated by a distance greater than 2ρ to prevent overlap.

This method, which involves generating a new direction directly from the monomer of the mother chain, is simple and will be the default method used in this manuscript, unless otherwise

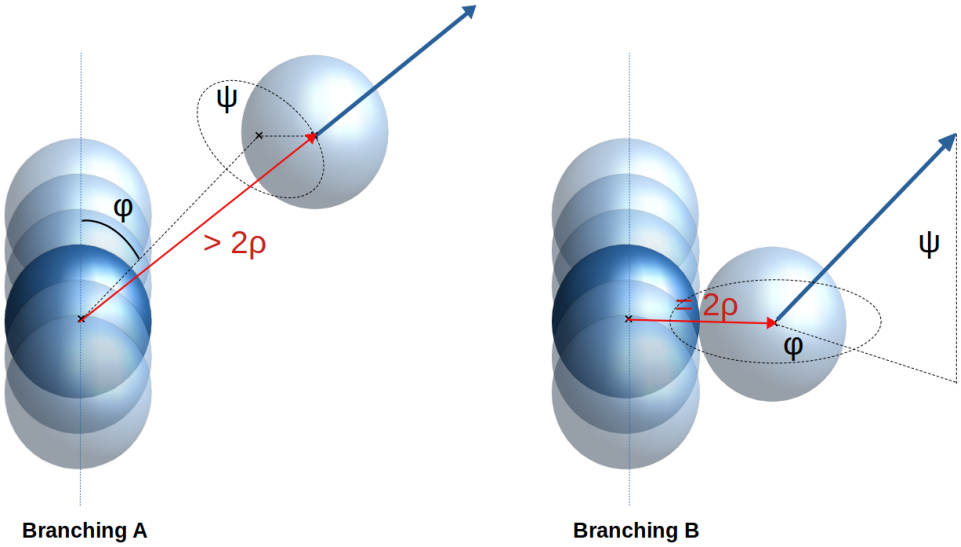

Figure 2.3: **Schematic of two models for branching. Left:** Model referred to as "Branching A". In this model, two angles ϕ and ψ are randomly picked from a uniform distribution to provide a direction from the monomer on the "mother chain". This model is used by default in this manuscript. Perpendicular directions are favored because the first monomers can potentially overlap with the "mother chain" if the generated direction is too parallel to the "parent" direction. **Right:** Model referred to as "Branching B". An alternative way to perform branching. In this model, we first randomly generate a vector perpendicular to the "mother chain" (red). From this direction, we pick ϕ and ψ from a uniform distribution to provide the direction of the new branch. It allows for more possible new directions.

specified. Its main drawback is the number of rejections it causes when attempting to simulate branching. When a branching attempt occurs, if the generated direction is too close to the mother direction (blue dotted line), the first daughter monomers will overlap with the mother chain, which is not allowed. This leads to the rejection of the proposal and increases calculation time. Furthermore, the accepted chains will tend to be rather perpendicular to the parent direction, which can potentially decrease molecular density locally.

Possible improvements

To overcome this, another possible branching model is described in Figure 2.3 (right panel, Branching B). In this case, a vector is first generated perpendicular to the parent chain, and then from the first "daughter" monomer, the direction of the chain is generated. Practically, we first generate a random vector from a uniform distribution. Then, we subtract its component co-linear to the mother direction and normalize it.

if $\overrightarrow{r}$ is vector generated randomly, it can be written as follow:

$$\overrightarrow{r} = a \cdot \overrightarrow{u} + \overrightarrow{v}$$

where $a \in \mathbb{R}$, $\overrightarrow{u}$ is a direction vector of the mother chain, and $\overrightarrow{v}$ is the perpendicular part to it. We can then writes:

$$a = \frac{\overrightarrow{r} \cdot \overrightarrow{u}}{||\overrightarrow{u}||^2} \tag{2.2}$$

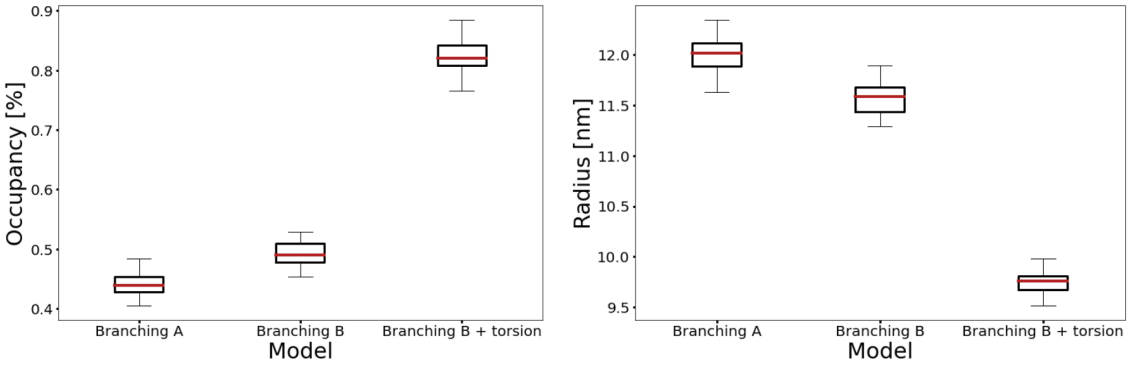

Figure 2.4: **Comparison of branching models.** Boxplot of 30 simulations per model. All simulations are stopped when $N = 10,000$ monomers have been synthesized. **Left:** Occupancy of the different models. While the second way to branch does not significantly increase the granule density, the addition of possible torsion to the chain does. **Right:** Radius of the granule for each model. The radius of the granule decreases with the introduction of possible torsion.

To further develop this coarse-grained branching approach, one could think of allowing flexibility to the chains. Linear polymers have a persistence length, a concept introduced by de Gennes [Gennes, 1979]. It measures the polymer stiffness by calculating how much the direction, defined by a pair of monomers, is correlated over a certain distance. Therefore, incorporating the ability for chains to bend may prove useful in future work. As a proof of concept, we tested the case where, if volume exclusion rejects the addition of one glucose unit, the chain will find the closest new location for this unit, minimizing the quadratic sum of the angles between each monomer constituting the chain. This has the effect of spreading the total torsion on each consecutive monomer. The occupancy during the synthesis is now way higher than in the first branching model (Figure 2.4). As it is hard to visualize in 3D, figure 2.5 shows the result in a 2D version of the model.

2.1.2 The enzyme mechanistic

In the previous subsection, we presented how we modeled the 3D organization of glucose molecules. In this subsection, we will explain how we modeled the "mechanical" action of

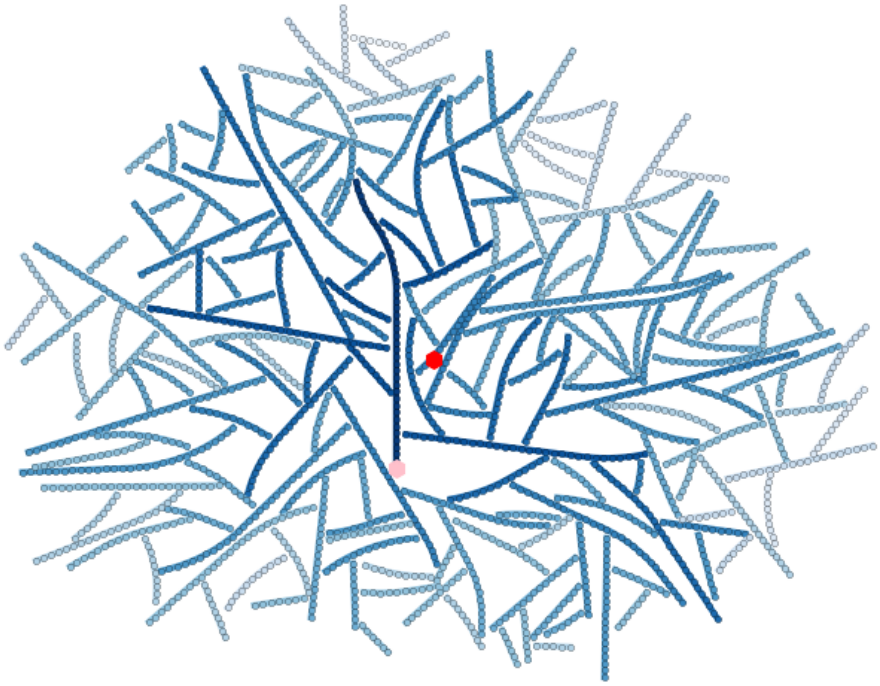

Figure 2.5: **Example of the 2D variation of the model with flexible chains.** The pink marker shows the starting position. The red marker shows the center of mass. The blue gradient shows the connectivity distance from the initial chain.

enzymes, which we will refer to interchangeably as "enzyme mechanistic" or "reaction mechanistic". In this general term, we include everything that pertains to both the characteristics and specificity of the substrate (in this case, glucose chains), as well as the specificity of the enzymes involved, particularly the branching enzyme (GBE) and the disconnection enzyme (GDE).When we refer to substrate specificity, we are primarily referring to the lengths of the glucose chains involved in the reactions, as measured by their degree of polymerization (DP).

Elongation reaction catalyzed by Glycogen Synthase (GS)

This reaction is one of the simplest to parameterize, as only one length is involved, which we refer to as ${L_{\mathrm{GS}}}^{\mathrm{min}}$. This length corresponds to the minimal degree of polymerization required for the elongation reaction to take place. By default, we set this length to 4 glucose units, as it has been observed biochemically that the enzyme cannot react on shorter chains [Berg et al., 2018].

$$ {L_{\mathrm{GS}}}^{\mathrm{min}} = 4 $$

The reaction of a chain with degree of polymerisation of n glucose (DP_n) writes:

$$ \mathrm{DP}_n + \mathrm{UDP{-}Glc} \xrightarrow{\mathrm{GS}} \mathrm{DP}_{n+1} + \mathrm{UDP}, \quad \text{with} \quad n \geq {L_{\mathrm{GS}}}^{\mathrm{min}} \tag{2.3} $$

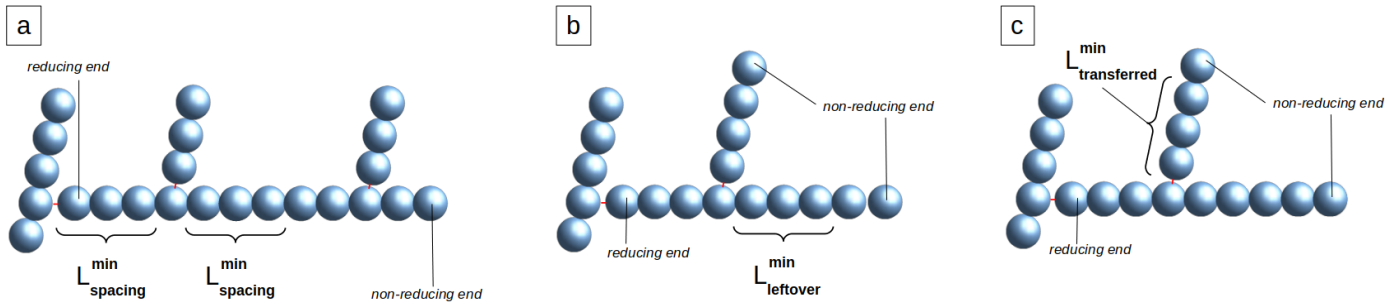

Figure 2.6: **Schematic explaining $L^{\mathbf{GBE}}_{\mathbf{spacing}}$, $L^{\mathbf{GBE}}_{\mathbf{leftover}}$ and $L^{\mathbf{GBE}}_{\mathbf{transferred}}$.**

Reduction reaction catalyzed by Glycogen Phosphorylase (GP)

As for GS, here again only one length is necessary to characterize the reduction of a chain by GP. We will call it ${L_{\mathrm{GP}}}^{\min}$. Unless otherwise stated, we will fix this length in relation to ${L_{\mathrm{GS}}}^{\min}$, so that

$$ {L_{\mathrm{GP}}}^{\min} = {L_{\mathrm{GS}}}^{\min} + 1. $$

It prevents too much reduction (e.g. DP $= 3$, with ${L_{\mathrm{GS}}}^{\min} = 4$) where the chain would becomes chemically "inert", blocking the synthesis of the molecule. This assumption can however be relaxed if eventually, the disconnecting enzyme can compensate and use a chain shorter than ${L_{\mathrm{GS}}}^{\min}$. The reaction of a chain with degree of polymerisation of n glucose (DP_n) writes:

$$ \mathrm{DP}_n + \mathrm{P} \xrightarrow{\mathrm{GP}} \mathrm{DP}_{n-1} + \mathrm{Glc{-}1\text{-}phosphate}, \quad \text{with} \quad n \geq {L_{\mathrm{GP}}}^{\min} \tag{2.4} $$

Glycogen Branching Enzyme (GBE) catalyzed reduction reaction

This reaction is the most complex to model because it involves different chain lengths. The enzyme will first cleave a chain and branch the "cut" part of the chain into $\alpha - 1, 6$. We always model this reaction in an intra-molecular way, to simplify our study. To describe the specific length chains involved in the reaction, we introduce 3 lengths:

$$ L^{\mathrm{GBE}}_{\mathrm{spacing}}, L^{\mathrm{GBE}}_{\mathrm{leftover}}, L^{\mathrm{GBE}}_{\mathrm{transferred}} $$

$L^{\mathrm{GBE}}_{\mathrm{spacing}}$ is the minimum distance in glucose residue that can separate two branches, or that can separate a branch from the the reducing end (Fig 2.6.a. $L^{\mathrm{GBE}}_{\mathrm{leftover}}$ is the minimum distance in glucose residue that can remain on the chain after a new branch has been added. In other terms it is the minimal chain length between the last branch and the non-reducing end (Fig 2.6.b). Finally $L^{\mathrm{GBE}}_{\mathrm{Transferred}}$ is the minimal length that can be transferred (Fig 2.6.c). To fulfil these 3 criteria, a chain has to be strictly greater than the sum of these 3 minimal length in order to be considered as substrate for the branching:

$$ L^{\min}_{\mathrm{GBE}} = L^{\mathrm{GBE}}_{\mathrm{spacing}} + L^{\mathrm{GBE}}_{\mathrm{leftover}} + L^{\mathrm{GBE}}_{\mathrm{transferred}} \tag{2.5} $$

The "strictly greater" come from the glucose residue that is attached to the new branch, which is not considered in any of these length in our description. Fig 2.7 shows an example of the possible outcomes when we parameterized these lengths with $L^{\text{GBE}}_{\text{spacing}} = 2, L^{\text{GBE}}_{\text{leftover}} = 2, L^{\text{GBE}}_{\text{transferred}} = 2$. For a DP of 7 glucose residues there is only one possible outcome, while for DP9 we get 6 possibilities.

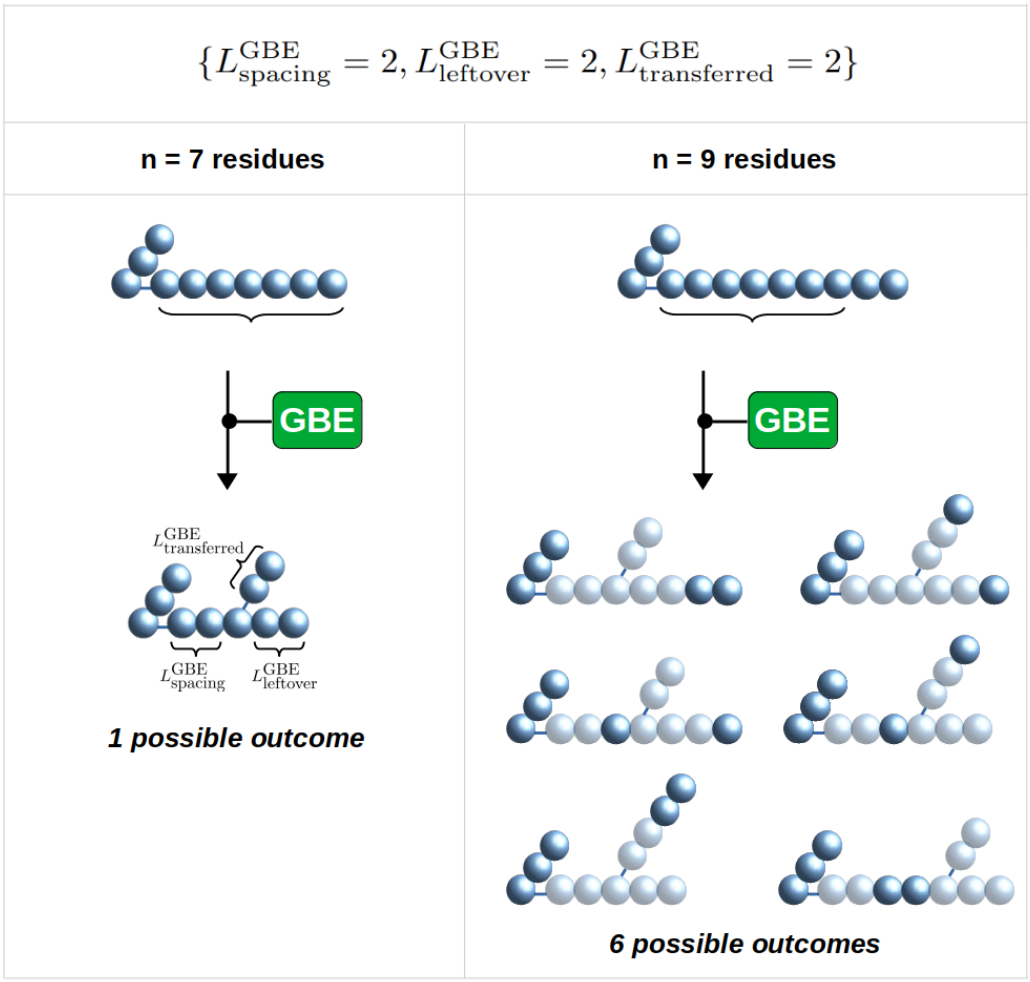

Figure 2.7: **Illustration of the potential outcomes by GBE branching with $\{L^{\text{GBE}}_{\text{spacing}} = 2, L^{\text{GBE}}_{\text{transferred}} = 2, L^{\text{GBE}}_{\text{leftover}} = 2\}$.** With these minimal lengths, the minimal DP required for a branching to occur is DP = 7. If the chain length is longer, the number of possible outcomes increases. **Left:** With a substrate of DP = 7, only one outcome is possible. **Right:** With a substrate of DP = 9, up to 6 distinct outcomes are possible.

These different possibilities are described here in a qualitative way in the sense that we do not describe the possibility of variability in the probabilities of these outcomes. In the chapter of

this thesis, we test two different ways of choosing among these different outcomes. The first one, called "the flexible location model", consists in not privileging any configuration. Thus, all cleavage and branching points allowed by $L^{\text{GBE}}_{\text{spacing}}, L^{\text{GBE}}_{\text{leftover}}, L^{\text{GBE}}_{\text{transferred}}$ are equiprobable. A second model, named "strict location model" is also tested. In this one, the cleavage point is also selected according to a uniform probability distribution, while the branching point is always the same and is located closest to the non-reducing end. Supported by the results of this paper, and unless otherwise stated, we will keep the "flexible location model" as the default model.

The reaction of a chain with degree of polymerisation of n glucose (DP_n) writes:

$$\text{DP}_n \xrightarrow{\text{GBE}} \text{DP}_{n-m} + \text{DP}_m, \quad \text{with} \quad n > L^{\text{min}}_{\text{GBE}} \quad \text{and} \quad m \geq L^{\text{GBE}}_{\text{transferred}} \tag{2.6}$$

Glycogen Debranching Enzyme (GDE)

This reaction, like the one catalised by GBE, takes place in two steps. The substrate chain has to be an A chain, which means that it does not contain any branch. The length criteria here, ${L_{\text{GDE}}}^{\text{max}}$ is a maximal DP instead of a minimal, as it is used in the other reactions. When an A chain becomes short enough ($\text{DP}(n)$, $n \leq {L_{\text{GDE}}}^{\text{max}}$), A chain of $\text{DP}(n-1)$ will be transferred to the so-called "mother chain". The remaining glucose will be hydrolysed an release in the system. Overall we have:

$$\text{DP}_n + \text{DP}_m \xrightarrow{\text{GDE}_{\text{step 1}}} \text{DP}_1 + \text{DP}_{m+n-1} \tag{2.7}$$

$$\text{DP}_1 + \text{DP}_{m+n-1} \xrightarrow{\text{GDE}_{\text{step 2}}} \text{DP}_{m+n-1} + \text{Glc} \tag{2.8}$$

2.1.3 Stochastic approaches to model the dynamics

In the previous subsection we have introduced how we model the 3-dimensional organisation of glucose in a glycogen granule as well as how the different enzymes act and perform reaction on it. We will now explain how we model the evolution of glycogen structure dynamically, in relation to the different enzymes present in the system. We have to know which single reaction occurs and when, in order to make the corresponding changes to the granule structure, reaction by reaction. To do this, we model the dynamics of the system using a Gillespie algorithm [Gillespie, 1976]. Moreover, the number of enzymes involved in glycogen synthesis and degradation is small compared to the number of glucose contained in a granule. As such, temporal fluctuations can have a significant impact that is difficult or even impossible to characterize using deterministic methods.

Gillespie

Initially introduced by Daniel T. Gillespie [Gillespie, 1976, 1977a] with the purpose of numerically simulating the temporal evolution of a chemical reaction system for gases, this algorithm

is more powerful than a simple Monte-Carlo approach. Indeed, it follows from rigorous probabilistic considerations that not only allow it to provide an exact solution to the system, but also to take into account the "true" fluctuations of the system. This subsection will consist in explaining and discussing the foundations of this algorithm, which will first be used to explain how our simulations work, and secondly, to set the basis for a hybrid algorithm that is both stochastic and deterministic that we have developed and that we will describe later.

Gillespie's algorithm follows from the following assumptions: the system is spatially homogeneous and in thermal equilibrium. Let us consider a system composed of N species involved in M chemical reactions R_μ ($\mu = 1, \cdots, M$), we will seek to estimate the probability $\mathcal{P}_\mu$ that reaction $\mathcal{R}_\mu$ occurs in a time interval Δt:

$$\begin{aligned} \mathcal{P}(\tau, \mu)\ d\tau \ &= \ \text{probability that } R_\mu \text{ will happens in the next infinitesimal} \\ &\quad \text{time interval}(t+\tau, t+\tau+d\tau) \end{aligned} \tag{2.9}$$

This quickly leads the author to rewrite the equation in the form:

$$\mathcal{P}(\tau, \mu)\ d\tau \ = \ \mathcal{P}_0(\tau) a_\mu d\tau \tag{2.10}$$

where $\mathcal{P}_0(\tau)$ is the probability that not a single event occurs during time τ and $a_\mu d\tau$ the probability that reaction $\mathcal{R}_\mu$ occurs during the next interval time $d\tau$. After carefull derivation, $\mathcal{P}_0(\tau)$ can be put in the form

$$\mathcal{P}_0(\tau) \ = \ \exp\left(-\sum_{\nu=1}^{M} a_\nu \tau\right) \tag{2.11}$$

Inserting 2.11 into 2.9 give us

$$\mathcal{P}(\tau, \mu) \ = \ a_\nu \exp\left(-\sum_{\nu=1}^{M} a_\nu \tau\right) \tag{2.12}$$

Now that we have an expression for $\mathcal{P}(\tau, \mu)$, the author describes two methods for generating events according to this probability distribution. Here we will briefly explain the so-called "The direct method" which is used in the context of our study.

The method consists of rewriting equation 2.12 using conditional probabilities. $\mathcal{P}(\tau, \mu)$ can be decomposed as follows:

$$\mathcal{P}(\tau, \mu) \ = \ \mathcal{P}(\tau) \cdot \mathcal{P}(\mu|\tau) \tag{2.13}$$

Where $\mathcal{P}(\tau)$ is the probability that a reaction takes place between times $t+\tau$ and $t+\tau+d\tau$, whatever it may be, and $\mathcal{P}(\mu|\tau)$ is the probability that the next reaction is $\mathcal{R}_\mu$ given that the

reaction takes place at time $t + \tau$. The probability $\mathcal{P}(\tau)$ is nothing more than the sum of the probabilities of each of the reactions:

$$\mathcal{P}(\tau) = \sum_{\mu=1}^{M} \mathcal{P}(\mu, \tau) \tag{2.14}$$

By substituting 2.14 into 2.13, the following is obtained:

$$\mathcal{P}(\mu|\tau) = \frac{\mathcal{P}(\mu, \tau)}{\sum_{\nu=1}^{M} \mathcal{P}(\nu, \tau)} \tag{2.15}$$

We now set $a = \sum_{\mu=1}^{M} a_\mu$ and using Equation 2.13, we have the following:

$$\mathcal{P}(\tau) = a \cdot \exp(-a\tau) \tag{2.16}$$
$$\mathcal{P}(\mu|\tau) = a_\mu / a \tag{2.17}$$

Where $\mathcal{P}(\tau)d\tau$ is the probability of a reaction occurring between $t + \tau$ and $t + \tau + d\tau$, it can be written that:

$$F(t) = \int_0^t \mathcal{P}(\tau) \tag{2.18}$$

as the probability distribution function. It will give us the probability that the event "reaction" has occurred at a time $t' < t$. To generate a random time in accordance with this distribution we will use what the author calls "the inversion generating method". This amounts to picking r according to a uniform law between 0 and 1 such that $t = F^{-1}(r)$ (see Fig 2.8.a).

with:

$$\mathcal{P}(t) = a \cdot \exp(-at) \tag{2.19}$$
$$\Rightarrow F(t) = 1 - \exp(-at) \tag{2.20}$$

with $t = F^{-1}(r)$ and the fact that picking r or $1 - r$ from a uniform distribution is equivalent, r writes:

$$r = 1 - \exp(-a \cdot t) \tag{2.21}$$
$$t = \frac{1}{a} \ln(1/r)$$

Thus, by generating a random number $r_1 \in [0, 1]$, we can generate the time interval Δt that had to be waited for the next reaction to occur.

Once this is done, it is now necessary to choose which reaction has taken place. The probability $\mathcal{P}(\mu|\Delta t)$ is nothing more than the ratio between the probability a_μ that the reaction $\mathcal{R}_\mu$ takes

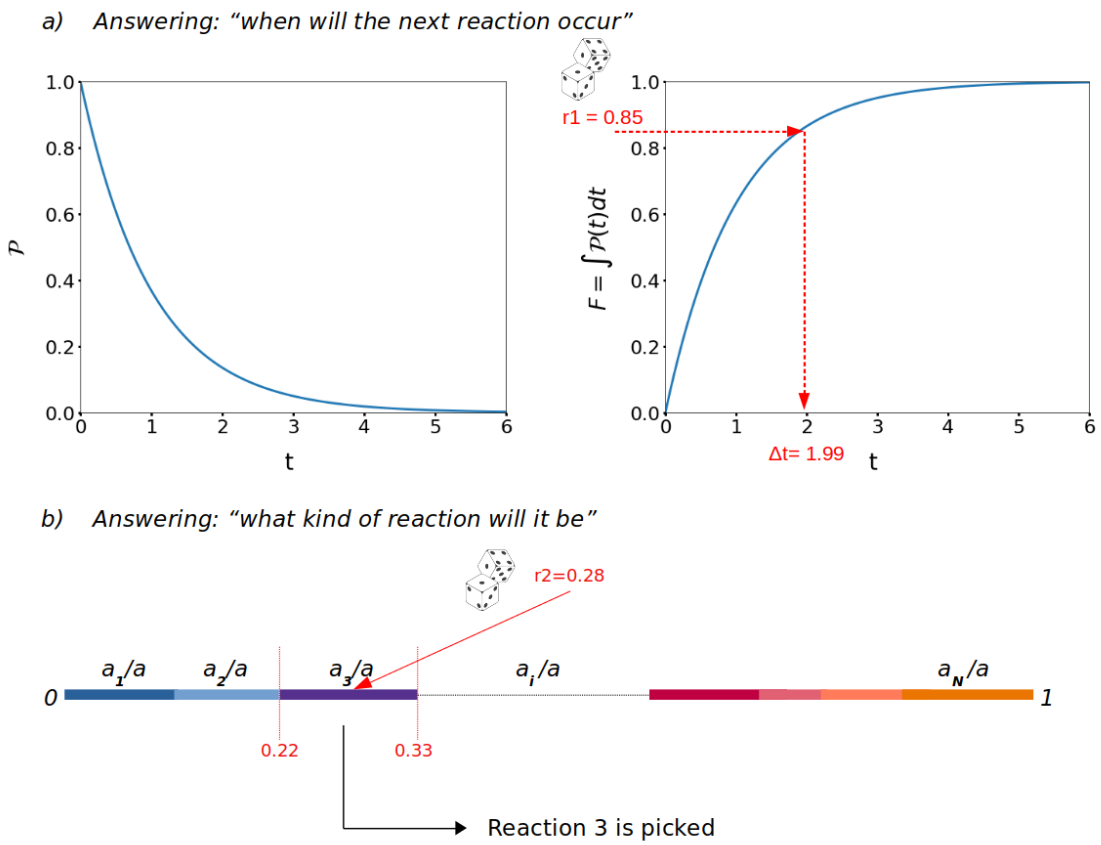

Figure 2.8: **Schematic view of the random numbers involved in the Gillespie direct method. a:** The probability $\mathcal{P}$ of a reaction to occurs and the corresponding probability distribution function. Picking a random number between 0 and 1 allows to returns the elapsed time. **b:** The reaction is selected by drawing a random number between 0 and 1, and looking on which segment, defined by the normalized propensities, this random number falls.

place in the next infinitesimal interval dt, and the sum of all probabilities of seeing a reaction take place in the interval dt, we have:

$$\mathcal{P}(\mu|\Delta t) = a_\mu/a \tag{2.22}$$

Thus, we only need to draw a second random number $r_2 \in [0,1]$ and look in which interval $i \in [1,M]$ it falls such that:

$$\sum_{j=1}^{i} a_j \leq r_2 < \sum_{j=1}^{i+1} a_j \ ,$$

as illustrated in Fig. 2.8.b.

2.1.4 Equivalence between deterministic rate-laws and stochastic propensities

The strength of Gillespie's approach lies primarily in the evaluation of stochastic reaction rates, which are meant to reflect the actual course of events occurring in a system. While it is possible to access these reaction rates in the case of "simple" chemical kinetics (i.e. Mass-action kinetics) as shown cleverly in the original paper from Gillespie [Gillespie, 1976], it becomes rapidly really difficult for more complex kinetics. For gases in thermal equilibrium, this amounts to estimating the probability that reactants collide within the infinitesimal time interval dt, within a volume V. Considerable efforts have been produced in applying Gillespie type algorithms to enzymatic reactions kinetics such as Michaelis-Menten [Petzold, 2011].

The rates laws used in ODEs systems often reflect complex kinetics. We can always use the deterministic reaction rate and convert it into a stochastic rate, with the only difference being that the populations will be a discrete number of molecules rather than a concentration. In doing so, we do not gain precision as would have been the case with a stochastic description of all elementary reactions, but we have an effective way of reproducing the right kinetics in a stochastic manner.

To do so, we take the flux v_i of a given reaction i and multiply it by the Avogadro number and the volume V of the system we consider:

$$\phi = \mathcal{N}_a \cdot V \cdot v \tag{2.23}$$

ϕ_i is the number of reactions i that occurs within V per unit time, therefore a true reaction rate (in opposition to a reaction rate measured in concentration per unit time).

The product of this reaction rate by an infinitesimal time dt gives us the probability that the reaction of type i will take place in this time interval. This is precisely the definition given to propensity in 2.10.

Unless otherwise mentioned, in what follows we will always associate the stochastic propensity a_i with ϕ_i, the reaction rate derived from the reaction flux.

Test case

In this paragraph, we will test different reaction kinetics for the same system and discuss the non-optimized approach of the Gillespie algorithm. Consider the system described in Fig. 2.9

We can write the system as:

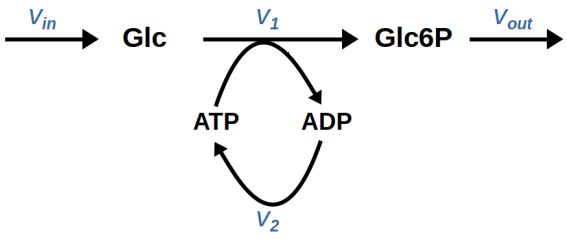

Figure 2.9: **A toy system to compare to compare stochastic and deterministic solving.** The system is composed by 4 reactions and 4 metabolites. Each reaction flux is highlighted in blue. Glucose (Glc) enter the system at rate v_in. Glucose react with one ATP to form Glucose-6-phosphate (Glc6P) and ADP at rate v_1. ADP is converted again into ATP at rate v_2. Finally Glucose-6-phosphate escape the system at rate v_{out}

$$\begin{aligned}
\frac{d[\text{Glc6P}]}{dt} &= v_{in} - v_1 \\
\frac{d[\text{Glc}]}{dt} &= v_1 - v_{out} \\
\frac{d[\text{ATP}]}{dt} &= -v_1 + v_2 \\
\frac{d[\text{ADP}]}{dt} &= v_1 - v_2,
\end{aligned}$$

At first, we assume each of the four reactions to follow a mass-action kinetics. Thus, one can write the four fluxes as follows:

$$\begin{aligned}
v_{in} &= k_{in} \\
v_1 &= k_1 \cdot [\text{Glc}] \cdot [\text{ATP}] \\
v_2 &= k_2 \cdot [\text{ADP}] \\
v_{out} &= k_{out} \cdot [\text{Glc6P}],
\end{aligned}$$

One classical way to get the propensities is to calculate the number of molecules N involved in the system and to convert deterministic kinetic constants into stochastic ones. Following Gillespie's paper, the 4 reactions above would lead to the following propensities:

$$\begin{aligned}
a_{in} &= k_{in} \cdot \mathcal{N}_a \cdot V \\
a_1 &= \frac{k_1}{\mathcal{N}_a \cdot V} \cdot N_{\text{Glc}} \cdot N_{\text{ATP}} \\
a_2 &= k_2 \cdot N_{\text{ADP}} \\
a_{out} &= k_{out} \cdot N_{\text{Glc6P}},
\end{aligned}$$

where the stochastic kinetic constant of each reaction will depend on the stoichiometry and the reaction type. If we denote ${K^s}_i$ as the stochastic kinetic rate of reaction $\mathcal{R}_i$, we would have: ${K^s}_{in} = k_{in} \cdot \mathcal{N}_a \cdot V$, ${K^s}_1 = \frac{k_1}{\mathcal{N}_a \cdot V}$, ${K^s}_2 = k_2$ and ${K^s}_{out} = k_{out}$.

In living systems, enzyme catalyzed reactions are often described with more complex kinetic rate-laws than mass-action kinetics. It usually consist in several elementary reactions, in which enzymes concentration can become limiting not following a mass-action kinetic anymore. One common description of enzymatic reactions is provided by the Michaelis-Menten (MM) kinetics equations, but sometimes reactions are subject to even more complex kinetics (allosteric activation/deactivation, product or substrate inhibition, ping-pong mechanism etc.). Thus transforming the corresponding kinetic constants of these rate laws becomes less and less trivial.

Here we propose to simply uses equation 2.23 to calculate the propensities from phenomenological rate laws used in deterministic system. We found again the propensity as calculated above:

$$\begin{aligned} a_{in} &= \mathcal{N}_a \cdot V \cdot k_{in} \\ a_1 &= \mathcal{N}_a \cdot V \cdot k_1 \cdot [\text{Glc}] \cdot [\text{ATP}] \\ a_2 &= \mathcal{N}_a \cdot V \cdot k_2 \cdot [\text{ADP}] \\ a_{out} &= \mathcal{N}_a \cdot V \cdot k_{out} \cdot [\text{Glc6P}], \end{aligned}$$

We can thus apply the Gillespie's direct method described in the previous section 2.1.3. We simulated the time course of the four metabolites until the system reaches steady-states. The initial (unrealistic) concentrations are the following: $[\text{Glc}]_0 = 0.1\ M$, $[\text{Glc6P}]_0 = 0.1\ M$, $[\text{ATP}]_0 = 0.4\ M$ and $[\text{ADP}]_0 = 0.4\ M$. The results of the simulation are presented in Figure 2.10. The colored lines are the deterministic approximations using an ODE's solver. Here we used Modelbase [Aalst et al., 2020] which uses the Assimulo solver [Andersson et al., 2015]. The black lines are the result from the Gillespie method using our equivalence between fluxes and propensities. As the volume V used for the conversion between concentration and number of particles is increased, the noise is drastically reduced and eventually becomes negligible. However, this comes at a cost in terms of computation time. Indeed, the simulation presented at the bottom right of the figure is obtained for $V = 10^{-18}\ m^{-3}$, a volume for which the initial conditions of $0.4\ M$ correspond to $240,000$ molecules. At steady state the system is composed of almost 10^6 molecules. Since each reaction is proceed successively and would change the number of each metabolite by ± 1 or 0 depending on the stoechiometry, the 20 seconds time-course is obtained in a bit more than a minute (1 minute and 18 seconds) in term of computational time on a Intel(R) Xeon(R) W-2135 CPU @ 3.70GHz.

Now, as a proof of concept, we changed reaction $\mathcal{R}_1$ by a Michaelis-Menten-rate law in the system presented above. The rate law can be written as:

$$v_1 = v_{\max} \cdot \frac{[\text{Glc}]}{K_M \cdot [\text{Glc}]}$$

.

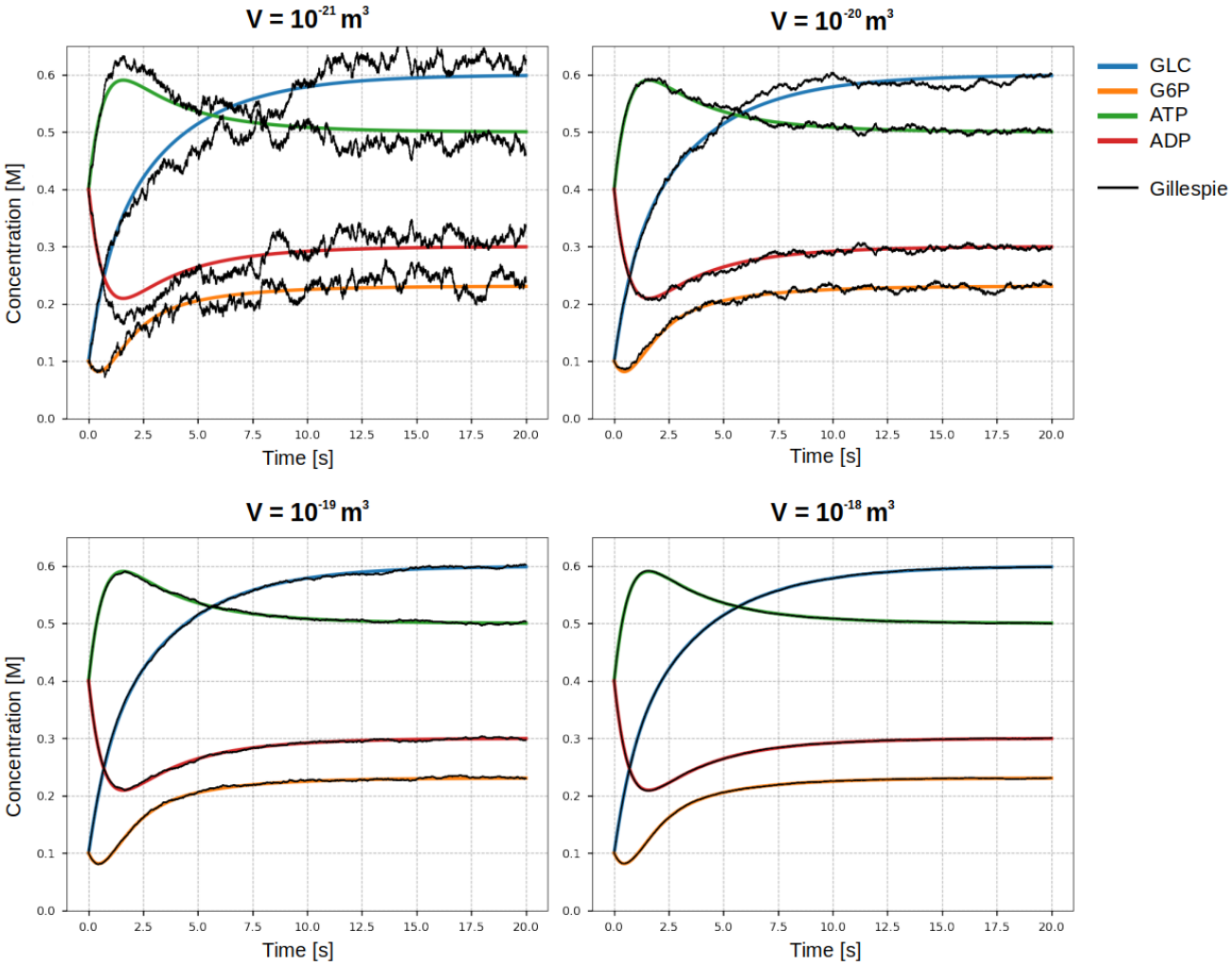

Figure 2.10: **Time courses of the species involved in the system for the deterministic and the stochastic approaches.** The colored lines show the result from the deterministic approach and the black lines show the result from the Gillespie algorithm for different volume V. When V increases, the fluctuations are reduced until the number of particles is big enough they becomes negligible.

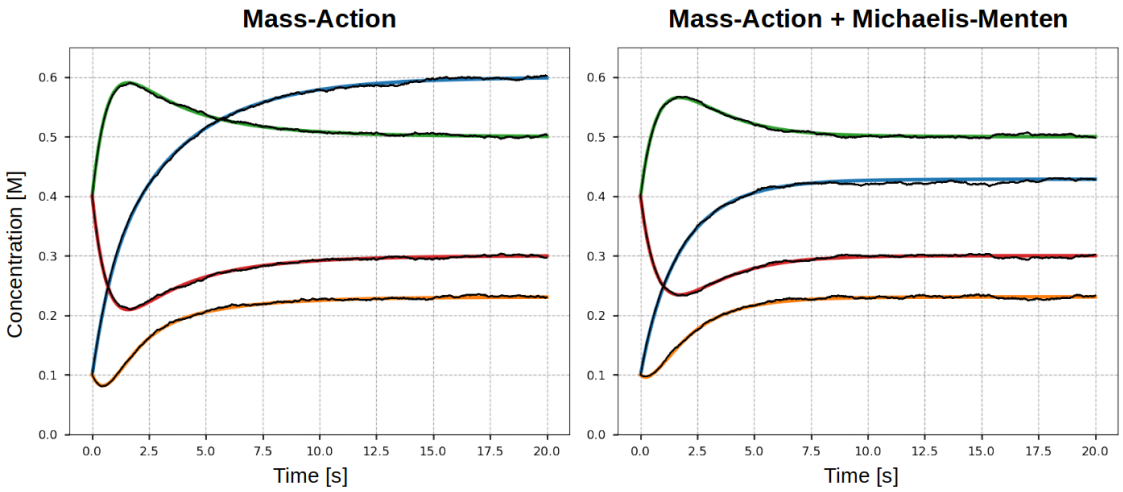

Figure 2.11: **Test case of the stochastic conversion of a Michaelis-Menten kinetic** Time courses of the species involved in the system for the deterministic and stochastic approaches.

Using equation 2.23 again, a_1 becomes simply:

$$a_1 = \mathcal{N}_a \cdot V \cdot v_{\max} \cdot \frac{[\text{Glc}]}{K_M \cdot [\text{Glc}]}.$$

We tested it with the same initial concentration, and setting $v_{\max}$ and K_M to 1. We obtained the result presented in Fig 2.11.

2.2 Numerical procedure

2.2.1 Glycogen granule object

In this section, we explain how the simulation of glycogen synthesis and degradation take place. To accomplish this, extensive information about the granule is required to make the molecule evolve in time and space, as well as the subsequent analytical treatment. In the preceding subsections (2.1.1, 2.1.2, and 2.1.3), we have introduced all the necessary requirements to construct our glycogen structural model. The first part accounts for the structure's geometry, including volume effects, and requires storing the positions of the glucose units and the introduction of a a chain identity. The chain identity allows for distinguishing between members and non-members of a chain and permitting or denying overlap between monomers belonging to different chains (see subsection 2.1.1). The second part of the model (subsection 2.1.2) pertains to enzyme mechanistics. It requires access to chain lengths and branch positions to determine the degree of "free" polymerization, which determines whether a chain is a substrate for one of the reactions. The model also necessitates accessing chain connectivity to branch

and unbranch chains. Creating a new "daughter" branch reduces the degree of "free" polymerization of the mother chain and must be updated accordingly. Therefore, it is essential to access the genealogy of the chains. Similarly, when a chain is unbranched, we must update the chain that loses a branch. The third part of the model (subsection 2.1.3) models the kinetics of the reactions using the Gillespie algorithm, which also requires knowledge of the chain lengths and the "free" portion. The free DP determines whether a chain is a substrate for a reaction, enabling the calculation of reaction propensities. Finally, all post-simulation analysis is based on the aforementioned information

To allow an easy access to this information, we store in a Python dictionary each of the chains, themselves being dictionaries. A chain will thus contain the information depicted in figure 2.12.

2.2.2 Main numerical steps to simulate synthesis and degradation

The simulations begin by initializing a granules object, consisting of two chains of DP 4 on opposite directions on the z axis, separated by the diameter of the glycogenin core. Initially, these chains are only substrates for elongation, but they rapidly grow in length, providing substrate chains for all other reactions. At each simulation step (as shown in Fig 2.13), the structure information is used to determine the number of substrate chains for each reaction. Propensities are calculated from this number to determine which reaction occurs and the corresponding elapsed time.

If GS is selected, the algorithm chooses one substrate chain and tries to elongate it in the direction of the chain. If the volume at the new location is not occupied by another branch, the reaction is successful, and the monomer is added to the structure. If the volume is occupied, the algorithm tries to add the monomer on another branch until the reaction is successful.

For GBE, the algorithm selects one substrate chain and possible cleaving points according to the rules introduced in subsection 2.1.2. From the selected cleaving points, the algorithm generates a new direction that will have the potential new branches. For all potentially transferred monomers, the algorithm checks if the space is not occupied by other glucose units. If the space is not occupied, the reaction is successful, and the change is implemented in the granule object, updating the mother chains and adding a new chain with a new identifier and the position of the glucose units. If the chains overlap, the algorithm restarts from selecting a substrate chain.

If GP is selected, a chain substrate for the reaction is chosen, and its DP is reduced by 1.

For GDE, a corresponding substrate chain is selected, and by default, $n - 1$ glucose units are transferred to the mother chain while the remaining glucose is removed. The mother chain is updated, and the information on the debranched chain is erased.

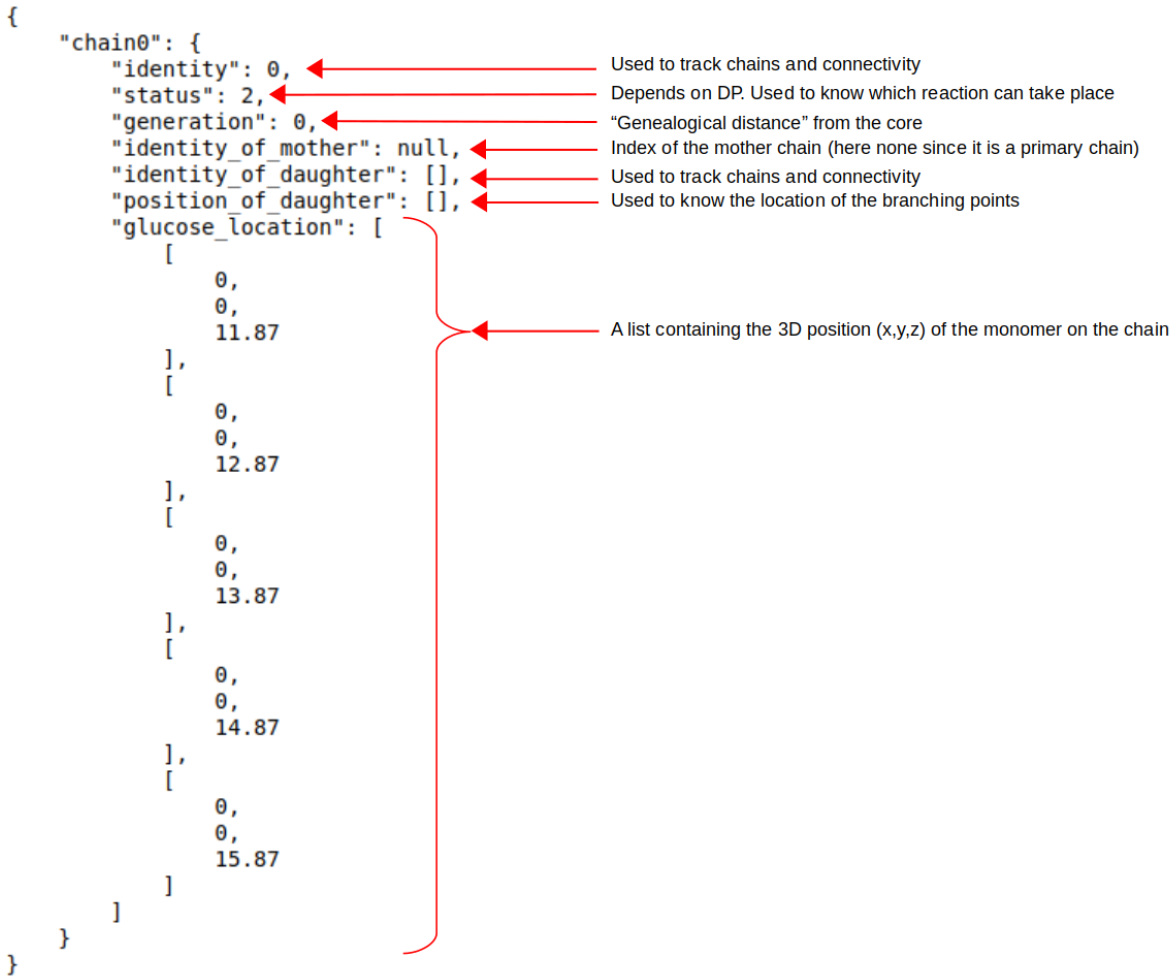

Figure 2.12: **Python dictionary containing the information needed in the simulation.** This example shows the initial information contained for a simulation starting with a single chain. The status is "2", meaning that the chain is subject to elongation (GS) and degradation (GP) only. There is no "mother" chain because it is the primary chain attached to the glycogenin core. There is no branching point (empty list for identity and position of the "daughter"). Finally, by default the chains are initialized to DP5 on the z axis, starting at a distance r_{GN} from the center (radius of the glycogenin).

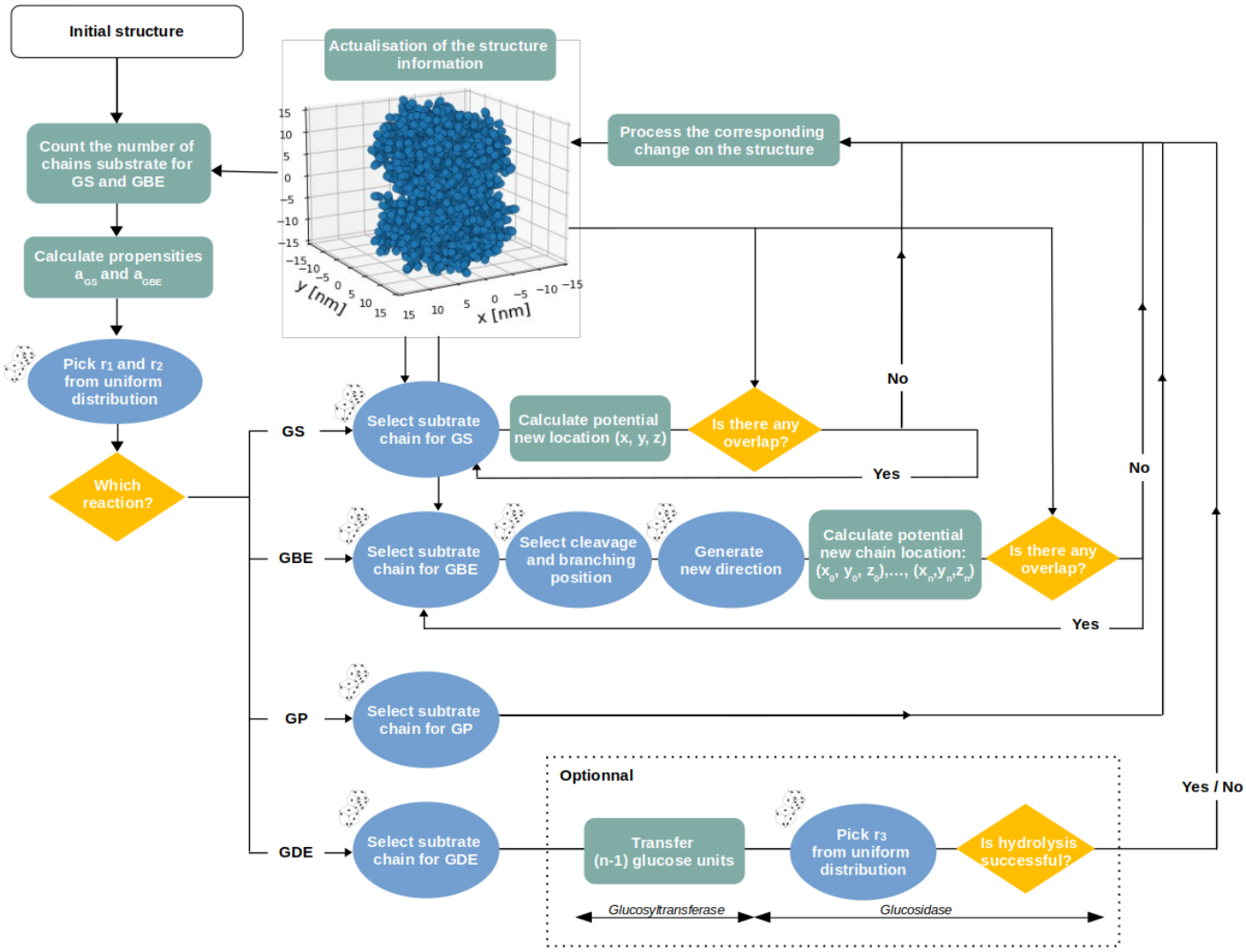

Figure 2.13: **Algorithm flow diagram.** This scheme shows some of the main steps of our algorithm. The green squares represent deterministic calculations or actions. The blue ovals involve random numbers. Finally, the yellow diamonds represent questions or "if" statements.

Chapter 3

A structural model for glycogen synthesis and degradation

This section consists in an article that have been accepted in PLOS Computational Biology. I was the main contributor to the conceptualization, formal analysis, investigation, methodology, programming, results visualization, and draft writing.

Summary

The article presents the structural glycogen model that we have developed, along with an in-depth analysis of its ability to facilitate glycogen synthesis in the presence of only Glycogen Synthase and Glycogen Debranching Enzyme. Our study is focused on analyzing the impact of the ratio of glycogen synthase to branching enzyme activities on the granule's structure. To begin with, we examine the model using generic parameters before validating it against experimental data obtained from mice.

Our investigation into the mechanism of branching leads us to parameterize it using distinct lengths. We consider various sets of values for these lengths and different rules for their application, ultimately showing how these parameters can be combined to fine-tune the structure of glycogen macromolecules. By comparing our model to experimental data, we demonstrate that we can accurately reproduce glycogen chain length distributions in wild type mice. Furthermore, the additional structural properties obtained from this fit are consistent with values reported in the literature.

We find that the mechanism of branching may be more flexible than previously thought. Our study suggests that the chain length distribution is an indicator of the branching activity and mechanism, and our model provides a theoretical basis for quantifying these effects. It can be applied to any glycogen data set, and could potentially contribute to characterizing the mechanisms at work in glycogen storage disorders.

We further extend the investigation and provide a characterization of the A to B ratio. Our findings indicate that if the branching reaction is symmetrical in a way such that the transferred chain is of the same length as the remaining chain, the expected A:B ratio is one. However, when the branching reaction differs, we demonstrate that different A:B ratios can be obtained,

thereby explaining the variability in ratios observed among polysaccharides. Additionally, we quantify the effects of steric hindrance on the granule and demonstrate that this addition to the model is necessary to accurately predict the distribution of chain lengths.

Stochastic modelling of a three-dimensional glycogen granule synthesis and impact of the branching enzyme

Yvan Rousset[1], Oliver Ebenhöh[1,2], Adélaïde Raguin[2,3*]

1 Institute for Quantitative and Theoretical Biology, Heinrich-Heine University, Düsseldorf, Germany
2 Cluster of Excellence on Plant Sciences (CEPLAS), Heinrich-Heine University, Düsseldorf, Germany
3 Institute for Computational Cell Biology, Heinrich-Heine University, Düsseldorf, Germany

* adelaide.raguin@hhu.de

Abstract

In humans, glycogen storage diseases result from metabolic inborn errors, and can lead to severe phenotypes and lethal conditions. Besides these rare diseases, glycogen is also associated to widely spread societal burdens such as diabetes. Glycogen is a branched glucose polymer synthesised and degraded by a complex set of enzymes. Over the past 50 years, the structure of glycogen has been intensively investigated. Yet, the interplay between the detailed three-dimensional glycogen structure and the related enzyme kinetics is only partially characterised and still to be fully understood. In this article, we develop a stochastic coarse-grained and spatially resolved model of branched polymer biosynthesis following a Gillespie algorithm. Our study largely focusses on the role of the branching enzyme, and first investigates the properties of the model with generic parameters, before comparing it to *in vivo* experimental data in mice. It arises that the ratio of glycogen synthase over branching enzyme activities drastically impacts the structure of the granule. We deeply investigate the mechanism of branching and parametrise it using distinct lengths. Not only do we consider various possible sets of values for these lengths, but also distinct rules to apply them. We show how combining them finely tunes glycogen macromolecular structure. Comparing the model with experimental data confirms that we can accurately reproduce glycogen chain length distributions in wild type mice. Additional granule properties obtained for this fit are also in good agreement with typically reported values in the experimental literature. Nonetheless, we find that the mechanism of branching must be more flexible than usually reported. Overall, our model provides a theoretical basis to quantify the effect that single enzymatic parameters, in particular of the branching enzyme, have on the chain length distribution. Our generic model and methods can be applied to any glycogen data set, and could in particular contribute to characterise the mechanisms responsible for glycogen storage disorders.

Author summary

Glycogen is a granule-like macromolecule made of 10,000 to 50,000 glucose units arranged in linear and branched chains. It serves as energy storage in many species, including humans. Depending on physiological conditions (hormone concentrations, glucose level, etc.) glycogen granules are either synthesised or degraded. Certain

metabolic disorders are associated to abnormal glycogen structures, and structural properties of glycogen might impact the dynamics of glucose release and storage. To capture the complex interplay between this dynamics and glycogen structural properties, we propose a computational model relying on the random nature of biochemical reactions. The granule is represented in three dimensions and resolved at the glucose scale. Granules are produced under the action of a complex set of enzymes, and we mostly focus on those responsible for the formation of new branches. Specifically, we study the impact of their molecular action on the granule structure. With this model, we are able to reproduce structural properties observed under certain *in-vivo* conditions. Our biophysical and computational approach complements experimental studies and may contribute to characterise processes responsible for glycogen related disorders.

1 Introduction

Management of energy resources is crucial for an organism to survive, since nutrient availability may vary considerably with time. Moreover, organisms face numerous other stresses that may temporarily increase energy demand. To react to such changes in energy supply and demand, it is apparent that some internal stores are necessary. These stores are filled when nutrients are abundant and depleted when demand exceeds available supply. As direct substrate for catabolic pathways, glucose plays a central role in energy metabolism in most organisms [1]. While plants have found their way to store glucose as starch, animals, fungi, and most bacteria store glucose as glycogen. Both starch and glycogen are branched polymers consisting of glucose monomers, linked into linear chains by α-1,4 bonds, and branching points by α-1,6 bonds. However, these two macromolecules exhibit rather different physico-chemical properties. In contrast to glycogen, starch is insoluble in water under physiological conditions, and contains high density crystalline regions. These different properties are reflected by distinct branching patterns and chain length distributions (CLD). Functionally, starch and glycogen can be compared to capacitors in electric circuits. The latter are able to store and release electrons depending on current and voltage. Thus, they can be used to stabilise a fluctuating electric signal. Analogously, glycogen and starch can be seen as two different capacitors that both contribute to glucose homeostasis by managing energy resources through time.

While the fine structure of starch has been widely investigated over the past 50 years, less is known on that of glycogen [2]. The characterisation of the detailed structure of glycogen, as well as the interplay between its structural properties and the dynamics of glycogenesis (synthesis) and glycogenolysis (degradation) is unclear. Yet, both glycogen structure and dynamics are of utmost interest for understanding glycogen metabolism and the impact of related genetic variances. For human health, this is particularly important considering the increasing prevalence of glycogen storage diseases (GSDs), as well as diabetes, and other glycogen related disorders.

So far, different structures of glycogen have been proposed [3–6], but a structural model that emerges from the detailed underlying enzymatic mechanisms of synthesis and degradation is still lacking. Understanding and precisely describing the biochemistry of glycogen is challenging. With a molecular weight of 10^6 to 10^7 g·mol^{-1} [7–9], glycogen is a large molecule, even compared to the enzymes involved in its dynamics [10–12]. Thus, enzymes synthesising, degrading, or otherwise altering glycogen, can only access certain branches near its surface, while many glucose residues near the centre of the molecule remain 'hidden'. At the macroscopic scale, the structure of glycogen is well known. Drochmans [13] observed two populations of glycogen granules in rat liver. The so-called α granules are aggregates of the smaller β

granules [14–16]. The latter have a radius in the range from 10 to 20 nm, while the radius of α granules ranges between 40 and 300 nm [13]. Here, we focus on β granules, whose synthesis and degradation both involve a relatively small number of enzymes (see Fig 1). For a comprehensive review, see [17].

To initiate a new glycogen molecule, a chain of 5 glucose units is synthesised and bound to a glycogenin (GN) protein, which forms the core of the final granule [18–21]. Once initiation is completed, the granule is expanded by the two enzymes Glycogen Synthase (GS) and Glycogen Branching Enzyme (GBE). GS is an elongating enzyme that adds one glucose residue to the non-reducing end of an α-1,4 linear chain, thereby forming a new α-1,4 glucosydic bond. GBE cleaves a small part of a linear chain and creates a new branch by forming an α-1,6 glucosydic bond. We call "daughter" the newly formed chain and "mother" the one it is branching from.

Besides synthesis, granules are subject to degradation, that is performed by two other enzymes. Glycogen Phosphorylase (GP) and Glycogen Debranching Enzyme (GDE) respectively shorten and debranch glycogen branches.

Depending on the relative enzymatic activities of these four enzymes (GS, GBE, GP, and GDE) the overall size of the glycogen granule can either increase or decrease. In this article, we choose to focus on glycogen synthesis, and more specifically the role of the branching enzyme GBE.

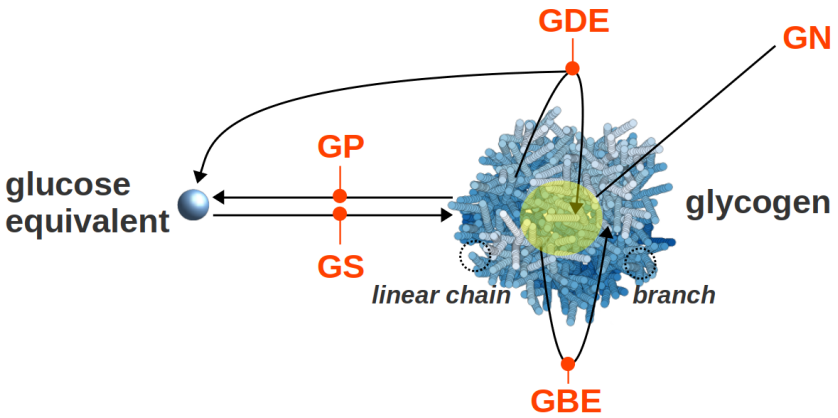

Fig 1. Main enzyme reactions involved in the synthesis and breakdown of glycogen. *In vivo*, the GS and GBE enzymes synthesise glycogen, while the GP and GDE degrade it. Besides, GN is the initial precursor of the granule and stands in its core. Enzymes are noted in orange, glucose residues are in blue, and GN is highlighted with a yellow sphere.

Experimental observations [22, 23] of glycogen show an average chain length (CL) around 13 glucose units, which depends on the organism type. A typical peak is observed at low degree of polymerisation (DP), around DP 8, and almost no chains are detected above DP 40 [22–25]. The degree of branching is defined in two ways in the literature. It is most commonly defined as the ratio of α–1,6 to α–1,4 linkages, but sometimes also as the average number of α–1,6 bonds per chain [6]. We will apply the first definition throughout the paper. This ratio is in the range $0.02 - 0.05$ in amylopectin [26–28] from starch and $0.06 - 0.10$ in glycogen [9, 29]. In 1956, Peat et al. [30] introduced the concept of A and B chains. An A chain does not carry any branch, while a B chain does. The A:B ratio is an important characteristic of glycogen granules and an indicator of the branching pattern. Early studies reported an A:B ratio of 1 in glycogen [22, 31], while it is usually greater than 1 in amylose, and ranging

from 1.5 to 2.6 in amylopectin [31, 32].

As early as in the 1940s various hypotheses have been formulated aiming at explaining macroscopic features of glycogen granules. One of them has become known as the Whelan model, which assumes that every chain carries on average two branches. Based on the Whelan model, the nowadays widespread idea emerged that glycogen can be described as a fractal molecule [6, 33, 34]. A fractal glycogen structure is indeed attractive because it can reproduce various structural properties of glycogen. Moreover, it provides a mechanism explaining why glycogen granules seem to have a maximum size. With this model, glycogen becomes too dense at the surface due to an exponential increase of the number of non-reducing ends with the distance from the centre. Thus, steric hindrance prevents synthesis to continue. However, Manners [2] stressed in 1990 that there is no clear evidence that supports a regular branching and therefore a fractal pattern. More recently, further arguments and results against a fractal structure have been raised [35–37]. Besides, independently from any fractal structure considerations, Zhang et al. [38] proposed a mathematical model based on a Monte Carlo approach to numerically simulate glycogen biosynthesis, aiming to support the common idea that steric hindrance limits granule growth. In this model, glucose units are placed on a three-dimensional grid, and the granule biosynthesis is simulated by adding glucose units on any of the 26 neighbouring positions around the end of the growing chain. As a result, limited growth is an emergent property of this model.

Here, we propose a mechanistic model for glycogen synthesis, focussing on the impact of the branching enzyme on the granule structure. We aim at explaining macroscopic and experimentally observable quantities, such as chain length distributions (CLDs), from the underlying enzymatic mechanisms. Such CLDs can be predicted theoretically using numerical inversion of Laplace transforms [39] or kinetic equations [40–42]. However, besides the fact that CLDs are only one of the many quantities that can be measured in the glycogen structures we simulate, our model accounts for complex features for which an analytic treatment is no longer feasible, e.g. the complete connectivity of the structure, complex enzyme mechanisms, simple helix structure of linear glucan chains, and steric-hindrance effects. Also, in contrast to the model proposed by Zhang et al. [38] based on stochastic simulations, we do not restrict glucose units to a grid, and instead reflect the helical structure of glucose chains, using parameter values derived from biophysical properties. Therefore, our model provides a more flexible and more realistic representation of the three-dimensional granule structure. Specifically, it is designed to study the effect of the enzymatic activity on the structure, and to infer mechanism parameters from measurable quantities, while, Zhang et al. [38] focus on explaining the limited growth of glycogen granules. To our knowledge, these scientific questions have so far not been addressed by computational approaches. In this article, we first detail the geometrical and structural features, and enzymatic reactions, taken into consideration in our model. Then, we analyse distinct properties of the model with a specific focus on the effect of the branching enzyme. Finally, we compare the model to experimental data and discuss the parameter values resulting in a best fit, in relation to typical values reported in the experimental literature. In addition, several complementary results justifying our modelling choices are reported in the extensive Supplementary Material.

2 Results

2.1 The model

Glycogen structure and geometry

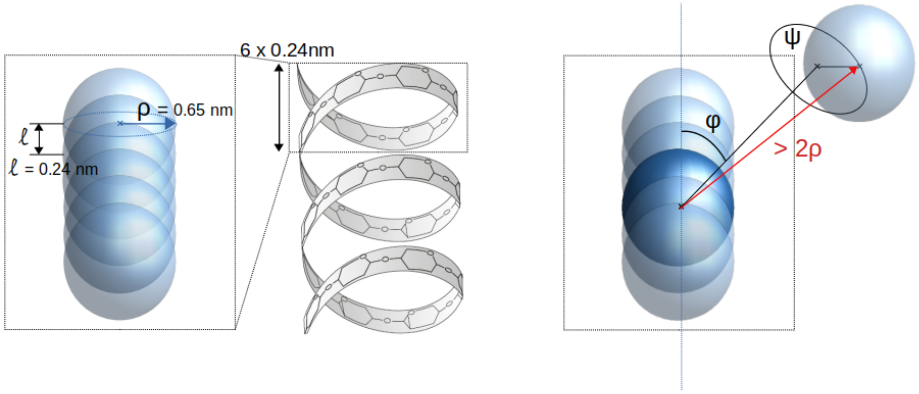

Fig 2. Geometrical description of glucose chains. Left: Coarse-grained linear chain. Assuming helical chains, glucose units are described as interpenetrated spheres with radius $\rho = 0.65$ nm. Two consecutive glucoses are distant by $l = 0.24$ nm, which is the radial contribution to the chain length of one glucose in a helical structure. **Right: Description of a branching point.** We generate the direction of the new branch by randomly picking two angles φ and ψ. The first monomer of the new branch will be located at a distance greater than 2ρ to insure no overlapping between the mother and the daughter branches.

We represent glycogen granules as simplified three-dimensional structures, in which every glucose monomer is characterised by its position in space (see Fig 2). To describe the branched tree-like structure, we generalise the simple representation of linear self-avoiding polymers. Using X-ray experiments, Goldsmith et al [33] characterised in detail how glucose molecules are arranged into helical α-1,4 linear chains. The cross-section of the helix has been calculated to be 1.3 nm^2 with 6 to 7 residues per turn of length 1.4 nm. The radius of the circular cross-section is thus $\rho = 0.65$ nm, and each glucose residue contributes to the chain length by $l = 0.24$ nm. Inspired by these findings, we propose that monomers are described as overlapping spheres with radius $\rho = 0.65$ nm, equal to that of the helix. The validity of this assumption and its impact on our results are presented in detail in the Supplementary Material (see Fig 12). Besides, the center of consecutive monomers are distant by $l = 0.24$ nm, which corresponds to the contribution of one glucose unit to the chain length, but also involves that the coarse-grained monomer spheres overlap. However, to realistically reflect self-avoidance, two different chains cannot overlap. With these spatial considerations, we ensure that the contribution to the volume by one glucose unit in the model is similar to that of the real helical chain. This way, we provide a description which is simple enough to be easily represented in a computer model, but still realistic enough to reflect the spatial properties of linear chains arranged into helices.

Describing branches formed by α-1,6 linkages requires additional geometrical considerations. As illustrated in Fig 2 (Right panel), a branch point is defined by the monomer on the mother chain to which it is attached, and two angles defining the

direction of the daughter chain. Besides, the anchoring monomer on the mother chain and the first one of the daughter chain are distant by more than 2ρ, ensuring that they do not overlap.

Glycogen synthesis is initiated by GN that is located in the centre of the molecule [43]. It contributes to the total volume of glycogen, and thus we also consider self-exclusion between GN and any glucose residue of the granule. For the sake of simplicity, we assume that it is a sphere of density 1.3 $\mathrm{g \cdot cm^{-3}}$ [44], which is the typical density of a protein. Accounting for its two sub-units of 38 kDa each [11], we approximate its radius by $\rho_{\mathrm{GN}} \approx 2.85$ nm. Together with GS, which is responsible for elongation, GN may initiate more than a single primary chain, possibly either 2 or 4 [21]. We model this initial core structure with two primary chains pointing out of the GN sphere in opposite directions.

2.1.1 Enzymatic reactions

Two enzymes (GS and GBE) are directly involved in glycogen synthesis. Their role is illustrated in Fig 3.

GS binds the non-reducing end of an α-1,4 linear chain and elongates it by adding one glucose residue. It is commonly assumed that GS needs a glucose chain primer as a glucose acceptor [45]. In the model, we call L_{min}^{GS} the minimal chain length of this required primer by GS. As elongation takes place, the chain becomes long enough to be cleaved and branched. This reaction is catalysed by GBE. Just like GS, the action of GBE is also characterised by specific substrate and product lengths as shown in Fig 3. Since less is known for GBE, we tested two models for its mechanism, namely the strict location model and the flexible location model. These are detailed in the Supplementary Material (see Figs 13 and 14). While comparing the two models to experimental data, we observed that the flexible location model provides a considerably better fit (see Figs 15 and 16 in the Supplementary Material). Thus, throughout the paper, we choose to use the flexible location model and will vary the GBE associated parameters. As illustrated in Fig 3, we consider a minimal chain length for the substrate (noted $L_{\mathrm{min}}^{\mathrm{GBE}}$) such that GBE is able to bind. We ensure that no daughter branch stands between the binding point and the non-reducing end of the branch. After binding, GBE cleaves at least $L_{\mathrm{transferred}}^{\mathrm{GBE}}$glucose units. Finally, GBE must attach the cleaved chain on the initial substrate, following an intramolecular process, and creating a new A chain. To precisely describe this last step, we define two additional lengths. First, the new α-1,6 branching point must not be closer than $L_{\mathrm{spacing}}^{\mathrm{GBE}}$ from either the first glucose of the chain, or an above α-1,6 branching point. Second, the new branching point must not be closer than $L_{\mathrm{leftover}}^{\mathrm{GBE}}$ from the non-reducing end of the substrate chain, which is the original position of cleavage. Thus, for a new branch to be created by GBE, the substrate branch must have a chain length greater than $L_{\mathrm{min}}^{\mathrm{GBE}}$, verifying:

$$L_{\mathrm{min}}^{\mathrm{GBE}} > L_{\mathrm{spacing}}^{\mathrm{GBE}} + L_{\mathrm{transferred}}^{\mathrm{GBE}} + L_{\mathrm{leftover}}^{\mathrm{GBE}}. \quad (1)$$

To illustrate the impact of these minimal lengths on the reaction outcomes, in Fig 4 we detail the case of $\{L_{\mathrm{spacing}}^{\mathrm{GBE}}, L_{\mathrm{transferred}}^{\mathrm{GBE}}, L_{\mathrm{leftover}}^{\mathrm{GBE}}\} = \{2, 2, 2\}$. If the substrate reaches a length of DP equal 7, the condition (Eq 1) is fulfilled and the reaction may take place. If this reaction occurs, there is a single possible outcome (Fig 4, left panel). If the branching reaction occurs on a longer chain than just the minimal one, several outcomes are possible, all fulfilling the set of rules specified by the triplet $\{2, 2, 2\}$. Fig 4 (Right panel) depicts the case of a substrate chain with 9 residues, which results in 6 possible outcomes.

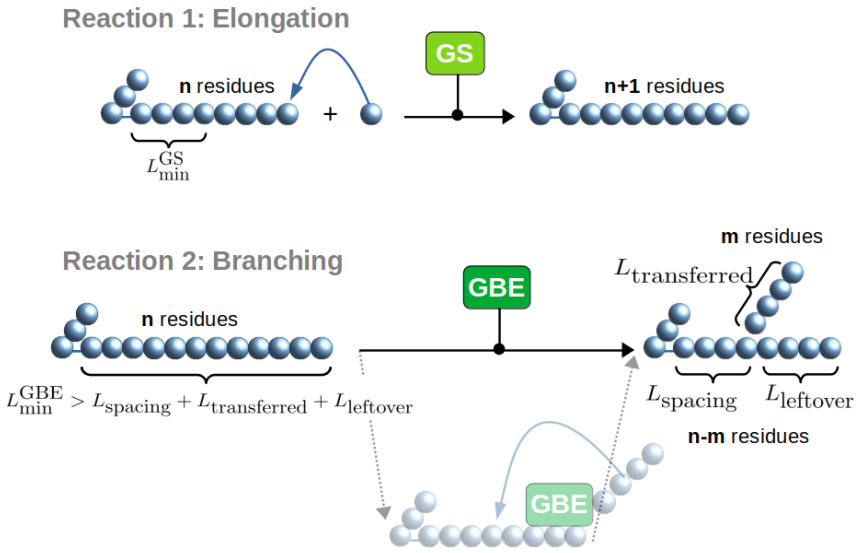

Fig 3. Glycogen synthesis reactions. Glycogen Synthase (GS) catalyses the elongation reaction. It needs a branch with a minimal DP as a substrate and a glucose unit to react. Glycogen Branching Enzyme (GBE) catalyses the branching reaction if the substrate's DP is greater than the sum of 3 different minimal lengths.

2.2 Analysis of the model's properties with generic parameters

2.2.1 Elongation to branching ratio

When the simulation starts, the system is composed of a GN core with two primary chains standing in opposite directions in the center of the simulation space. Two enzymes (GS and GBE) modify the structure of the glycogen granule. To quantify that, we define Γ as the ratio of the GS over GBE reaction rates. For $\Gamma \approx 0$, branching dominates over elongation, and *vice versa* when $\Gamma \gg 1$. Fig 5 (Top) shows two simulated glycogen structures obtained with $\Gamma = 0.2$ and $\Gamma = 10$, respectively. We observe that a low Γ corresponds to a tightly packed structure, while a high Γ leads to a sparsely packed structure, with further elongated chains. Both simulations have been performed with a high N value, such that the number of monomers is not a limiting factor. The simulations end when the total number of monomers incorporated into the growing granule is $N = 50{,}000$. As can be seen in Fig 5 (Top), when Γ increases, for the same number of glucose units incorporated into the granule, the length of the chains increases while their number decreases, and so does the number of non-reducing ends (represented by green spheres).

Fig 5 (Bottom) shows the chain length distribution (CLD) for the two simulated structures. It is computed as the abundance of each chain length, taking into consideration all the chains of the structure. With $\Gamma = 0.2$ (light grey histogram) the average DP is 12.2 with a peak in the range $[4 - 8]$ DP. When elongation is stronger than branching ($\Gamma = 10.0$, black histogram) the distribution shifts to higher DPs, the mean is 38.8, and the intensity of the peak is much reduced. By analogy with the well-studied case of starch, a chain length distribution with abundant high DP, might be an indication of double helix formation [46]. We do not model double helices as such, but our results allow to determine the Γ range that might lead to double helix formation, and thus potential glycogen precipitation.

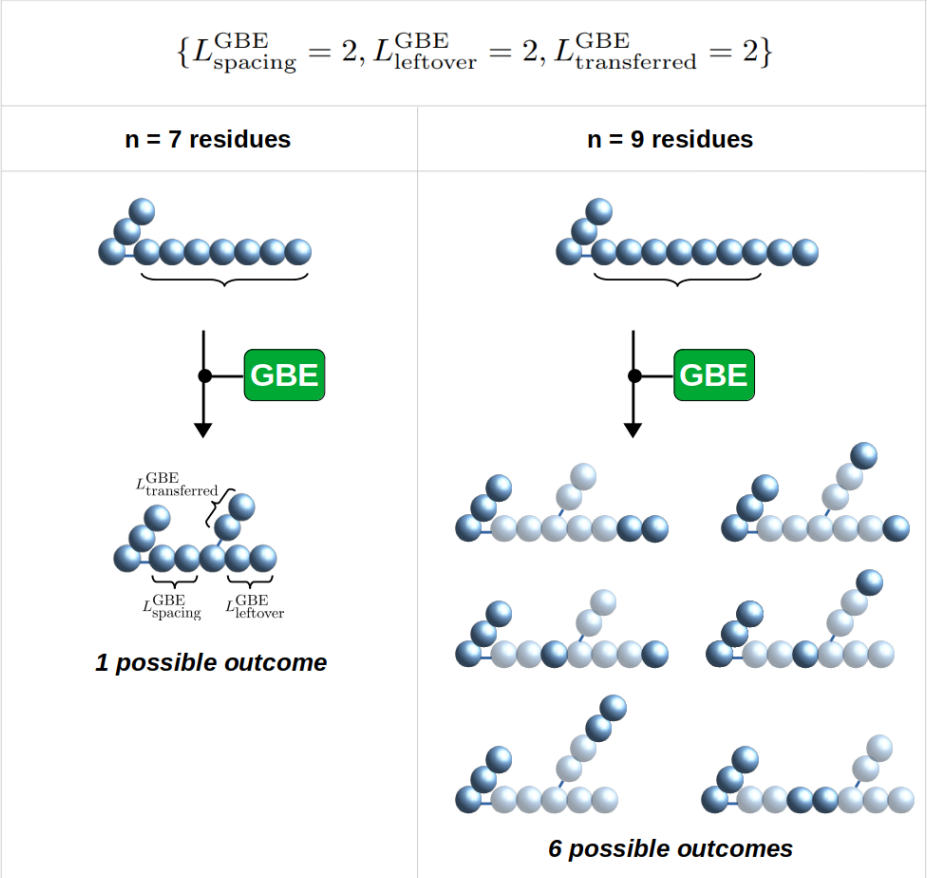

Fig 4. Illustration of the potential outcomes by GBE branching with $\{L^{\mathbf{GBE}}_{\mathbf{spacing}} = 2, L^{\mathbf{GBE}}_{\mathbf{transferred}} = 2, L^{\mathbf{GBE}}_{\mathbf{leftover}} = 2\}$. With these minimal lengths, the minimal DP required for a branching to occur is DP = 7. If the chain length is longer, the number of possible outcomes increases. **Left:** With a substrate of DP = 7, only one outcome is possible. **Right:** With a substrate of DP = 9, up to 6 distinct outcomes are possible.

2.2.2 Granule density

Our approach tracks the position (x, y, z) of each glucose unit in three-dimensions. This allows us to compute how densely the granules are packed. Granule packing is quantified by the occupancy Ω, which is defined by the volume occupied by glucoses V_{glucose} divided by the total volume V_{total},

$$\Omega = \frac{V_{\mathrm{glucose}}}{V_{\mathrm{total}}}. \tag{2}$$

To characterise an entire granule, we consider the total volume V_{total} to be a sphere of gyration radius R_g (defined in the section 5.1 of the Supplementary Material). To determine the occupancy at a given radius r from the center of the granule, we estimate the local occupancy in a spherical shell between radii r and $r + \Delta r$. Fig 6 displays how the occupancy Ω as a function of the radius r dynamically changes during granule synthesis for two different values of the elongation to branching ratio Γ. The left panel shows the formation of a tightly packed granule ($\Gamma = 0.2$), while the

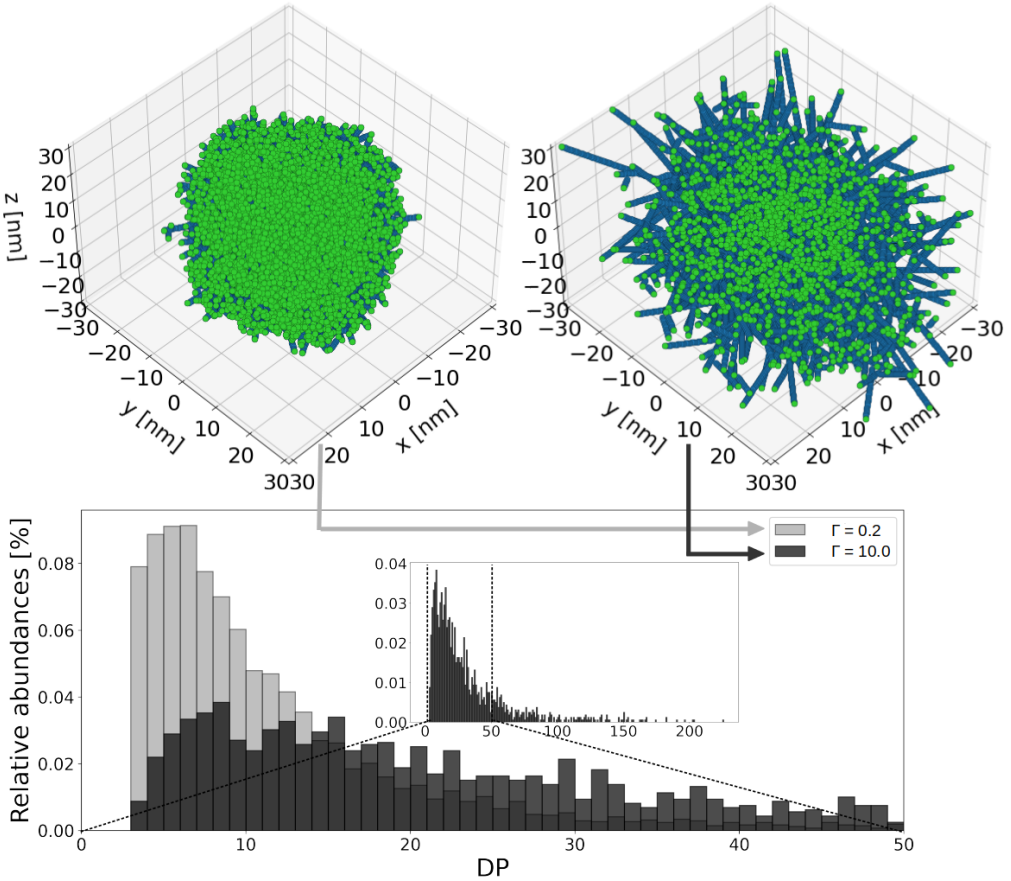

Fig 5. Simulated glycogen granule structures for different elongation to branching ratios (Γ). 50,000 glucose units are incorporated. **Top: 3D structures of glycogen granules.** Blue spheres represent the glucose units, green spheres the non-reducing ends. When $\Gamma = 0.2$, the structure of the granule is tightly packed. For $\Gamma = 10.0$, the structure of the granule is sparsely packed. **Bottom: Associated chain length distributions.** The light grey histogram shows the CLD for the tightly packed granule, while the black one shows that of the sparsely packed granule. The inset shows the full range of DP for $\Gamma = 10.0$. The longest chain is found to have a DP of 226.

right panel shows a sparsely packed structure ($\Gamma = 10.0$). Each line in the figure corresponds to the incorporation of 5,000 glucose units into the granule. It can be observed that granule synthesis proceeds in two phases. The first phase is characterised by an increase of the density close to the granule centre, while the radial extension increases only moderately. This can be explained by the fact that initially there is sufficient space to add new glucose units and there is hardly any steric hindrance among them. When steric hindrance constrains the synthesis (after around 10,000 glucose units have been incorporated), the system transits to the second phase. The latter is characterised by a radial expansion of the overall structure, while the density remains approximately constant around 30%. The two phases can be observed for both Γ values considered, but they are more pronounced for the tightly packed granule ($\Gamma = 0.2$). Our simulated occupancy profile is in agreement with recent experimental measurements of the radial density profile of glycogen granules [37].

There, the authors also observed a rather constant density in the inner part of the granule, and a decrease at its periphery.

The relatively low occupancy close to the granule center is due to the presence of the GN core, which is not counted in the occupancy, but its corresponding volume cannot be filled with glucose units. We observe that at most 30% of the granule volume is occupied by glucose. This value is rather low, which is expected, since in the model, branches are straight helices, without any flexibility, while in reality, branches can bend and form locally higher densities. Specifically, this occupancy value is below $\Omega = 0.45$, which we can estimate from previous studies [33]. There, the authors combine experimental and modelling approaches to conclude that granules of 55,000 glucoses have a radius of 21 nm [33]. Nonetheless, there are many uncertainties on the molecular masses experimentally measured. Thus, the occupancy value (i.e. $\Omega = 0.45$) that we deduce from their work might hold large errors. For instance, it is unclear how the water and protein molecules embedded in the granules contribute to the molecular masses experimentally measured.

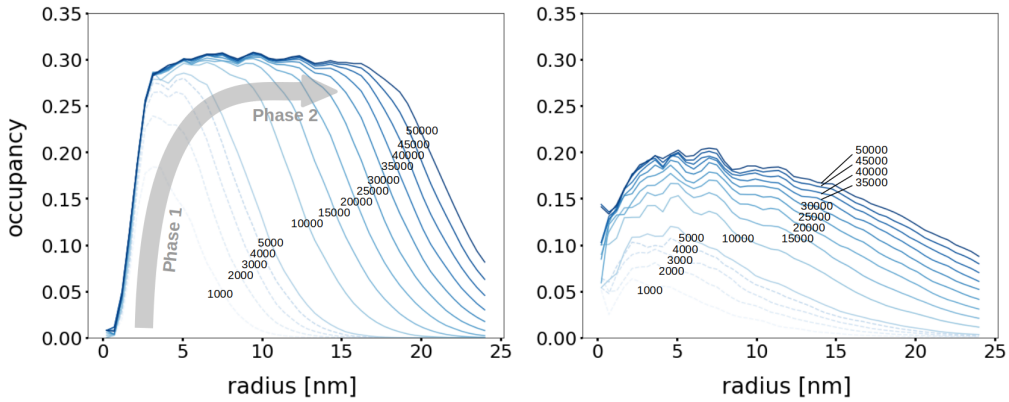

Fig 6. Dynamics of the occupancy profiles for a tightly ($\Gamma = 0.2$, left) and sparsely ($\Gamma = 10.0$, right) packed granule. Occupancy as a function of the radius at different simulation times. Each line corresponds to an incorporation of $N = 5{,}000$ glucose units. The simulation stops at $N = 50{,}000$. The grey arrow highlights the two phases of the granule synthesis dynamics. In phase 1, steric hindrance constrains are low, allowing occupancy to increase. In phase 2, i.e. after incorporation of ca. $N = 10{,}000$ glucose units, the occupancy reaches a plateau and the granule expands.

2.2.3 Effect of the branching enzyme on the CLD

The branching enzyme mechanism is characterised by a triplet of integer numbers, denoted $\{L^{\text{GBE}}_{\text{spacing}}, L^{\text{GBE}}_{\text{transferred}}, L^{\text{GBE}}_{\text{leftover}}\}$, which specifies a unique set of rules for the enzymatic reaction. These rules considerably impact the CLD. Fig 7 shows that when these minimal lengths increase, the peak of the CLD is less pronounced and the distribution spreads towards higher DPs. Also, a change in each minimal length has a specific effect on the CLD.

Increasing $L^{\text{GBE}}_{\text{spacing}}$ drastically reduces the peak and spreads the distribution, while making it bimodal. An increase in $L^{\text{GBE}}_{\text{spacing}}$ reduces the granule's number of possible configurations. Less configurations are possible, and $L^{\text{GBE}}_{\text{spacing}}$ further shapes the chain length distribution. Chains of DP that are combinations of $L^{\text{GBE}}_{\text{spacing}}$ and $L^{\text{GBE}}_{\text{leftover}}$ are favoured, resulting in local peaks. Increasing $L^{\text{GBE}}_{\text{leftover}}$ also decreases the peak and spreads the overall distribution towards higher DPs. Besides, $L^{\text{GBE}}_{\text{transferred}}$ has a

different effect on the structure. Since $L^{\text{GBE}}_{\text{transferred}}$ is the smallest DP that can be formed, it is found on the leftmost part of the CLD, and variations in $L^{\text{GBE}}_{\text{transferred}}$ shift the overall distribution by the corresponding exact amount.

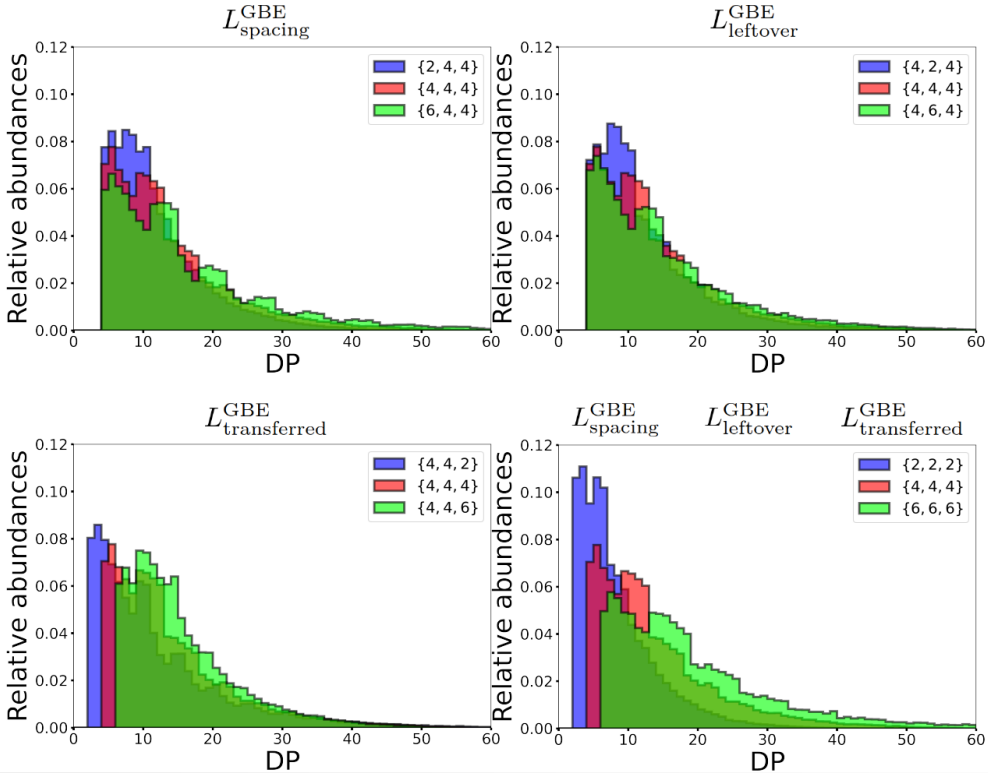

Fig 7. Effect of the chain length specificities on the CLD. Top-left: CLD for $L^{\text{GBE}}_{\text{spacing}} = 2$, 4, and 6, respectively. When $L^{\text{GBE}}_{\text{spacing}}$ increases, a multi-modal distribution emerges. **Top-right:** CLD for $L^{\text{GBE}}_{\text{leftover}} = 2$, 4, and 6, respectively. When $L^{\text{GBE}}_{\text{leftover}}$ increases, the peak is reduced and the overall distribution spreads towards higher DPs. **Bottom-left:** CLD for $L^{\text{GBE}}_{\text{transferred}} = 2$, 4, and 6, respectively. When $L^{\text{GBE}}_{\text{transferred}}$ increases, the distribution shifts towards higher DPs. **Bottom-right:** CLD for $L^{\text{GBE}}_{\text{spacing}} = L^{\text{GBE}}_{\text{transferred}} = L^{\text{GBE}}_{\text{leftover}} = 2$, 4, and 6, respectively. Varying these distinct minimal lengths concomitantly, combines the individual effects described above, when a single length is varied. Each CLD is the result of averaging 200 simulations of granules with 5,000 glucose units each.

It is important to notice, that these results are obtained when branching dominates over elongation. Increasing the elongation to branching ratio Γ systematically smoothens any multi-modal CLD, because it introduces flexibility in the branching location. It also flattens the peak and spreads the distribution towards higher DPs. Consequently, bi-modal distributions are obtained for high $L^{\text{GBE}}_{\text{spacing}}$ and low Γ values.

As discussed in the Elongation to branching ratio section, we can compare the synthesis process of glycogen and starch. Specifically, in starch the CLD is bi-modal and shifted towards higher DPs as compared to glycogen [47]. Based on our preceding remarks, it could mean that starch branching enzymes are characterised by large substrate specificity lengths, corresponding to a more constrained mechanism than for glycogen. An alternative explanation for the arising of multi-modal CLDs, also based on highly constrained branching, is discussed in the Supplementary Material (see Figs 13 and 14).

2.3 Comparison to experimental data

2.3.1 Parameter calibration

Based on our simulations, it is clear that the CLD is a signature of the branching enzyme mechanisms and substrate specificity. This signature results from the combination of many different parameters. Therefore, it is particularly difficult to analyse experimental data and infer parameter values without a complementary modelling approach. To extract useful information, and compare simulations to experimental data, we designed a fitting procedure, which is detailed in the Material and Methods section 3.2. With this strategy, we are able to determine parameter values that we compare to experimental results.

GBE's mechanism and substrate specificity are incompletely characterised, yet they drastically impact on the CLD, as shown in Fig 7. Therefore, we specifically focus on this enzyme. To this end, we use experimental data obtained by Sullivan et al. [25] for mice muscles, that we extracted from their publication using the software Engauge Digitizer [48]. After purification, the granule chains are debranched using isoamylase, and their degree of polymerisation is measured using size exclusion chromatography. Our fitting procedure can be applied to any glycogen data, obtained from any specie and tissue. The data by Sullivan et al. [25] present two major advantages for our study. First, they are quantitative measurements of good resolution. Second, the authors investigated the case of a defective GBE.

Fig 8 shows the heat-map containing the best fit we obtained with the model. Extensive scans of the parameter space (see Figs 15 and 16 in the Supplementary Material) have shown that best fits can be obtained for $L^{\text{GBE}}_{\text{transferred}} = L^{\text{GBE}}_{\text{leftover}} = 3$. We therefore fix these two values in the following analysis. On Fig 8, the parameter $L^{\text{GBE}}_{\text{spacing}}$ is represented on the Y-axis, ranging from 1 to 6. The elongation to branching ratio Γ is varied on the X-axis. We notice that only low values for $L^{\text{GBE}}_{\text{spacing}}, L^{\text{GBE}}_{\text{leftover}}$ and $L^{\text{GBE}}_{\text{transferred}}$ allow to fit the experimental data. However, one should remember that these are minimal lengths and that the positions at which GBE is able to cleave and branch is flexible beyond these minimal lengths (see Fig 4). Overall, not only have we ran our fitting procedure using both the flexible and the strict location branching models, but also considering distinct values for ρ (see Figs 15 and 16 in the Supplementary Material). Interestingly, the best fit is obtained for $\rho = 0.65$ nm, which reflects the realistic size of a glucose unit inside an helical chain. This highlights the role of the steric hindrance to mimic the granule's structural properties. The best fit is obtained for the following set of parameters:

$$\{\Gamma = 0.6, \quad L^{\text{GBE}}_{\text{spacing}} = 1, \quad L^{\text{GBE}}_{\text{leftover}} = L^{\text{GBE}}_{\text{transferred}} = 3\}.$$

Importantly, the parameter values for the branching enzyme inferred from the best fit are distinct from those reported in the field [33, 49, 50], especially for the typical spacing observed between two branches [33]. Also, based on our results, GBE is able to transfer less than 4 glucose units. Knowing that GS's chain length specificity has to follow the same rule, this questions the commonly assumed value of DP4 as the minimal length that can be elongated by GS.

In Fig 9, the CLD for the best fit is shown, together with simulation results for parameter values typically reported in the experimental literature, i.e. $\{L^{\text{GBE}}_{\text{spacing}} = 4, \ L^{\text{GBE}}_{\text{leftover}} = 4, L^{\text{GBE}}_{\text{transferred}} = 4\}$. For the latter, we observe a plateau from DP4 to DP10, while experimental data show a peak between DP6 and DP8. Additionally, longer chains (DP ≥ 15) are over-represented in our results. Noticeably, for this set of GBE minimal lengths, our model is not able to reproduce the experimental data by Sullivan and coworkers [25], even when varying the elongation to branching ratio Γ (see Fig 15 in the Supplementary Material).

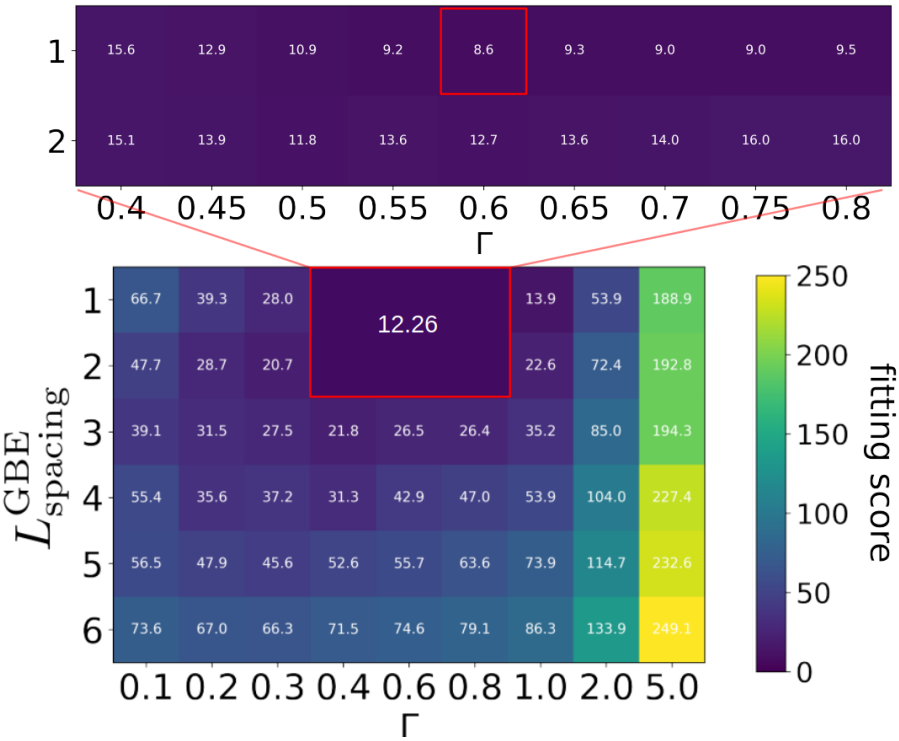

Fig 8. Heat-map showing fitting scores for various sets of parameters. The Y-axis shows $L^{\text{GBE}}_{\text{spacing}}$ ranging from 1 to 6. The X-axis shows the elongation to branching ratio Γ ranging from 0.1 to 5.0. A given cell corresponds to a parameter set $\{\Gamma, L^{\text{GBE}}_{\text{spacing}}, L^{\text{GBE}}_{\text{leftover}}, L^{\text{GBE}}_{\text{transferred}}, \rho\}$. Additional parameter sets are tested around good scores, i.e. the resolution on the elongation to branching ratio is increased, as well as the number of runs averaged. This area is surrounded by a red rectangle in which the average score is 12.26. Fitting scores are ranging from 8.6 to 369.0. The best score is 8.6 (red square in the inset heat-map) which corresponds to $\{\Gamma = 0.6, L^{\text{GBE}}_{\text{spacing}} = 1, L^{\text{GBE}}_{\text{transferred}} = L^{\text{GBE}}_{\text{leftover}} = 3, \rho = 0.65 \text{ nm}\}$.

2.3.2 Glycogen structure using the fitted parameters

In this section, unless otherwise specified, we assume that GBE chain length specificities are set to those of the best fit $\{\Gamma = 0.6, \quad L^{\text{GBE}}_{\text{spacing}} = 1, \quad L^{\text{GBE}}_{\text{leftover}} = L^{\text{GBE}}_{\text{transferred}} = 3\}$. With these parameters, we simulate the synthesis of glycogen granules, and compute their structural features and macroscopic characteristics (see Table 1). For each of those, the average values and standard errors are calculated over 30 granules with $N = 50{,}000$ glucose units. The number of non-reducing ends (noted N_{NREs}) is equal to the total number of chains, since there is one non-reducing end per chain.

Table 1. Summary of the granule structural features for the best fit parameters

	Structural features					**Macroscopic characteristics**	
	N_{NREs}	A:B	CL	Branching degree	Generation	Ω	R_g
$\Gamma = 0.6$	4135.1 ± 57.6	0.98 ± 0.02	12.09 ± 0.17	0.0901 ± 0.0014	21.8 ± 1.4	0.521 ± 0.012	$19.42 \text{ nm} \pm 0.15$
$\Gamma = 50.0$	430.9 ± 21.5	1.03 ± 0.06	116.32 ± 5.95	0.0087 ± 0.0004	11.3 ± 1.3	0.025 ± 0.007	$54.83 \text{ nm} \pm 7.69$

For $\Gamma = 0.6$, we find that on average, granules are made of 4,136 chains, with a

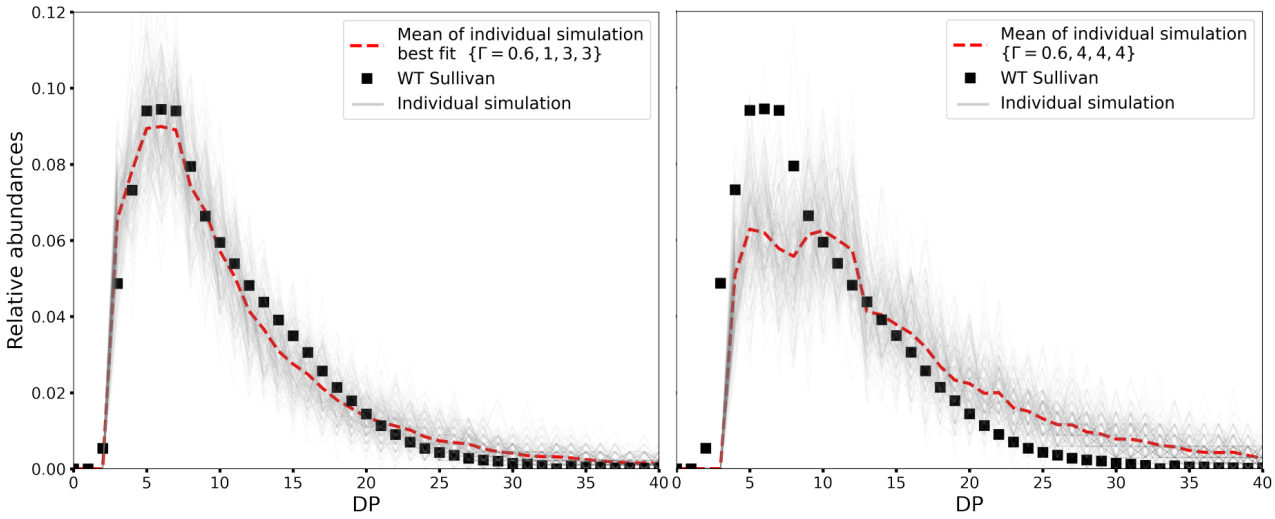

Fig 9. Comparison of simulated *versus* experimental CLDs. Experimental data are from Sullivan and coworkers [25] (black squares). In each simulation run 50,000 glucose units are incorporated in the growing granule (grey line). The average over 200 runs is represented as a red dotted line. **Left:** The CLD for the best fitting score ($\mathcal{S} = 8.6$) is obtained with $\{\Gamma = 0.6$, $L^{\text{GBE}}_{\text{spacing}} = 1$, $L^{\text{GBE}}_{\text{leftover}} = 3$, $L^{\text{GBE}}_{\text{transferred}} = 3$, $\rho = 0.65$ nm$\}$. Our best fit almost perfectly captures the experimental CLD. **Right:** CLD using parameter values typically assumed in the literature $\{\Gamma = 0.6$, $L^{\text{GBE}}_{\text{spacing}} = 4$, $L^{\text{GBE}}_{\text{leftover}} = 4$, $L^{\text{GBE}}_{\text{transferred}} = 4$, $\rho = 0.65$ nm$\}$ [49,50]. The simulated CLD differs a lot from the experimental CLD, with under-representation of small DPs, and over-representation of high DPs.

chain length of 12.09 glucoses. These two quantities are intrinsically linked since the total number of glucoses is fixed to $N = 50{,}000$. The higher the average chain length (noted CL), the lower the total number of non-reducing ends, since $CL \cdot N_{\mathrm{NREs}} = N$. The A:B ratio is found to be 0.98. With $N_{\mathrm{NREs}} = 4{,}135$, it means that ca. 2,047 chains are of type A, while 2,088 are of type B. The branching degree is found to be 9.01%. This is in good agreement with the range typically reported in glycogen (i.e. [5-10 %] [9, 29]), while starch exhibits much lower values (typically between 1 and 5% [26–28]). The average occupancy is $\Omega = 0.521$, which means that half of the total granule volume is filled with glucoses. The gyration radius is $R_g = 19.42$ nm which is consistent with radii reported for big granules (molecular weight ca. $10^7 \mathrm{g.mol}^{-1}$), like the ones considered here ($N = 50{,}000$) [13, 51].

It is interesting to notice that ca. 21 nm is the upper limit for the radius of a glycogen granule of ca. 50,000 glucoses, when considering the fractal structure depicted by Meléndez et al. [6]. Thus, with our approach, that is not based on a fractal structure, we determine a gyration radius close to their results. However, the glycogen structures we simulate are deeply distinct from theirs. In our simulations, we observe that the density is approximately constant with the radius, while a fractal glycogen granule has a density that increases exponentially with the radius, so resembling a rather empty shell. Also, we show that our simulations can reproduce characteristic quantities (summarised in Table 1 for $\Gamma = 0.6$), which are in good agreement with experimental results. Thus, our model appears to provide a more realistic depiction of glycogen granules.

For comparison, in Table 1, we also report the case of $\Gamma = 50.0$, while keeping all other parameters unchanged. By increasing Γ, we force elongation over branching. As expected, less chains are created but they are longer. Consistently, the branching degree is also lower. In our model, chains are rigid rods that cannot bend, which leads to a bigger radius of gyration. Since the total content of glucose is the same like for $\Gamma = 0.6$, the occupancy Ω decreases almost proportionally. *In vivo*, glycogen granules might follow a distinct behaviour, where the presence of long chains could trigger precipitation, like for instance reported in the Lafora and GBE diseases [25, 52].

The formation of an A chain results from branching either an A or a B chain. These different events can be represented as reactions and described as follows:

$$R_{A\to AB}: \quad \mathrm{A} \xrightarrow{\mathrm{GBE}} \mathrm{B} + \mathrm{A}_{\mathrm{new}},$$
$$R_{B\to AB}: \quad \mathrm{B} \xrightarrow{\mathrm{GBE}} \mathrm{B} + \mathrm{A}_{\mathrm{new}}\ .$$

To illustrate how different branch structures correspond to specific A:B ratios, we sketch extreme cases in Fig 10. If the $R_{B\to AB}$ and $R_{A\to AB}$ reactions are equally likely, a purely probabilistic approach predicts that, on average, the A:B ratio is equal to 1, independently of the relative activities of the enzymes (see Fig 17 in the Supplementary Material). Such a ratio is for instance observed in the first two rows of Table 1, with $\mathrm{A{:}B} = 0.98 \pm 0.02$ and $\mathrm{A{:}B} = 1.03 \pm 0.08$, which is in good agreement with reported values for glycogen ($\mathrm{A{:}B} = 0.6 - 1.2$ [31]). Other closely related branched polysaccharides can exhibit different A:B ratios. For example, amylopectin has a typical A:B ratio in the range from 1.5 to 2.6 [31, 32], depending on the organism. It is therefore interesting to investigate which factors influence this ratio.

Here, we identify two effects of the interplay between the branching mechanism and the dynamics of the granule structure, highlighting once again the importance of the branching enzyme on the emerging structural patterns. First, considering a small Γ, branching dominates over elongation, and the enzyme branches as soon as it can. Thus, if in addition $L^{\mathrm{GBE}}_{\mathrm{spacing}}$ is small, the A chains are closely stacked together and the granule is tightly packed. In this regime, overall, A chains are closer to other A chains

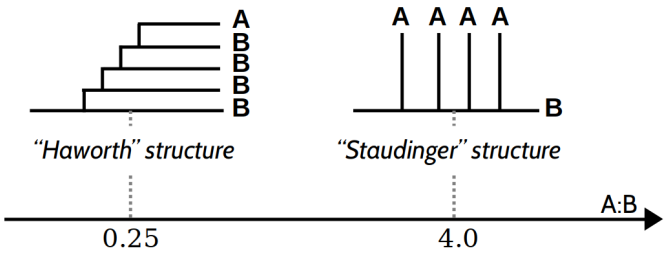

Fig 10. Two schematic branched structures with different A:B ratios. For a "Haworth"-like structure, branching reactions tend to occur on A chains, leading to a low A:B ratio, while for a "Staudinger"-like structure branching on B chains is favoured, leading to a high A:B ratio.

than to B chains. Due to the increased steric hindrance for A chains, B chains react more easily, which favours the $R_{B\to AB}$ reaction, and leaves A chains unchanged. Therefore, the A:B ratio increases, meaning that it shifts to the right along the axis in Fig 10. Second, if $L^{\mathrm{GBE}}_{\mathrm{transferred}} > L^{\mathrm{GBE}}_{\mathrm{leftover}}$, when a branching occurs, on either an A or a B chain, the newly formed A chain will on average be longer than the B (leftover) one. Therefore, the new A chain will react faster, favouring $R_{A\to AB}$, and lead to a reduction of the A:B ratio, meaning that it shifts to the left along the axis in Fig 10. We confirmed this effect for the specific case of $L^{\mathrm{GBE}}_{\mathrm{transferred}} = 7$ and $L^{\mathrm{GBE}}_{\mathrm{leftover}} = 3$. The resulting structures have an A:B ratio on average equal to 0.80 ± 0.02. If the values of the minimal lengths are exchanged, such that $L^{\mathrm{GBE}}_{\mathrm{transferred}} = 3$ and $L^{\mathrm{GBE}}_{\mathrm{leftover}} = 7$, the A:B ratio is instead equal to 2.01 ± 0.04. It appears that the $L^{\mathrm{GBE}}_{\mathrm{transferred}}$ and $L^{\mathrm{GBE}}_{\mathrm{leftover}}$ are specific mechanistic features that uniquely determine the A:B ratio. Consequently, they could cause the main structural differences observed between amylopectin (the prime constituent of starch) and glycogen.

3 Material and Methods

3.1 Simulation procedure

We present a model that simulates the dynamics of glycogen synthesis. We record all enzymatic reaction steps, the time at which they occur, and the full 3-dimensional details of the glycogen structure at any time point. The model specifically keeps track of each glucose position, the complete connectivity of the chains, and the position of each branching point. To account for this complexity, we implement the model using a stochastic algorithm. This approach also allows to specifically consider how changes in the glycogen structure enable or disable enzymatic reactions. For comparison, employing a deterministic approach, for instance based on systems of ordinary differential equations, would lead to unnecessarily complex simulation rules, that would also include additional ad-hoc assumptions. Besides, randomness has a stronger impact on a system as it involves small numbers, like the one we consider here. Indeed, the synthesis and breakdown of a single glycogen granule of ca. 50,000 glucose units involves only a small number of enzymes. For example, experiments indicate that on average a single glycogen synthase enzyme is active per granule [53]. As a result, a stochastic approach appears very natural to deal with this complex and spatially structured substrate.

We base our stochastic approach on the Gillespie algorithm [54, 55]. At each time step, it consists in randomly selecting both an enzymatic reaction and its duration, while systematically updating the structure of the glycogen molecule accordingly.

While the enzymatic reaction is selected from a uniform distribution, the time is chosen from an exponential distribution, such that the Gillespie algorithm allows to simulate the real time of the dynamics of the system, as far as the underlying enzymatic mechanisms and their kinetic parameters are known. Each reaction is associated to a specific propensity, that we choose as proportional to the concentration of enzymes and substrate, thereby following a typical mass-action kinetics.

The Gillespie algorithm accounts for all possible reactions, and keeps track of any modification of the granule structure, such that only possible reactions can be selected. Nonetheless, because of the complexity of the structure we simulate, to account for steric hindrance among monomers, we instead allow for certain reactions to be rejected. For these rejected reactions the time elapsed is not accounted.

We provide details about the Gillespie algorithm, and how it was employed to simulate the dynamic changes of the three-dimensional structure of the granule in the Supplementary Material. Moreover, the source code of our model, together with Jupyter Notebooks that recreate the main figures of this manuscript, are available on our gitlab repository (link provided below under Data Availability).

3.2 Best fit algorithm

The model contains various kinds of parameters. Some describe the physical properties of the glycogen structure, others relate to the enzymatic activity, including the enzyme substrate specificities. On the one hand, certain parameters are inferred from literature data, for instance the minimum DP for GS to act ($L_{\text{min}}^{\text{GS}}$). On the other hand, other parameters are free to be fitted by our in-house designed best fit algorithm. Here we choose to fit the minimum DP between two branches after a branching reaction ($L_{\text{spacing}}^{\text{GBE}}$), the minimum DP between the non-reducing end of the mother branch and the new branching point ($L_{\text{leftover}}^{\text{GBE}}$), and the ratio between the elongation and branching reaction rates (Γ).

The best fit algorithm consists in setting bounding ranges for the parameters to be fitted, and systematically scanning the parameter space thereby defined. For each parameter set tested, we run 50 simulations, take the average of the resulting CLDs, and compare it to our targeted experimental data set. For each parameter set, the comparison to the experimental data is measured by the score $\mathcal{S}$, defined as follows:

$$\mathcal{S} = \sum_{\text{DP}=1}^{n} \left(\mathcal{A}_{\text{exp}}(\text{DP}) - \bar{\mathcal{A}}_{\text{sim}}(\text{DP})\right)^2,$$

where $\mathcal{A}_{\text{exp}}(\text{DP})$ is the experimental CLD abundance for a given DP and $\bar{\mathcal{A}}_{\text{sim}}(\text{DP})$ the average simulated CLD abundance for a given DP. Therefore, $\mathcal{S}$ measures the difference between the average simulated and the experimental CLDs. The best fit is found for the parameter set that has the lowest score $\mathcal{S}$.

4 Conclusion

4.1 Discussion of the results

According to the World Health Organization, metabolic diseases are a burden in western countries. Among them, Glycogen Storage Diseases (GSDs), Lafora disease, Adult Polyglucosan Body Disease (APBD), and even diabetes directly or indirectly involve glycogen. Investigating the regulation of glycogen's structural properties could therefore strongly contribute to further understand such diseases. Computational models that encompass the complex interplay of both glycogen's structure and its metabolism allow to tackle this challenge but remain poorly exploited.

In this article, we present a stochastic structural model for glycogen synthesis. The model provides a precise description of both the structure of glycogen in 3 dimensions, that we can visualise, and the detailed dynamics and mechanistics of the underlying enzymatic reactions. For instance, both elongation and branching have precise rules regarding their substrate specificities. Modelling glycogen granules in 3 dimensions is made possible by a coarse-grained geometrical description at the glucose level, allowing us to track all the glucose units in space. This description also accounts for the steric hindrance effects resulting from the impossibility for the chains to overlap. Simulating reaction dynamics relies on a Gillespie algorithm, which determines when and where a reaction of branching or elongation takes place, thereby dictating the corresponding change in the 3-dimensional structure.

We show qualitatively how enzyme activity affects glycogen structure for generic sets of parameters. We highlight two different synthesis regimes, depending on Γ, the ratio between the elongation and branching reaction rates. By varying this ratio, either small and dense, or big and sparse granules can be simulated. Still, it can be expected that Γ depends on numerous factors, such as the organism under investigation, the cell type, and possibly the external conditions. In addition, a model that would consider chains bending and intermolecular interactions might lead to more complex results, in particular during the synthesis of sparse granules. In our results, the phenomenology of the two synthesis regimes is also confirmed by the occupancy profiles along the radius of the corresponding granules. In both synthesis regimes, first the center of the granule is filled around GN, before reaching a critical density preventing further internal reactions to occur due to steric hindrance. Then, the granule expands such that the density remains approximately constant within the sphere defined by the gyration radius. While this result is a consequence of the geometrical assumptions of our model, it is interesting to note that a glycogen fractal description would instead give a density exponentially increasing with the granule radius. Besides our own results and the various arguments exhibited against a fractal representation of glycogen, we would like to highlight here that, as soon as one considers even a single degree of freedom in the glycogen branching reaction, it would rapidly lead to the loss of any "fractal-like" structure. Such degrees of freedom are necessarily present in natural conditions. For instance, the dihedral angles defining the $\alpha - 1,6$ bonds may take various values, making a fractal pattern very unlikely *in vivo*. Beyond these spatial considerations, the model establishes a natural and clear connection between enzymes' reaction rates, and both the degree of polymerisation of the chains, and the number of non-reducing ends.

Our results show that the chain length distribution (CLD) of glycogen is highly sensitive to the branching reaction, predominantly its mechanism. Each of the three characteristic minimal lengths of the reaction has specific effects on the CLD. Additionally, if any of them increases, less branching outcomes are possible, eventually leading to a bi-modal or even multi-modal distribution. This effect is enhanced by high branching activity. When varied together, these minimal lengths show even more complex imprints on the CLD. In contrast, multi-modal distributions become less pronounced by increasing the elongation reaction rate, because rapid elongation increases the number of possible configuration outcomes. Additionally, increasing elongation leads to longer chains and results in a CLD spreading towards higher DPs. Altogether, we show that the CLD, and in particular peaks location and intensity, are subtly affected by several complex effects.

Guided by these findings, we propose to consider the CLD not only as an important structural feature of glycogen, but also as a signature of GBE. Thus, fitting experimental data with our model arises as a natural strategy to infer knowledge on the GBE mechanism. Not only did we illustrate the strength of our fitting procedure

on an experimental data set from Sullivan et al. [25], but we also extracted several macroscopic characteristic features from the resulting fit, which we compared to various literature sources, thereby confirming the validity of our results. Using this fit, at the microscopic level, we were able to discriminate between the two branching models we hypothesised, and selected the flexible location branching over the strict one. Besides, we could critically evaluate parameter values typically reported in the literature. For instance, it is often assumed that GBE transfers branches of DP 4 or 6, or even longer [33, 49, 50]. Similarly, it is typically reported that around 6 glucose units space two branches. Within the framework of our model and its underlying assumptions, it is impossible to reproduce CLDs of *in vivo* glycogen using the above mentioned values. Instead, our fitting procedure suggests that a high flexibility is necessary for both the branching mechanism and the minimal lengths involved. This finding fully confirms the importance of modelling glycogen synthesis using a stochastic approach.

Moreover, our customisable branching model shows that the A:B ratio is independent of the kinetic parameter Γ, but specifically determined by the difference between $L^{\mathrm{GBE}}_{\mathrm{transferred}}$ and $L^{\mathrm{GBE}}_{\mathrm{leftover}}$. Based on these observations, we hypothesise that the branching mechanism is chiefly responsible for the structural differences observed between starch and glycogen.

Throughout this study, our coarse-grained approach accounts for the contribution of individual glucoses to the overall granule structure, by considering them as spheres of radius 0.65 nm. It is interesting to notice that if we set the glucose volume to 0 nm^3, the CLD remains almost unchanged while other macroscopic properties of the granule, such as its overall volume, are dramatically impacted. This observation confirms once more that the glycogen CLD is primarily shaped by the enzyme mechanisms.

4.2 Outlook

It is important to keep in mind that our model contains limitations that may be circumvented by further refining the model assumptions. In our model, the simplified coarse-grained representation of glucoses assumes that all of them are arranged in single helices. This hypothesis implies that all $\alpha - 1,4$ glycosidic bonds have the same angle values. Yet, *in vivo*, this is highly unlikely, and instead, the dihedral angles of the $\alpha - 1,4$ bonds should be able to take various values. To take this into account, we could randomly pick the dihedral angles of the $\alpha - 1,4$ bonds using the Ramachandran plots of their energetically favourable regions. As a first trial, we could use that of maltose, that is well characterised [56]. We expect that, introducing such disorder in the angles, chains will appear longer. Besides, in the model, we consider that all chains are stiff. Thus, to improve the macroscopic representation of the chains, we could introduce the possibility that they bend when encountering steric hindrance. To do so, we would minimise their torsion energy, like it is done in polymer physics models [57]. Opposite to the change suggested for the $\alpha - 1,4$ bonds variability, accounting for the flexibility of the chains might lead to a higher granule density, and thus, a lower radius. The fact that these two effects might cancel each other, possibly explains why our simplified model is nonetheless able to capture realistic *in vivo* granule radii. Besides, in abnormal conditions, the potential formation of double helices may not only lead to glycogen precipitation, but also prevent enzymatic reactions. Thus, a later improvement of the model could include to tune the enzymes' reaction rate depending on the substrate chain configuration and length.

Throughout this article, we focus on glycogen synthesis, yet, simulating the degradation dynamics with the algorithms we developed would be straightforward. We expect residual degradation activity to only lightly modify the effective elongation to branching ratio Γ, and slow down the synthesis. In such a case, the CLD would be

slightly shifted to the left, and the first chain length detected would be the smallest value between $L^{\mathrm{GBE}}_{\mathrm{transferred}}$ and its GP counterpart. It could be particularly interesting to account for the degrading enzymes, beyond a residual activity, in case the synthesis is defective. Indeed, abnormal structures produced by defective synthesis enzymes could be degraded, and thereby corrected, by degrading enzymes. For instance, if $L^{\mathrm{GBE}}_{\mathrm{transferred}}$ becomes too short for GS to elongate the newly formed chain, GDE could unbranch the latter and thereby preserve the macroscopic properties of the granule. Last, following an approach analogous to the one taken in this article, one could choose to investigate the glycogen granules' breakdown in full depth, by first synthesising granules and then proceeding to their degradation. Although, for the sake of simplicity, we would then uncouple in time the synthesis and degradation processes, it would still be very interesting to study the mutual impact of distinct modes of synthesis and degradation on the overall glucose release and fixation.

In this article, we show that the availability of the substrate strongly influences the enzyme activity, leading to distinct chain lengths and number of non-reducing ends. Noticeably, other modelling approaches, such as kinetic models using systems of ordinary differential equations (ODEs), instead consider glycogen as a single metabolite, approximated by the sum of all glucose units that compose it. A direct consequence is that any structural aspects are neglected and these models cannot differentiate between a single chain of 50,000 glucoses, and an actual granule of the same weight. Still, for instance, the number of non-reducing ends available for elongation are drastically different in these two cases. Thus, coupling our model to glycogen metabolic ODE models, would constitute a hybrid approach that would include key structural details, while enlarging its biochemical scope. It would thereby open up a whole new range of modelling possibilities. For instance, it would allow to investigate the interplay between glycogen structure and the evolution in time of important metabolites, under various physiological conditions, including diseases scenarios. Using this approach, we shall be able to characterise the phenomenology of each glycogen storage disease, with a focus on the role of glycogen structure, and address questions such as glycogen accumulation, glucose cycling, glucose homeostasis, and even glycogen precipitation.

Data availability statement

The source code and data used to produce the results and analyses presented in this manuscript are available on a Gitlab repository at `https://gitlab.com/qtb-hhu/models/Stochastic-modeling-of-a-three-dimensional-glycogen-granule`. The version associated to this publication is tagged as "v1-updated".

Acknowledgments

This paper is supported by European Union's Horizon 2020 research and innovation program under the Marie Skłodowska-Curie grant agreement PoLiMeR, No 812616. O.E. and A.R. are supported by the Deutsche Forschungsgemeinschaft (DFG) under Germany's Excellence Strategy EXC 2048/1, Project ID: 390686111. The current position of A.R. is funded by the German federal and state programme Professorinnenprogramm III for female scientists.

References

1. Berg JM, Tymoczko JL, jr GJG, Stryer L. Stryer Biochemie. Springer Spektrum Berlin, Heidelberg; 2018.

2. Manners DJ. Recent developments in our understanding of glycogen structure. Carbohydrate Polymers. 1991;16(1):37–82. doi:https://doi.org/10.1016/0144-8617(91)90071-J.

3. Meyer KH, Bernfeld P. Recherches sur l'amidon V. L'amylopectine. Helvetica Chimica Acta. 1940;23(1):875–885. doi:https://doi.org/10.1002/hlca.194002301112.

4. Larner J, Illingworth B, Cori GT, Cori CF. Structure of Glycogens and Amylopectins: II. Analysis by Stepwise Enzymatic Degradation*. Journal of Biological Chemistry. 1952;199(2):641–651. doi:https://doi.org/10.1016/S0021-9258(18)38502-8.

5. Gunja-Smith Z, Marshall JJ, Mercier C, Smith EE, Whelan WJ. A revision of the Meyer-Bernfeld model of glycogen and amylopectin. FEBS Letters. 1970;12(2):101–104. doi:https://doi.org/10.1016/0014-5793(70)80573-7.

6. Meléndez-Hevia E, Waddell TG, Shelton ED. Optimization of molecular design in the evolution of metabolism: the glycogen molecule. Biochemical Journal. 1993;295(2):477–483. doi:10.1042/bj2950477.

7. Wanson JC, Drochmans P. Rabbit Skeletal Muscle Gglycogen. The Journal of Cell Biology. 1968;38(1):130–150.

8. Mordoh J, Krisman CR, Leloir LF. Further studies on high molecular weight liver glycogen. Archives of Biochemistry and Biophysics. 1966;113(2):265–272. doi:https://doi.org/10.1016/0003-9861(66)90186-X.

9. Edstrom RD. Structure of a Low Molecular Weight Form of Glycogen Isolated from the Liver in a Case of Glycogen Storage Disease. Journal of Biological Chemistry. 1972;247(5):1360–1367. doi:https://doi.org/10.1016/S0021-9258(19)45568-3.

10. Roach P, Depaoli-Roach A, Hurley T, Tagliabracci V. Glycogen and its metabolism: some new developments and old themes. Biochemical Journal. 2012;441(3):763–787. doi:10.1042/BJ20111416.

11. Pitcher J, Smythe C, Campbell DG, Cohen P. Identification of the 38-kDa subunit of rabbit skeletal muscle glycogen synthase as glycogenin. European Journal of Biochemistry. 1987;169(3):497–502. doi:https://doi.org/10.1111/j.1432-1033.1987.tb13637.x.

12. Thon VJ, Khalil M, Cannon JF. Isolation of human glycogen branching enzyme cDNAs by screening complementation in yeast. Journal of Biological Chemistry. 1993;268(10):7509–7513. doi:https://doi.org/10.1016/S0021-9258(18)53204-X.

13. Drochmans P. Morphologie du glycogène: Etude au microscope électronique de colorations négatives du glycogène particulaire. Journal of Ultrastructure Research. 1962;6(2):141 – 163. doi:https://doi.org/10.1016/S0022-5320(62)90050-3.

14. Marchand I, Chorneyko K, Tarnopolsky M, Hamilton S, Shearer J, Potvin J, et al. Quantification of subcellular glycogen in resting human muscle: granule size, number, and location. Journal of Applied Physiology. 2002;93(5):1598–1607. doi:10.1152/japplphysiol.00585.2001.

15. Roach P, Depaoli-Roach A, Hurley T, Tagliabracci V. Glycogen and its metabolism: some new developments and old themes. Biochemical Journal. 2012;441(3):763–787. doi:10.1042/BJ20111416.

16. Besford QA, Sullivan MA, Zheng L, Gilbert RG, Stapleton D, Gray-Weale A. The structure of cardiac glycogen in healthy mice. International Journal of Biological Macromolecules. 2012;51(5):887–891. doi:https://doi.org/10.1016/j.ijbiomac.2012.06.037.

17. Prats C, Graham TE, Shearer J. The dynamic life of the glycogen granule. Journal of Biological Chemistry. 2018;293(19):7089–7098. doi:https://doi.org/10.1074/jbc.R117.802843.

18. Whelan WJ. Pride and prejudice: The discovery of the primer for glycogen synthesis. Protein Science. 1998;7(9):2038–2041. doi:https://doi.org/10.1002/pro.5560070921.

19. Smythe C, Cohen P. In: Christen P, Hofmann E, editors. The discovery of glycogenin and the priming mechanism for glycogen biogenesis. Berlin, Heidelberg: Springer Berlin Heidelberg; 1992. p. 149–155. Available from: `https://doi.org/10.1007/978-3-642-77200-9_12`.

20. Roach PJ, Skurat AV. Self-Glucosylating Initiator Proteins and Their Role in Glycogen Biosynthesis. vol. 57 of Progress in Nucleic Acid Research and Molecular Biology. Academic Press; 1997. p. 289–316. Available from: `https://www.sciencedirect.com/science/article/pii/S0079660308602846`.

21. Zeqiraj E, Tang X, Hunter RW, García-Rocha M, Judd A, Deák M, et al. Structural basis for the recruitment of glycogen synthase by glycogenin. Proceedings of the National Academy of Sciences. 2014;111:E2831 – E2840.

22. Ggunja-Smith Z, Marshall JJ, Smith EE. Enzymatic determination of the unit chain length of glycogen and related polysaccharides. FEBS Letters. 1971;13(5):309–311. doi:10.1016/0014-5793(71)80248-X.

23. Wang L, Wise MJ. Glycogen with short average chain length enhances bacterial durability. Naturwissenschaften. 2011;98(9):719. doi:10.1007/s00114-011-0832-x.

24. Yoo SH, Lee BH, Moon Y, Spalding MH, Jane Jl. Glycogen Synthase Isoforms in Synechocystis sp. PCC6803: Identification of Different Roles to Produce Glycogen by Targeted Mutagenesis. PLOS ONE. 2014;9(3):1–9. doi:10.1371/journal.pone.0091524.

25. Sullivan MA, Nitschke S, Skwara EP, Wang P, Zhao X, Pan XS, et al. Skeletal Muscle Glycogen Chain Length Correlates with Insolubility in Mouse Models of Polyglucosan-Associated Neurodegenerative Diseases. Cell Reports. 2019;27(5):1334–1344.e6. doi:https://doi.org/10.1016/j.celrep.2019.04.017.

26. Feng L, Fawaz R, Hovde S, Gilbert L, Chiou J, Geiger JH. Crystal Structures of Escherichia coli Branching Enzyme in Complex with Linear Oligosaccharides. Biochemistry. 2015;54(40):6207–6218. doi:10.1021/acs.biochem.5b00228.

27. Nilsson GS, Gorton L, Bergquist KE, Nilsson U. Determination of the Degree of Branching in Normal and Amylopectin Type Potato Starch with 1H-NMR Spectroscopy Improved resolution and two-dimensional spectroscopy. Starch - Stärke. 1996;48(10):352–357. doi:https://doi.org/10.1002/star.19960481003.

28. Buléon A, Colonna P, Planchot V, Ball S. Starch granules: structure and biosynthesis. International Journal of Biological Macromolecules. 1998;23(2):85–112. doi:https://doi.org/10.1016/S0141-8130(98)00040-3.

29. Tolmasky DS, Krisman CR. The degree of branching in (1,4)-(1,6)-linked glucopolysaccharides is dependent on intrinsic properties of the branching enzymes. European Journal of Biochemistry. 1987;168(2):393–397. doi:https://doi.org/10.1111/j.1432-1033.1987.tb13432.x.

30. Peat S, Whelan WJ, Thomas GJ. 587. The enzymic synthesis and degradation of starch. Part XXII. Evidence of multiple branching in waxy-maize starch. A correction. J Chem Soc. 1956; p. 3025–3030. doi:10.1039/JR9560003025.

31. Marshall JJ, Whelan WJ. Multiple branching in glycogen and amylopectin. Archives of Biochemistry and Biophysics. 1974;161(1):234 – 238. doi:https://doi.org/10.1016/0003-9861(74)90256-2.

32. Takeda Y, Hizukuri S, Takeda C, Suzuki A. Structures of branched molecules of amyloses of various origins, and molar fractions of branched and unbranched molecules. Carbohydrate Research. 1987;165(1):139–145. doi:https://doi.org/10.1016/0008-6215(87)80089-7.

33. Goldsmith E, Sprang S, Fletterick R. Structure of maltoheptaose by difference Fourier methods and a model for glycogen. J Mol Biol. 1982;156(2):411–427. doi:10.1016/0022-2836(82)90336-9.

34. Meléndez R, Meléndez-Hevia E, Canela EI. The Fractal Structure of Glycogen: A Clever Solution to Optimize Cell Metabolism. Biophysical Journal. 1999;77(3):1327–1332. doi:https://doi.org/10.1016/S0006-3495(99)76982-1.

35. Rolland-Sabaté A, Mendez-Montealvo MG, Colonna P, Planchot V. Online Determination of Structural Properties and Observation of Deviations from Power Law Behavior. Biomacromolecules. 2008;9(7):1719–1730. doi:10.1021/bm7013119.

36. Besford Q, Zeng XY, Ye JM, Gray-Weale A. Liver glycogen in type 2 diabetic mice is randomly branched as enlarged aggregates with blunted glucose release. Glycoconjugate journal. 2016;33. doi:10.1007/s10719-015-9631-5.

37. Kim D, Duhamel J. Interior of glycogen probed by pyrene excimer fluorescence. Carbohydrate Polymers. 2023;299:120205. doi:https://doi.org/10.1016/j.carbpol.2022.120205.

38. Zhang P, Nada SS, Tan X, Deng B, Sullivan MA, Gilbert RG. Exploring glycogen biosynthesis through Monte Carlo simulation. International Journal of Biological Macromolecules. 2018;116:264–271. doi:https://doi.org/10.1016/j.ijbiomac.2018.05.027.

39. Miller NC, Toffolo RW, McAuley KB, McLellan PJ. Determining Polymer Chain Length Distributions Using Numerical Inversion of Laplace Transforms. Polymer Reaction Engineering. 1996;4(4):279–301. doi:10.1080/10543414.1996.10744477.

40. Mulders K, Beeftink H. Chain length distribution and kinetic characteristics of an enzymatically produced polymer. e-Polymers. 2013;13. doi:10.1515/epoly-2013-0124.

41. Gray-Weale A, Gilbert RG. General description of the structure of branched polymers. Journal of Polymer Science Part A: Polymer Chemistry. 2009;47(15):3914–3930. doi:https://doi.org/10.1002/pola.23458.

42. Wu AC, Gilbert RG. Molecular Weight Distributions of Starch Branches Reveal Genetic Constraints on Biosynthesis. Biomacromolecules. 2010;11(12):3539–3547. doi:10.1021/bm1010189.

43. Whelan WJ. Pride and prejudice: The discovery of the primer for glycogen synthesis. Protein Science. 1998;7(9):2038–2041. doi:https://doi.org/10.1002/pro.5560070921.

44. Hinz HJ. Thermodynamic Data for Biochemistry and Biotechnology, 1st Edition. Springer Berlin, Heidelberg; 1986.

45. Palm DC, Rohwer JM, Hofmeyr JHS. Regulation of glycogen synthase from mammalian skeletal muscle – a unifying view of allosteric and covalent regulation. The FEBS Journal. 2013;280(1):2–27. doi:https://doi.org/10.1111/febs.12059.

46. O'Sullivan AC, Perez S. The relationship between internal chain length of amylopectin and crystallinity in starch. Biopolymers. 1999;50(4):381–390. doi:https://doi.org/10.1002/(SICI)1097-0282(19991005)50:4<381::AID-BIP4>3.0.CO;2-W.

47. Jane J, Chen YY, Lee LF, McPherson AE, Wong KS, Radosavljevic M, et al. Effects of Amylopectin Branch Chain Length and Amylose Content on the Gelatinization and Pasting Properties of Starch. Cereal Chemistry. 1999;76(5):629–637. doi:https://doi.org/10.1094/CCHEM.1999.76.5.629.

48. Mitchell M, Muftakhidinov B, Winchen T, Wilms A, van Schaik B, badshah400, et al.. markummitchell/engauge-digitizer: Nonrelease; 2020. Available from: `https://doi.org/10.5281/zenodo.3941227`.

49. Voet D, Voet JG. Biochemistry, 4th Edition. Wiley; 2011.

50. Froese DS, Michaeli A, McCorvie TJ, Krojer T, Sasi M, Melaev E, et al. Structural basis of glycogen branching enzyme deficiency and pharmacologic rescue by rational peptide design. Human Molecular Genetics. 2015;24(20):5667–5676. doi:10.1093/hmg/ddv280.

51. Sullivan MA, Vilaplana F, Cave RA, Stapleton D, Gray-Weale AA, Gilbert RG. Nature of and Particles in Glycogen Using Molecular Size Distributions. Biomacromolecules. 2010;11(4):1094–1100. doi:10.1021/bm100074p.

52. Nitschke F, Sullivan MA, Wang P, Zhao X, Chown EE, Perri AM, et al. Abnormal glycogen chain length pattern, not hyperphosphorylation, is critical in Lafora disease. EMBO Molecular Medicine. 2017;9(7):906–917. doi:https://doi.org/10.15252/emmm.201707608.

53. Smythe C, Watt P, Cohen P. Further studies on the role of glycogenin in glycogen biosynthesis. European Journal of Biochemistry. 1990;189(1):199–204. doi:https://doi.org/10.1111/j.1432-1033.1990.tb15477.x.

54. Gillespie DT. A general method for numerically simulating the stochastic time evolution of coupled chemical reactions. Journal of Computational Physics. 1976;22(4):403–434. doi:https://doi.org/10.1016/0021-9991(76)90041-3.

55. Gillespie DT. Exact stochastic simulation of coupled chemical reactions. The Journal of Physical Chemistry. 1977;81(25):2340–2361. doi:10.1021/j100540a008.

56. Gallant D, Bouchet B, Buléon A, Pérez S. Physical characteristics of starch granules and susceptibility to enzymatic degradation. European journal of clinical nutrition. 1992;46 Suppl 2:S3–16.

57. Pethrick R. Polymer physics. Edited by Michael Rubinstein and Ralph H Colby Oxford University Press, Oxford, 2003. ISBN 019852059X. pp 440. Polymer International. 2004;53(9):1394–1395. doi:https://doi.org/10.1002/pi.1472.

5 Supplementary materials

5.1 Symbols definition

α and β	Different forms of glycogen granules as described by [13]. The first one being aggregates of the second.
ρ	Radius of a monomer unit described as a sphere in the model.
l	Distance between two overlapping spheres.
N	Number of glucose units incorporated in a granule.
$L^{\text{GBE}}_{\text{spacing}}$	Minimum distance (in glucose units) between two chains after branching.
$L^{\text{GBE}}_{\text{transferred}}$	Minimum chain length (in glucose units) cleaved and transferred.
$L^{\text{GBE}}_{\text{leftover}}$	Minimum distance (in glucose units) between the non-reducing end of the mother chain and the newly created branching point.
Γ	Ratio of the elongation and branching reaction rates.
R_g	Radius of gyration (i.e. radius of a glycogen granule in nm) $$R_g = \sqrt{\frac{1}{N}\sum_{i=1}^{N}(\overrightarrow{r_i} - \overrightarrow{r_{\text{mean}}})^2},$$ where N is the total number of glucose units in the granule, r_i the spatial coordinate of the i^{th} glucose unit, and r_{mean} that of the center of mass of the granule.
$A:B$	Ratio of the number of A chains (that do not carry any daughter chain) over B chains (that carry at least one).
V_{glucose}	Volume effectively occupied by all the glucose units incorporated in the granule.
V_{tot}	Volume of a glycogen granule (i.e. $4/3 \cdot \pi \cdot {R_g}^3$).
Ω	Occupancy, defined as the ratio of V_{glucose} and V_{tot}.
$\mathcal{S}$	Score to measure the distance between simulated and experimental CLDs, defined as the mean square difference between the two curves.

5.2 Abbreviations list

GS	Glycogen synthase
GN	Glycogenin
GBE	Glycogen branching enzyme
NREs	Number of non-reducing ends
DP	Degree of polymerisation
A:B	Ratio between the number of A chains and the number of B chains
CL	Chain length
ODE	Ordinary differential equation

5.3 Numerical procedures

5.3.1 A brief reminder of the Gillespie "direct" method

In this subsection, we summarise the main steps of the so-called "direct" Gillespie method that has been introduced by Daniel Gillespie in 1976 [54] and further popularised in 1977 [55]. It consists in splitting the probability:

$$\mathcal{P}(\mu,\tau) = \text{Probability for the next reaction to be the reaction } \mathcal{R}_\mu \text{ and to occur between time } \tau \text{ and } \tau + d\tau,$$

into two terms, using conditional probability. $\mathcal{P}(\mu,\tau)$ can be written as follows:

$$\mathcal{P}(\mu,\tau) = \mathcal{P}(\tau) \cdot \mathcal{P}(\mu|\tau) \tag{3}$$

where $\mathcal{P}(\tau)$ is the probability that no reaction occurs between $t = 0$ and $t = \tau$, and $\mathcal{P}(\mu|\tau)$ is the probability that, given that a reaction occurs during the infinitesimal time τ and $\tau + d\tau$, it is the reaction $\mathcal{R}_\mu$. Gillespie showed that:

$$\mathcal{P}(\tau) = a \cdot \exp(-a\tau), \tag{4}$$

where a is the sum of all propensities a_μ, and $a_\mu \cdot dt$ is the probability for the reaction μ to occur in the next infinitesimal time interval dt. Additionally,

$$\mathcal{P}(\mu|\tau) = \frac{a_\mu}{a}. \tag{5}$$

The idea underpinning the algorithm is to draw random numbers that determine the next reaction and reaction time, according to these probabilities. (Eq 4) can be integrated to get the probability distribution function $F(t)$ from the probability density function $\mathcal{P}(t)$.

$$F(t) = \int_0^t \mathcal{P}(t') \cdot dt', \tag{6}$$

such that any $F(t_0)$ is the probability that t is less than t_0. Using the inversion generating method described in details in the Appendix of [55], Gillespie showed that the time τ can be generated using a pseudo-random number r_1 following a uniform distribution in $[0, 1]$, such that:

$$\tau = \frac{1}{a} \cdot \ln(1/r_1). \tag{7}$$

To determine which reaction takes place at time τ, one draws a second pseudo-random number r_2, and looks in which of the following intervals it falls. For N possible reactions, the reaction $\mu \in (1, .., N)$ will be selected if:

$$\sum_{j=1}^{\mu} \frac{a_j}{a} \leq r_2 < \sum_{j=1}^{\mu+1} \frac{a_j}{a}. \tag{8}$$

The couple $\{\mu,\tau\}$ determines the reaction μ and the increment of time $t \leftarrow t + \tau$ to be implemented.

The propensities of the distinct reactions of the system are computed based on rate laws. As a first assumption, in our model, we use mass-action kinetics with rate constants $k = 1$. Therefore, the respective propensities for the elongating enzyme (a_{GS}) and the branching enzyme (a_{GBE}) write:

$$a_{\mathrm{GS}} = c_{\mathrm{GS}} \cdot S_{\mathrm{GS}}$$
$$a_{\mathrm{GBE}} = c_{\mathrm{GBE}} \cdot S_{\mathrm{GBE}},$$

where c_{GS} and c_{GBE} are parameters that can force the system into either elongation or branching, depending on the previously introduced ratio $\Gamma = \frac{c_{GS}}{c_{GBE}}$. Besides, S_{GS} and S_{GBE} are the number of available substrate chains for GS and GBE, respectively. It is important to recall that we are interested in the phenomenology of the system. In this context, simply assuming mass-action kinetics is sufficient to investigate the different elongation and branching regimes of interest.

5.3.2 Overall simulation algorithm

The main steps of our algorithm are outlined in Fig 11. The overall loop, in which the biochemical reactions are simulated, begins by analysing the granule structure and identifying the available substrate chains for GS and GBE. Based on this information, the propensity for each reaction is calculated, using a mass-action kinetic approach as previously described in subsection 5.3.1. The Gillespie direct method randomly determines the next reaction and its duration. If the reaction is an elongation, the code selects one of the potential substrate chains at random, and analyses the granule structure to check for overlaps. If there is none, the glucose monomer is added to the substrate chain. Otherwise, another substrate chain is selected at random, and the overlap criterion is checked again. If the reaction is a branching, the code proceeds in the same way, but it verifies that there is no overlap for the entire daughter chain.

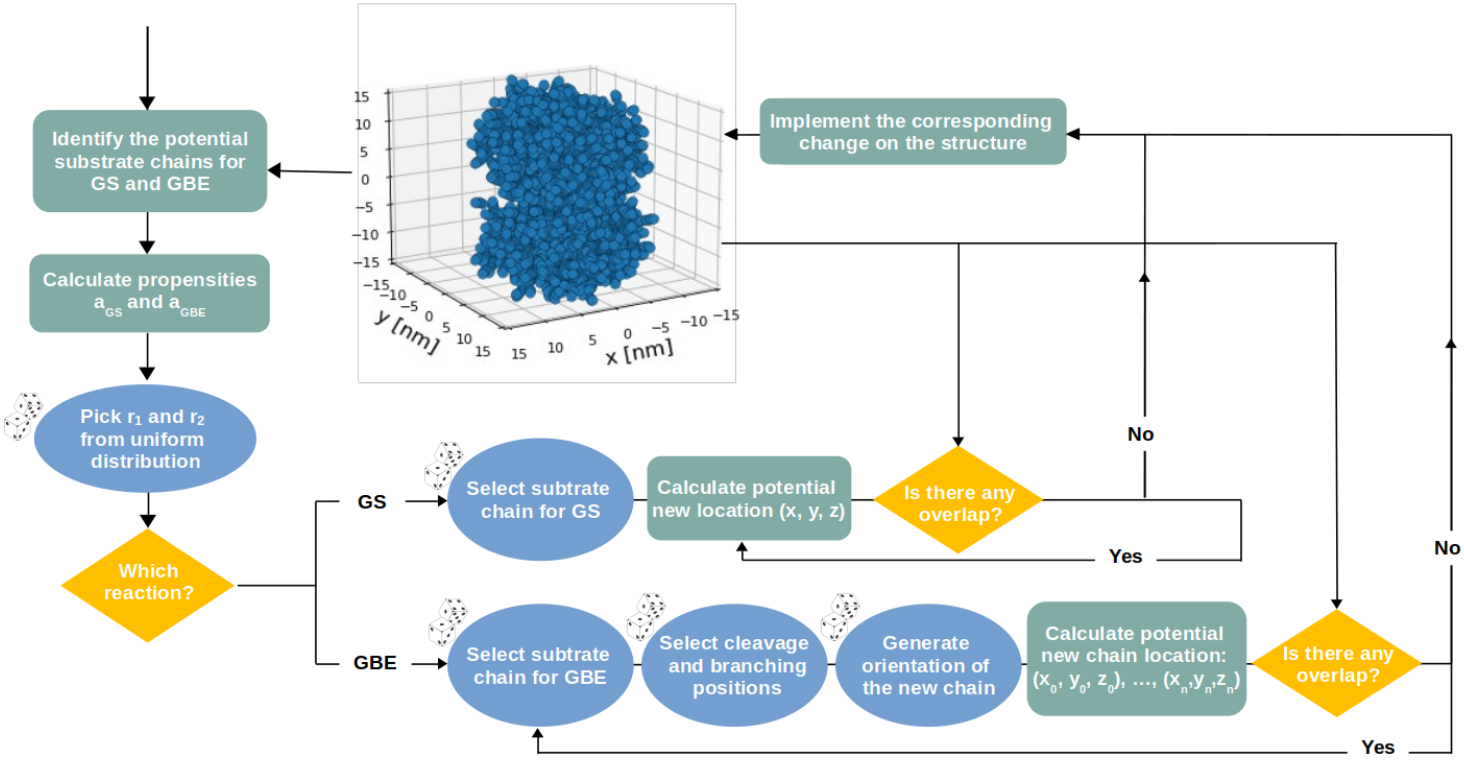

Fig 11. Algorithm flow diagram. This scheme shows the main steps of our algorithm. The green round rectangles denote deterministic calculations or actions. The blue ovals involve stochastic steps relying on random numbers. Finally, the yellow diamonds are conditional decisions (or "if" statements).

5.4 Effect of self-exclusion

Modelling the 3D structure of glycogen, while considering self-exclusion of glucoses, allows investigating glycogen chain patterns, molecular density, size of the granules, and crowding at the surface. But is this detailed approach strictly required for studying glycogen's macroscopic properties? In the following, we test various radius values for the spheric description of the glucose units and look whether they affect the macroscopic properties of glycogen.

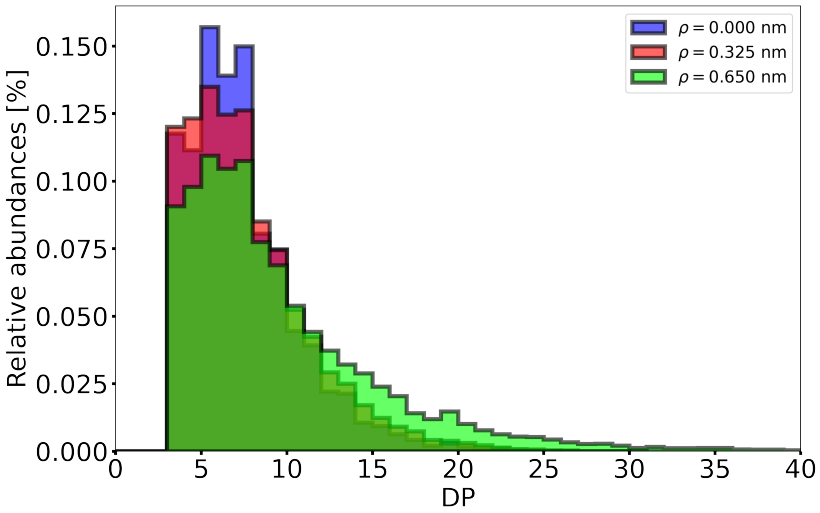

Fig 12. Effect of the steric hindrance on the CLD. Chain length distributions for $\rho = 0$ nm (blue), $\rho = 0.325$ nm (red), and $\rho = 0.650$ nm (green). For $\rho = 0$ nm, the distribution exhibits a higher peak at DP [6-8] than for $\rho = 0.325$ nm and $\rho = 0.650$ nm. Opposite, for $\rho = 0.325$ nm and $\rho = 0.650$ nm, the higher DPs (from DP 12) are over-represented as compare to $\rho = 0$ nm.

Three scenarios are investigated: $\rho = 0$ nm, $\rho = 0.325$ nm, and $\rho = 0.650$ nm. For $\rho = 0$ nm, the glucose units have no volume and thus no steric hindrance, allowing chains to overlap. Instead, $\rho = 0.325$ nm is equal to half of the helix's van der Waals radius, while $\rho = 0.650$ nm is its total radius. By comparing these scenarios, we can evaluate the impact of steric hindrance on the substrate availability during the synthesis of glycogen.

Fig 12 shows that short chains are more abundant for $\rho = 0$ nm than for the other two values of ρ. Opposite, long chains are more abundant for non-zero ρ. As soon as $\rho > 0$ nm, since elongation involves adding a single glucose unit, while branching means transferring an entire piece of a branch, it is easier to find the necessary space around the substrate for allowing elongation to take place. This is reinforced by the fact that when adding a new glucose unit at the non-reducing end of a branch, we do not elongate the substrate by the total length of a glucose unit, but only its radial contribution, which is $l = 0.24$ nm (see Fig 2). As a consequence, steric hindrance stronger impacts branching than elongation. In other words, the number of branching attempts rejected due to steric hindrance is higher than that of elongation. If we would define effective branching and elongation rates that respectively account for these rejections, the branching effective rate would reduce much more than that of elongation. This would lead to an effective elongation to branching ratio Γ_{eff} which

increases with the effect of steric hindrance. In the section Elongation to branching ratio, we concluded that as Γ increases, the overall CLD shifts towards higher DPs and the distribution peak decreases. $\Gamma_{\text{eff}}^{\rho=0} < \Gamma_{\text{eff}}^{\rho=0.325} < \Gamma_{\text{eff}}^{\rho=0.650}$ therefore explains both the over-representation of high DPs and the peak reduction as ρ increases in Fig 12. Despite these changes in the CLD, as the steric hindrance increases, we can remark that not only is the number of peak conserved (here unimodality) but their location too.

5.5 Flexible location *versus* strict location branching models

As briefly introduced in the section The model, paragraph Enzymatic reactions, two assumptions for GBE can be made. In the flexible location branching model (used throughout the article) all glucose units that are in the acceptable range can equiprobably be cleaved, and similarly for those that can potentially be transferred. Opposite, in the strict location branching model (introduced here for comparison) GBE always branches at a precise location, i.e. at a given distance from the non-reducing end. For the sake of clarity, these two models and their associated mechanisms are fully detailed in Fig 13. Their patterning impact on the CLD is presented in Fig 14.

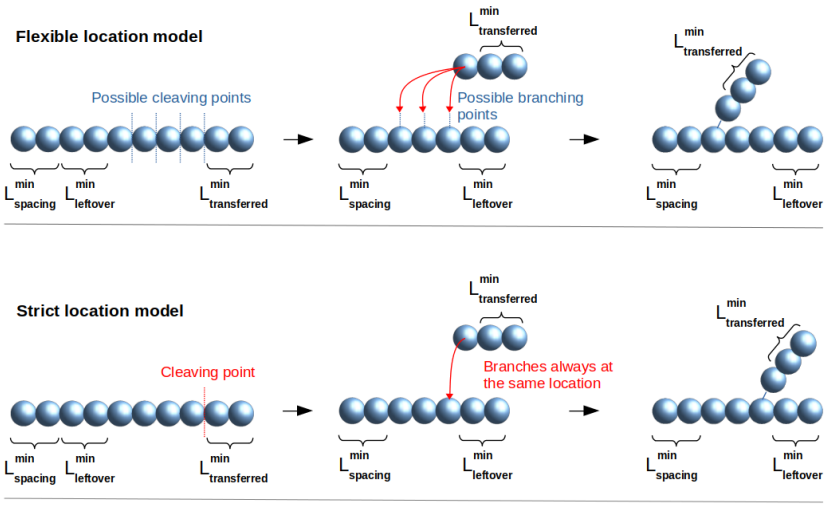

Fig 13. Mechanisms of the flexible location and the strict location branching models. Both models fulfill the minimal lengths requirement (Eq 1). In the flexible location branching model, the cleaving position is randomly picked from a uniform distribution, and so is that of branching. Instead, in the strict location branching model, both the cleaving and the branching always occur at a specific distance from the non-reducing end.

In Fig 14, we compare the two branching models, and vary Γ and $\{L_{\text{spacing}}^{\text{GBE}}, L_{\text{leftover}}^{\text{GBE}}, L_{\text{transferred}}^{\text{GBE}}\}$, which results in 12 different plots. From the Elongation to branching ratio section, we learnt that increasing Γ spreads the CLD towards higher DPs and reduces the distribution peak typically observed for short chains. In the section Effect of the branching enzyme on the CLD, we observed that increasing the minimal lengths $\{L_{\text{spacing}}^{\text{GBE}}, L_{\text{leftover}}^{\text{GBE}}, L_{\text{transferred}}^{\text{GBE}}\}$ modifies the shape of the CLD that becomes bi- or even multi-modal. When decreasing these lengths, the multi-modality is weakened while it is reinforced by a small Γ.

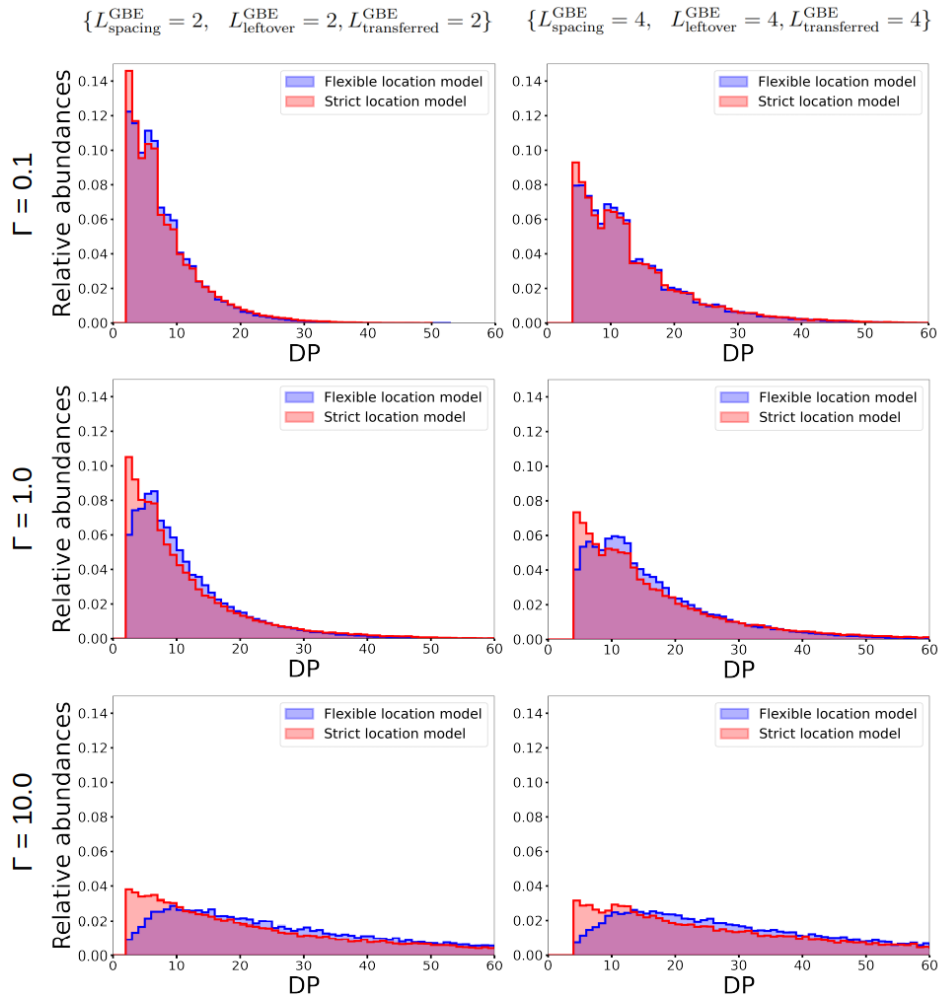

Fig 14. Comparing the impact of the flexible location and the strict location branching models on the CLD. The CLDs for the flexible location and the strict location branching models are shown in blue and red, respectively. **Top-left:** CLD obtained for small Γ (branching regime), with small minimal lengths for GBE. **Top-right:** CLD obtained for small Γ (branching regime), with longer minimal lengths for GBE. Increasing the lengths reinforces the multi-modality. **Middle-left:** CLD obtained for $\Gamma = 1$ (intermediary regime), with small minimal lengths for GBE. Both models loose their multi-modalities. **Middle-right:** CLD obtained for $\Gamma = 1$ (intermediary regime), with increased minimal lengths for GBE. Multi-modality is restored in both models. **Bottom-left:** CLD obtained when increasing Γ, with small minimal lengths for GBE. Both conditions contribute to weaken the multi-modality. **Bottom-right:** CLD when increasing both Γ and minimal lengths for GBE. The first condition tends to weaken multi-modality while the latter one instead reinforces it. As a consequence, multi-modality is observed for the strict location branching model only.

For $\Gamma = 0.1$, the strict location branching model (red) matches the flexible one (blue). That is not surprising, since in this particular regime, when branching strongly dominates over elongation, branching occurs as soon as possible (DP $> L^{\mathrm{GBE}}_{\mathrm{min}}$), making the two models equivalent.

When Γ increases (second and third lines), the two models are not anymore equivalent. In the strict location model, the length of the transferred branch is always equal to $L^{\mathrm{GBE}}_{\mathrm{transferred}}$, while in the flexible location model, longer chains can be transferred. In both cases, the smallest transferable DP is equal to $L^{\mathrm{GBE}}_{\mathrm{transferred}}$ (see section 2.2.3). For the strict location model the latter is the most abundant DP, unlike for the flexible one.

5.6 Scope of the parameter space

We apply our fitting procedure to all possible combinations of $\Gamma \in \{0.1,\ 0.2,\ 0.3,\ 0.4,\ 0.6,\ 0.8,\ 1.0,\ 2.0,\ 5.0\}$, $L^{\mathrm{GBE}}_{\mathrm{spacing}} \in \{1,\ 2,\ 3,\ 4,\ 5,\ 6\}$, $L^{\mathrm{GBE}}_{\mathrm{transferred}} = L^{\mathrm{GBE}}_{\mathrm{leftover}} \in \{2,\ 3,\ \ 4,\ 5\}$, and $\rho \in \{0\text{ nm},\ 0.325\text{ nm},\ 0.650\text{ nm}\}$, while considering two different branching scenarios, i.e. the flexible location and the strict location branching models. It sums up to a total of 1,296 tested parameter sets. The corresponding heat-maps are shown in Fig 15 for the flexible location and Fig 16 for the strict location branching models. In both figures, ρ varies across the columns, while $L^{\mathrm{GBE}}_{\mathrm{leftover}} = L^{\mathrm{GBE}}_{\mathrm{transferred}}$ does over the successive rows. On each heatmap, the Y and X axes correspond to $L^{\mathrm{GBE}}_{\mathrm{spacing}}$ and Γ, respectively.

In Fig 15, all scores below 20 are highlighted with white squares, and correspond to good fits (arbitrary cut-off chosen as up to twice the best-fit). Noticeably, various parameter sets fulfill this criterion and all of them show a small $L^{\mathrm{GBE}}_{\mathrm{spacing}}$ value. As presented in section Comparison to experimental data, paragraph Parameters calibration, the best score is obtained with $\rho = 0.65$ nm (third column). This supports the fact that steric hindrance plays a role in the chain length distribution of real glycogen, although good matches also exist with $\rho = 0$ nm, in which the CLD is not impacted by steric hindrance. When focussing on the good scores (dark blue cells), it appears that changing the elongation to branching ratio Γ can be compensated by varying $L^{\mathrm{GBE}}_{\mathrm{spacing}}$, or the two other minimal lengths.

Opposite, when considering parameter values that are typically reported in the literature, we systematically obtain very poor (i.e. high) scores, even upon varying Γ. This case is highlighted in red and discussed in the section Comparison to experimental data, paragraph Parameters calibration. In general terms, for each heatmap, we observe that $L^{\mathrm{GBE}}_{\mathrm{spacing}} \geq 4$ corresponds to poor scores, i.e. $\mathcal{S} \geq 30$. This effect is even more pronounced if $L^{\mathrm{GBE}}_{\mathrm{spacing}} \geq 6$, i.e. $\mathcal{S} \geq 60$ (66.3 and above). Thus, with these parameter sets, our model is not able to reproduce the experimental CLD data by Sullivan and coworkers used for fitting throughout this article [25]. Regarding the enzyme mechanism, this suggests that GBE is able to branch much closer than 4 glucose units away from an existing chain. Similarly, if $L^{\mathrm{GBE}}_{\mathrm{transferred}} = L^{\mathrm{GBE}}_{\mathrm{leftover}} \geq 3$, we are not able to fit the experimental data by Sullivan and coworkers. Therefore, these minimal lengths must be shorter than those typically reported in the literature.

In Fig 16, we present analogous results for the strict location model. Overall, the fitting procedure returns much poorer scores than for the flexible location branching model (see Fig 15). This is due to the fact that strict branching locations almost systematically lead to a multi-modal distribution, which is not the case of the experimental CLD data set fitted here [25]. Although multi-modality can be cleared by increasing Γ, this would lead to a reduction of the distribution peak and a shift of the overall distribution towards higher DPs. For the experimental data considered, we could not find a good trade-off, corresponding to good fitting scores. As a result, the strict location model seems to be very unlikely, that is why we instead choose the flexible location branching model for this study.

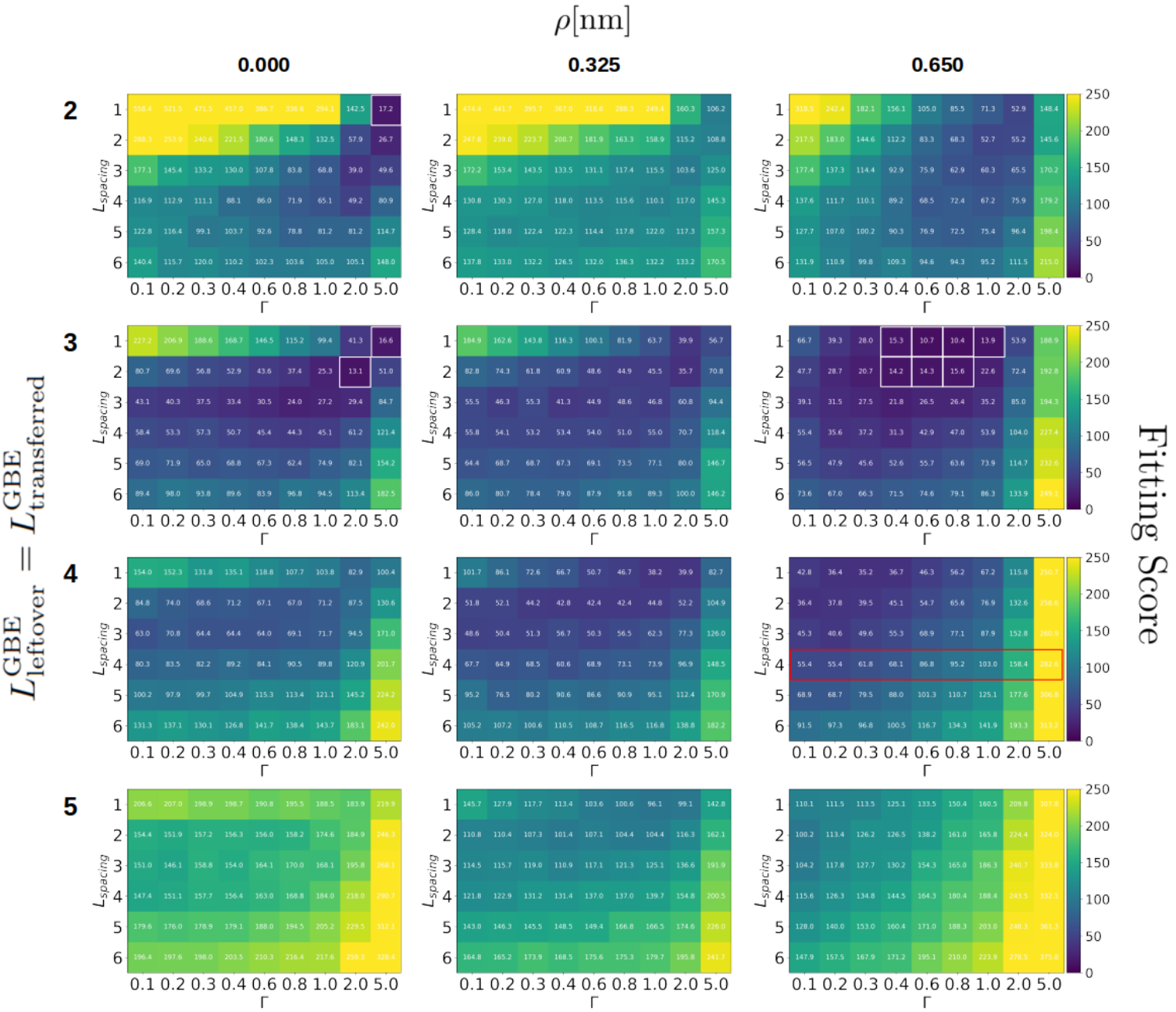

Fig 15. Heat-maps showing fitting scores for the flexible location branching model. The three columns show 3 different radii for the glucose units ($\rho = 0$ nm, 0.325 nm, and 0.650 nm). The 4 rows correspond to distinct values of $L^{GBE}_{transferred}$ (2, 3, 4, and 5) with $L^{GBE}_{leftover} = L^{GBE}_{transferred}$. For each heat-map, the values for $L^{GBE}_{spacing}$ are on the Y axis, while the Γ ratio is on the X axis. Each cell is characterised by a unique set of parameters $\{\rho, \Gamma, L^{GBE}_{spacing}, L^{GBE}_{leftover}, L^{GBE}_{transferred}\}$. White squares show good scores ($\mathcal{S} \leq 20$), while the red rectangle ($\mathcal{S} \geq 31.3$) shows the scores obtained for minimal length values as typically assumed in the literature.

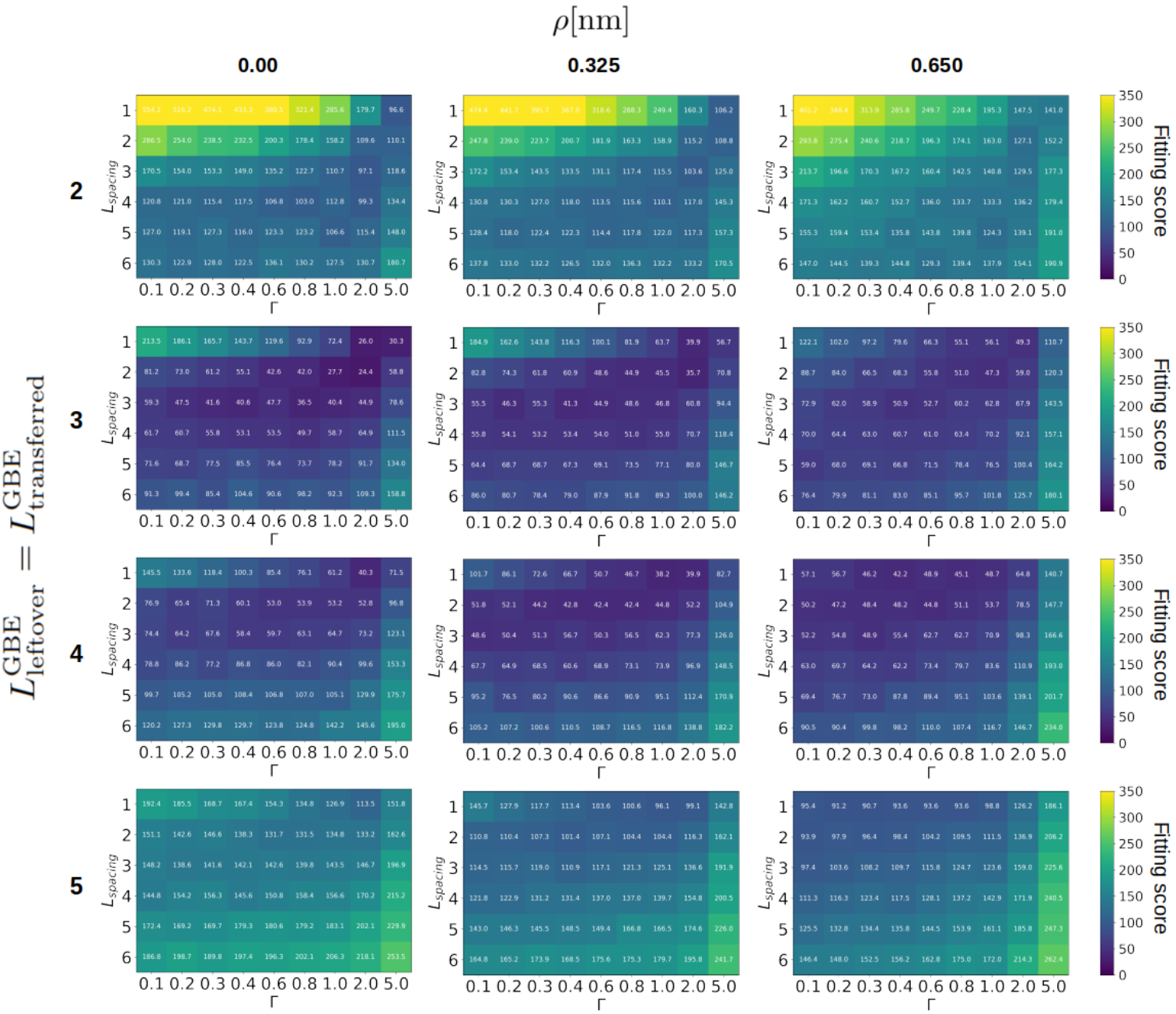

Fig 16. Heat-maps showing fitting scores for the strict location branching model. The three columns show 3 different radii for the glucose units ($\rho = 0$ nm, 0.325 nm, and 0.650 nm). The 4 rows correspond to distinct values of $L^{\mathrm{GBE}}_{\mathrm{transferred}}$ (2, 3, 4, and 5) with $L^{\mathrm{GBE}}_{\mathrm{leftover}} = L^{\mathrm{GBE}}_{\mathrm{transferred}}$. For each heat-map, the values for $L^{\mathrm{GBE}}_{\mathrm{spacing}}$ are on the Y axis, while the Γ ratio is on the X axis. Each cell is characterised by a unique set of parameters $\{\rho, \Gamma, L^{\mathrm{GBE}}_{\mathrm{spacing}}, L^{\mathrm{GBE}}_{\mathrm{leftover}}, L^{\mathrm{GBE}}_{\mathrm{transferred}}\}$. Overall, the scores are much higher than those for the flexible location branching model (see Fig 15), indicating poorer fits. Specifically, no scores below the threshold $\mathcal{S} = 20$ are found.

5.7 Probabilistic approach to the A:B ratio

In this section, we describe why the A:B ratio must be equal to 1 when no biological nor biophysical properties of the system are considered, but pure probabilities. As presented in Section 2.3.2, we note as two distinct reactions, branching on either an A or a B chain:

$$R_{A\to AB}: \quad \mathrm{A} \xrightarrow{\text{GBE}} \mathrm{B} + \mathrm{A}_{\text{new}}$$
$$R_{B\to AB}: \quad \mathrm{B} \xrightarrow{\text{GBE}} \mathrm{B} + \mathrm{A}_{\text{new}} \,.$$

In this purely mathematical framework, A and B chains have the same chance to react, so that the probability for a reaction to occur only depends on the respective number of A (noted N_A) and B (noted N_B) chains. We sketch the associated probability tree in Fig 17. The tree is symmetric, and the number of paths to a given state follows the

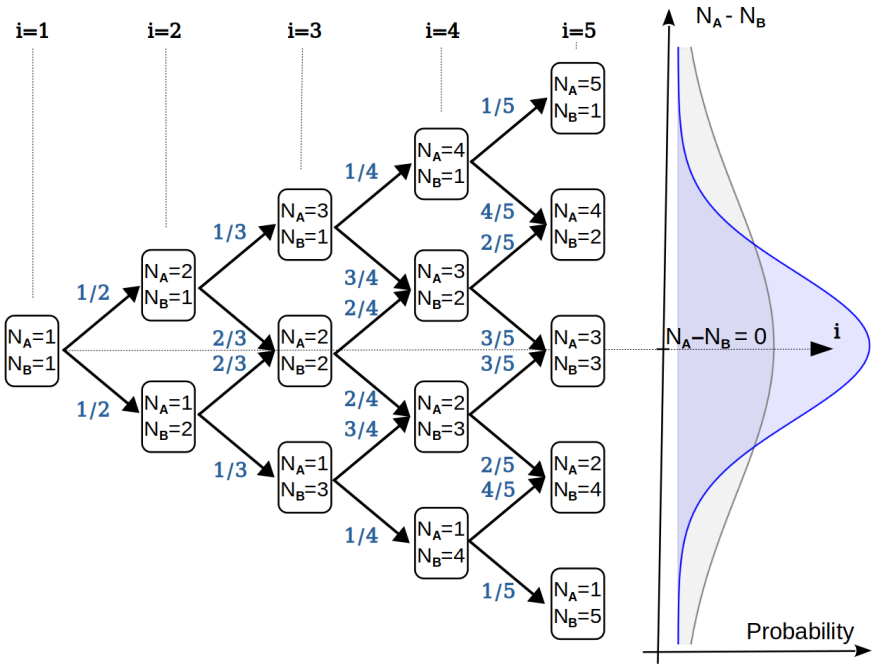

Fig 17. Probability tree of forming a given number of chains of type A (N_A) and B (N_B). Directions that spread the tree are unfavoured, while those oriented towards the center of the tree (reducing the absolute value $|N_A - N_B|$) are favoured, proportionally to the difference $|N_A - N_B|$. Therefore, for high numbers of branching reactions (noted i) the distribution (blue) is centered around $N_A - N_B = 0$, and is thinner than a binomial distribution (grey).

binomial coefficients. The horizontal dotted line corresponds to $N_A - N_B = 0$, when there are as many A as B chains. We know that in the example of flipping a coin i times, the central limit theorem tells us that the distribution of the difference in the number of heads and tails, tends to a normal distribution centered in 0 when i tends to infinity. This case corresponds to heads and tails having equal probabilities. In our case, a given state of the probability tree (N_A, N_B) leads either to the state $(N_A, N_B + 1)$ with probability $p(R_{A\to AB}) = \frac{N_A}{N_A+N_B}$, or to $(N_A + 1, N_B)$ with probability $p(R_{B\to AB}) = \frac{N_B}{N_A+N_B}$. It means that the probability to go from one state

to the next one depends on the state of the system, in a way that the probability to come closer to $N_A - N_B = 0$ (horizontal dotted line) is always higher than that of spreading away. Additionally, the distribution of the probabilities remains symmetric with respect to the case $N_A = N_B$. Based on these considerations, the mean value $\overline{N_A - N_B} = 0$ and the distribution is even more peaked than in the simple case of flipping a coin. Since the mean value $\overline{N_A - N_B} = 0$, the mean ratio $\overline{\frac{N_A}{N_B}} = 1$, in other words the A:B ratio is 1.

Chapter 4

The interplay between glycogen structure and glycogen metabolism

In the previous chapter's final section, we demonstrated how the model may serve as a valuable tool in characterizing the structural aspects of specific glycogen-related disorders. However, the scope of this approach is notably limited from a metabolic perspective. Thus far, we have considered a really limited system that comprises solely glucose units and four enzymes. Additionally, the kinetic laws applied, which served to describe the structures, held no physiological relevance.

While we could potentially establish a kinetic model that encompasses the principal pathways associated with glycogen, such as glycogen synthesis, glycogenolysis, or glyconeogenesis, in an ODE system, we would fail to account for the effects related to glycogen structure. Such models do not regard glycogen as a full metabolite; rather, it is viewed as the amount of glucose stored in the form of glycogen. In fact, these models solely encompass glycogen through its concentration and make no differentiation between distinct structures. For instance, a linear chain of 10,000 glucose molecules would not be distinguished from a highly branched molecule of equivalent molecular weight. These two structures should not possess identical synthesis and degradation rates, given that GS or GP react at the non-reducing end.

Furthermore, without a model that takes structural features into account, we are unable to characterize the roles of branching and debranching enzymes (GBE and GDE) in metabolism. Specifically, we cannot account for the fact that GDE directly releases glucose without conversion of glucose intermediate glucose-1-phosphate and glucose-6-phosphate.

The aforementioned examples illustrate the possible effects of structure on the dynamics of the metabolites within the system. We may also envision how metabolic reactions, even distant ones, may impact the molecule's structure. It is precisely this mutual interaction between structure and the metabolic network that piques our interest and that we aim to investigate.

In the opening section of this chapter, we introduce novel algorithmic approaches that allow for the integration of both deterministic and stochastic methods in solving a chemical reaction system. This approach permits the preservation of the stochastic nature of certain reactions essential for investigating glycogen structure, while modeling other reactions via a conventional

approach to avoid the excessive computational demands commonly associated with large reaction networks when utilizing purely stochastic methods.

Subsequently, we present a model that describes glycogen synthesis and degradation in response to hormonal signals. This model provides a simplified representation of glucose's pathway from blood plasma to glycogen and vice versa. With this model, we may apply the algorithmic approach introduced in the prior section and examine the relationship between glycogen structure and its metabolism.

Finally, armed with this newly developed tool, we conduct a comprehensive investigation of specific glycogen-related disorders described in Chapter 1 to gain deeper insights into their underlying mechanisms.

4.1 Coupling deterministic and stochastic methods

As outlined in chapter 2, the Gillespie algorithm [Gillespie, 1976, 1977a,b] yields results identical to solving an ODE system using any integrating methods if the propensities are properly expressed. The system depicted in figure 2.9 will be reused and will serve as a proof of concept in the following section, where we introduce new methods for coupling both deterministic and stochastic approaches.

4.1.1 Coupling stochastic and deterministic reactions

The classical Gillespie algorithm (Gillespie, 1977a) is limited in its ability to model complex systems consisting of a large number of molecules. This is because the algorithm generates the time elapsed reaction by reaction, and the total simulation time, T, is the sum of all the time intervals, δt_i, which have passed from one reaction to the next. When the system being modeled increases in the number of molecules, the time interval between successive reactions decreases drastically. For a mass action kinetics, the number of reactions in a system with N particles is proportional to N^2. Thus, an increase of a factor of 10 in the number of molecules considered will result in, on average, δt_i being 100 times less than the number of molecules considered. Consequently, the algorithm will have to iterate 100 times more to produce a simulation of the same total time, T.

To address this issue, methods derived from the classical algorithm have been developed, including the tau-leaping method [Gillespie, 2001]. This approach segments time and approximates the Gillespie solutions within the time intervals. Although more efficient, the tau-leaping method is only an approximation of the exact approach of the classical algorithm and depends on the characteristic time, τ, of the algorithm. A larger τ does not allow for a good approximation. Theoretical arguments on the appropriate value of τ for a good balance between performance and accuracy are discussed in the article [Gillespie, 2001]. However, such an algorithm is not applicable in situations where access to the events reaction by reaction is necessary.

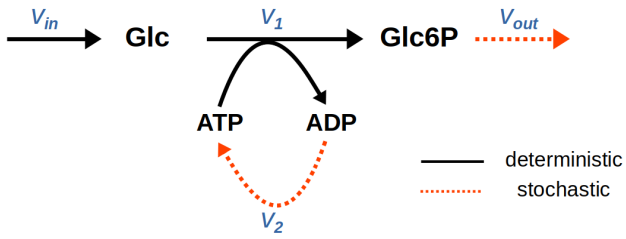

Figure 4.1: **A toy hybrid system to test the coupling method.** The system is composed by 4 reactions and 4 metabolites. Each reaction flux is highlighted in blue. Glucose (Glc) enter the system at rate v_in. Glucose react with one ATP to form Glucose-6-phosphate (Glc6P) and ADP at rate v_1. ADP is converted again into ATP at rate v_2. Finally Glucose-6-phosphate escape the system at rate v_{out}

Nonetheless, the idea of a time larger than that of an individual reaction, to approximate other reactions, is applicable and will be explored further in our study.

The Gillespie algorithm faces a second challenge, which is related to the first problem discussed above. Specifically, the algorithm struggles to model systems with vastly different reaction rates. In such cases, the algorithm will spend most of its time returning elapsed times for species involved in fast reactions, causing significant slowdowns in the simulation. To address this issue, a variant of the Gillespie algorithm, known as the Slow-Scale Stochastic Simulation Algorithm (SSSSA) has been proposed [Cao et al., 2005; Haseltine and Rawlings, 2002]. The SSSSA approach separates fast and slow reactions in the system and employs a stochastic description only for the slow reactions. During simulation, the same reaction can be simulated in both a deterministic and stochastic way.

However, in our specific case, we desire a stochastic description for the reactions that directly affect glycogen structure to allow us to track each reaction individually and take appropriate action. Therefore, we propose a variation of the SSSSA algorithm to address this requirement.

Our approach seeks to combine both the conventional deterministic approach and the stochastic Gillespie direct method approach. Specifically, we will solve the system described in section 2.1.4 using deterministic methods for two reactions (v_{in} and v_1), while using the Gillespie approach for two other reactions (v_2 and v_{out}). The red dotted arrows in figure 4.1 represent the two reactions that we will model using the Gillespie approach.

The proposed approach involves utilizing the elapsed time returned by the Gillespie algorithm at each iteration as the integration time for the solver to solve the deterministic equations. There are two possible scenarios. In the first case, the method is efficient when the stochastic reaction has the lowest flux, meaning that it occurs the least in the system. In such a scenario, the time returned by Gillespie, denoted by δt, will be significantly larger compared to the characteristic times of other reactions. As a result, the integration of the deterministic ODE system will capture the change in metabolite concentrations during this time interval. In

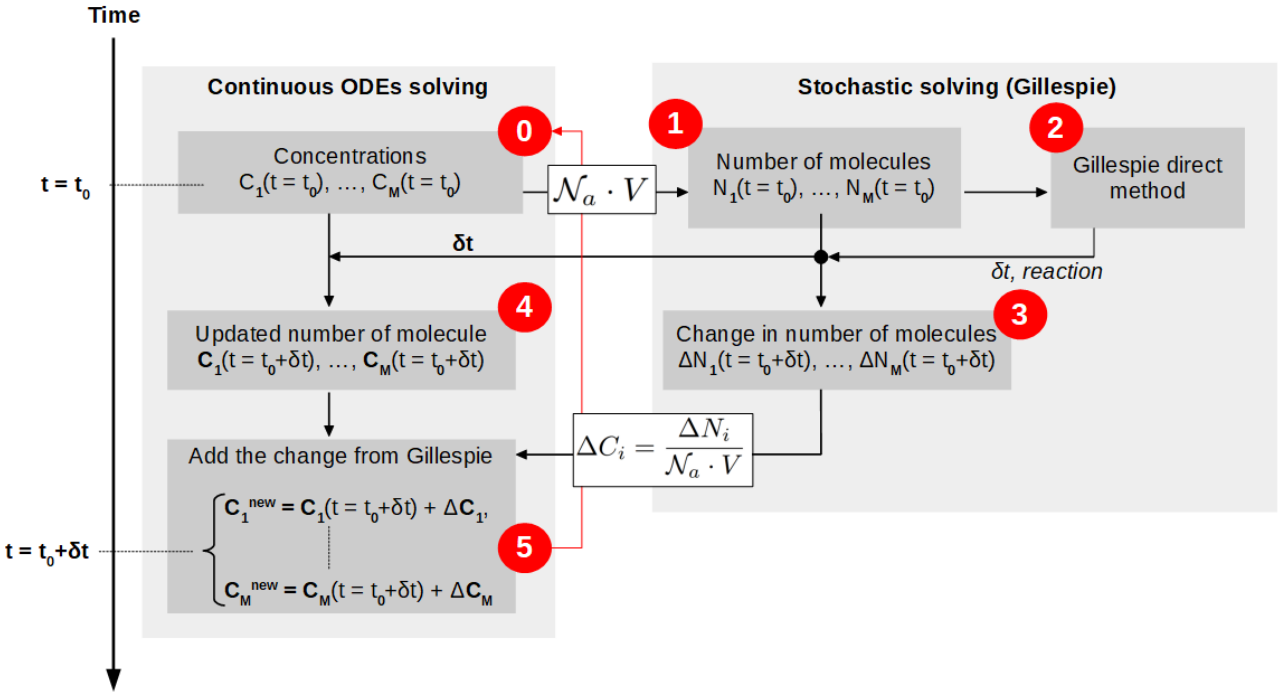

Figure 4.2: **Flow diagram of the hybrid algorithm.**

the second case, where the stochastic reactions occur more frequently, the time δt will be considerably smaller compared to the characteristic time of other reactions. In such a scenario, the deterministic system will have little time to evolve during this short time interval. The algorithmic workflow is summarized in Figure 4.2.

The algorithm consists of several steps to simulate the evolution of a system over time. At step 0, the initial state of the system is defined, including the concentrations C_i of the M metabolites at time t_0. These concentrations are then converted into the number of particles N_i at step 1, using the Avogadro number and the volume of the system V. The direct Gillespie method is then applied at step 2, where the corresponding propensities are calculated and a reaction and the time δt necessary for the reaction to occur are returned. The numbers of molecules involved in the reaction are then updated at step 3. After this, the ODE module is called at step 4 to integrate the initial concentrations based on the time obtained from the Gillespie algorithm, δt. This produces the concentrations at time $t = t_0 + \delta t$ as if the system evolved without taking into account the stochastically modeled reactions. At step 5, the variation in concentrations δC_i due to the stochastically modeled reaction is added to these concentrations. Finally, the algorithm returns to step 0 with the new state of the system. By iteratively performing these steps, the algorithm can simulate the evolution of the system over time, taking into account both deterministic and stochastic processes.

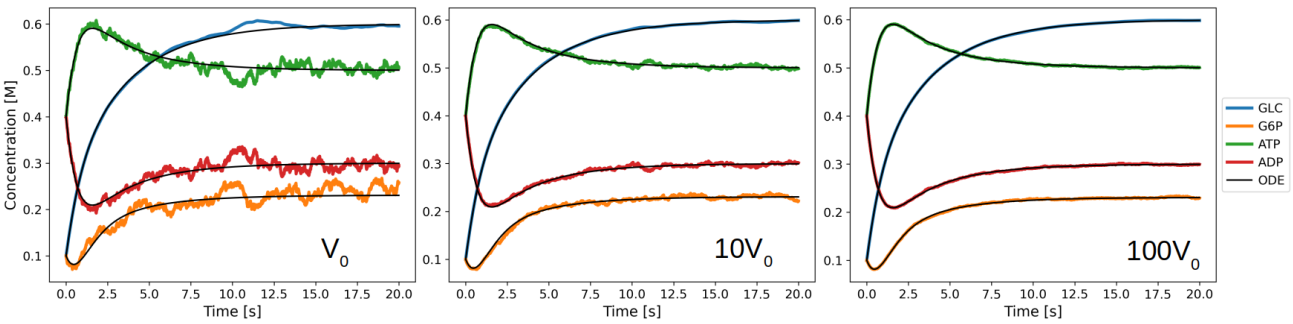

Figure 4.3: **Time course of the toy system.** From left to right, the volume of the system increases (V_0, $10V_0$, $100V_0$), reducing the fluctuation in the stochastic algortihm

It converges and match the solutions obtained by a regular ODE solving method (Fig 4.3). Different volumes for the system are considered and, as in section 2.1.4, we see the fluctuation diminishing as the volume increases. One can notice a visual difference in the left plot on Fig 4.3, compare to the plot from the full Gillespie approach from Fig 2.10. The Glucose time course fluctuate less than the others. It is because this is the only metabolite in the toy system which is not involved in any stochastic reaction. While the method converges,, the computational time for $V = 100V_0$ with $V_0 = 10^{-19}\text{m}^3$, is approximately 37 times higher than a simple Gillespie approach (see table 4.1), which motivates us to find an improvement to the approach to reduce this calculation time.

4.1.2 Toward a more efficient algorithm: Periodic-Coupling Algorithm

In this part, we will try to make the previous algorithm more efficient in terms of computation time. The previous algorithm is slow because at each stochastic reaction, we have to reintegrate the deterministic differential equation system. The idea here is to try to decouple slightly the two modules (stochastic and deterministic), so that they work independently, in parallel for a certain time interval greater than the δt time of a single stochastic reaction. The two modules will then communicate once this time is exceeded and allow the system to be updated. The figure 4.4 shows the main steps of this algorithm. The state of the system is described by all the concentrations C_i, $i \in \{1, .., M\}$ at time $t = t_0$ (step 0). As in the algorithm presented above 4.1.1, the concentrations C_i are then converted in number of particles N_i (step 1), with $N_i = C_i \cdot \mathcal{N}_a \cdot V$, where $\mathcal{N}_a$ is the Avogadro number and V the volume of the considered system. Step 2 consist in running the Gillespie algorithm several times on the sub-system made of stochastic reactions. The elapsed time is tracked by summing all the Gillepspie time δt and the number of molecules are updated independently of the deterministic reaction. After a certain number of iterations i, the elapsed time $\sum_{k=1}^{i} \delta t_k$ exceeds the threshold value of the communication time T. When this condition is reached, we record the elapsed time in the stochastic module $\Delta t = \sum_{k=1}^{i} \delta t_k$ (step 4). The fact that we stop after T ensure that at least

a single event occurs with $\delta t_k > T$. In this precise case the communication time becomes δt_k and the algorithm is a special case of the algorithm 1 shown in the previous section.This Δt is later used by the deterministic module to integrate its system of equations (step 5). Finally the two module communicate. The concentration of the deterministic module is changed by the variation in concentration due to the stochastic module (step 6). The new concentrations become the input of the next iteration (step 0).

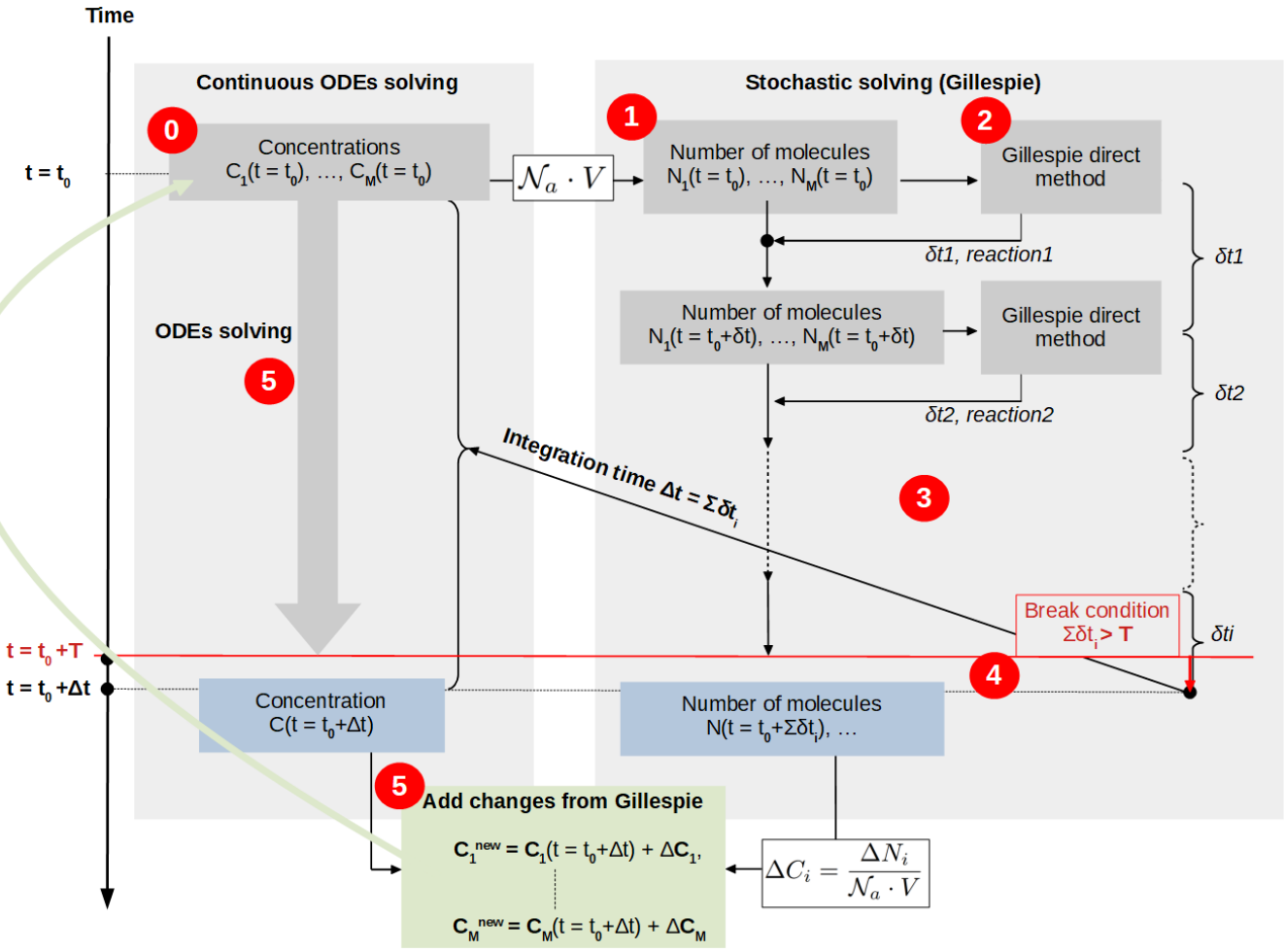

Figure 4.4: **Flow diagram of the Periodic-Coupling algorithm .**

We applied this algorithm to the toy model used before 4.1. The convergence is also obtained with this new algorithm, but only when the communication between the two module is frequent enough, or equivalently when T is small enough. We summarize the different time-course obtain with different volume V and communication time T in Fig 4.5. From left to right, the volume of the system increases reducing the fluctuation in the concentrations. From Top to bottom, we reduce the communication time between the two modules. For $T = 0.1$ we observe that

the algorithm do not converge toward the exact solution given by the deterministic solving. However, when we decrease the communication time ($T = 0.01$ and $T = 0.001$), the algorithm converge for all Volumes.

The algorithm ensure that we always have $\Delta T \geq \delta t$ and $\Delta T \approx \delta t$ when we set the communication time T below the typical Gillespie time. The algorithm will therefore solve the deterministic ODE system way less often, ensuring a faster computational method. In table 4.1 we compare the simulation times for all the methods with different volumes. We observe that the periodic-coupling method outperform the basic simple-coupling method by consequent time (18 and 58 times faster for all volumes. When the volume becomes bigger, the algorithm also outperformed the simple Gillespie approach. Another argument in favor of this algorithm is that these results are obtained for a simple system of 4 reactions. As the number of reactions considered increases, it is expected that the number of molecules to be considered will increase as well, deteriorating the computational time of the simple Gillespie approach, and not influencing, or very little, the periodic coupling algorithm. Indeed, adding non-stochastic reactions will increase the computation time in the deterministic module, which is not the limiting part here.

We will therefore use this new algorithm it the next sections.

Table 4.1: Computational time from different methods

Model	Volume of the system			
	V_0	$10 \cdot V_0$	$100 \cdot V_0$	$1000 \cdot V_0$
ODE	0.11 s	-	-	-
Gillespie	0.08 s	0.69 s	7.58 s	74.10 s
Algo 1: Direct coupling	3.40 s	29.41 s	281.30 s	3009.77 s
Algo 2: Periodic coupling ($T = 0.01$)	1.05 s	1.61 s	5.02 s	40.95 s

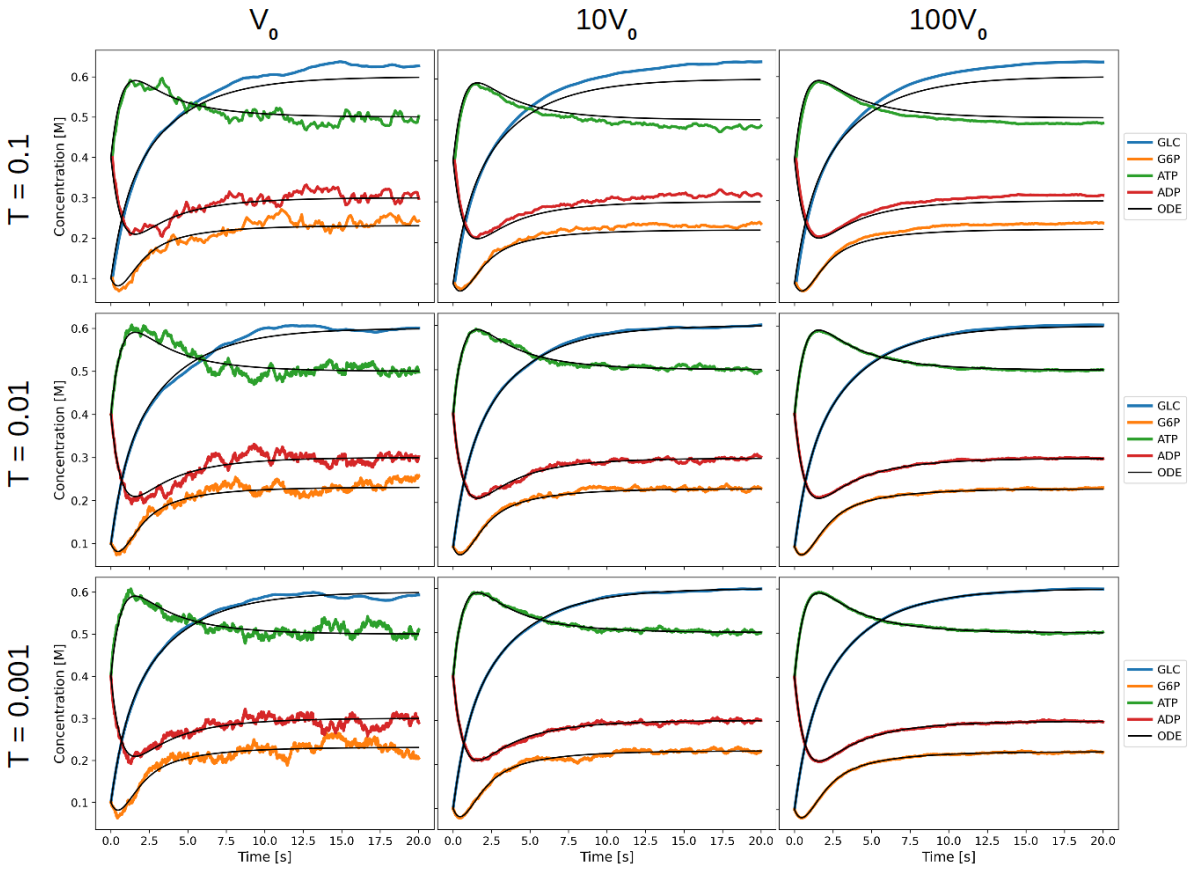

Figure 4.5: **Time-course obtained with the periodic-coupling algorithm. All algorithms were run on an Intel(R) Xeon(R) W-2135 CPU @ 3.70GHz** The colored lines correspond the algorithm solution while the black lines are the "exact" solution obtained from the ODE system. **Left to right:** the volume of the system increase which reduce the fluctuations. **Top to bottom:** The communication time T decrease between the two module of the algorithm. It results in the convergence of the solutions toward the exact solutions.

4.1.3 Plugging the glycogen module to the Periodic-Coupling algorithm

In this last subsection we will apply the algorithm developed to a system together with our glycogen module developed in Chapter 2 and 3. The system is depicted in Fig 4.6 and consist of 8 reactions. All reactions follow a mass-action kinetics. The stochastic sub-system of the Periodic-Coupling algorithm will consist of the 4 reactions that act on the glycogen granules (GS, GBE, GP and GDE). The other ones will be modeled deterministically.

The determinstc sub-system will be described by the following mass action rate laws:

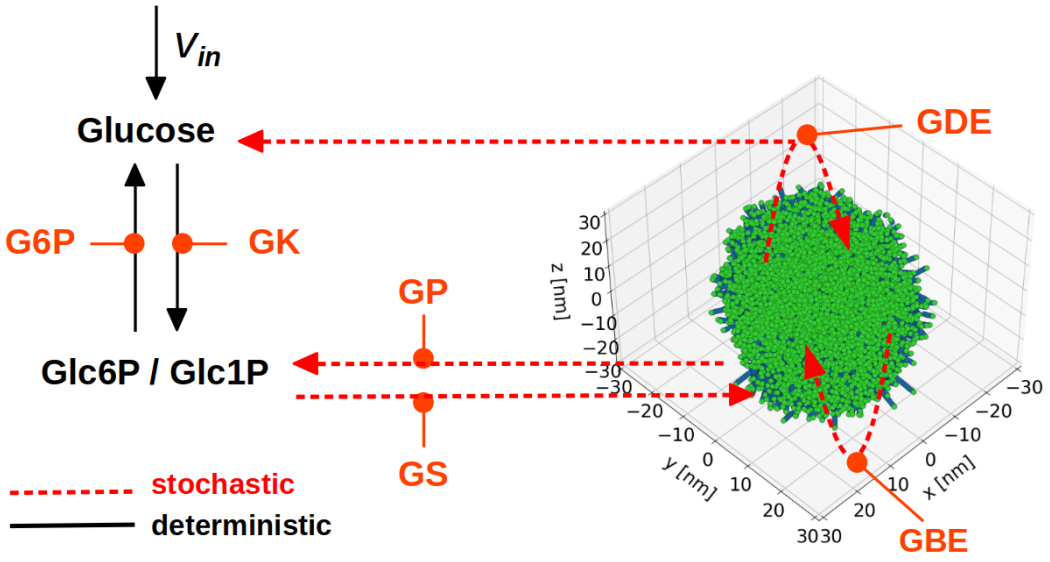

Figure 4.6: **Toy model with glycogen module.** The deterministic sub-system consists in the 3 reactions with black arrows. The import and export of glucose through v_{in} and v_{out} can be time dependant. The stochastic sub-system consists in the 4 reaction detailed in section 3. These reactions directly impact the glycogen granule structure.

$$\begin{aligned} v_{in} &= v_{in}(t) \\ v_{GK} &= k_{GK} \cdot [\text{Glc}] \\ v_{G6P} &= k_{G6P} \cdot [\text{G6P}] \\ v_{out} &= v_{out}(t) \end{aligned}$$

For the propensities of the stochastic subsystem, we want the glycogen glucose chains to be the substrate of the reactions. As each of the enzymes involved has different length specificities, we assign different status i, depending on their length, to each of the chains. To do this we calculate the degree of free polymerization DP^{free} of each of these chains. This DP^{free} is the degree of branchless polymerization between the non-reducing end of the chain and the last branch $\alpha - 1,6$ 2.1.2.

Thus when

$$\text{DP}^{\text{free}} < \min(L_{\text{GS}}, L_{\text{GDE}}),$$

the chain will be assigned the status $i = 0$. These chains are too short to react with the enzymes in presence. This status will be useful later when we introduce defects in the enzymatic mechanisms.

For

$$L_{\mathrm{GS}} \leq \mathrm{DP}^{\mathrm{free}} < \mathrm{L}_{\mathrm{GS}} + 1$$

the chain will be given the status $i = 1$. These chains are the shortest and can only react with GS and potentially GDE. GP will not be able to react on these strings. This avoids reducing a chain beyond the specific length for GS, which would make it impossible for the chain to react and would block the evolution of the structure.

For

$$\mathrm{L}_{\mathrm{GS}} + 1 \leq \mathrm{DP}^{\mathrm{free}} < \mathrm{L}_{\mathrm{GBE}},$$

the chain will be given the status $i = 2$. These chains are of intermediate length and will be able to react in an elongation reaction (GS) but also in a reduction reaction (GP). For

$$\mathrm{L}_{\mathrm{GBE}} \leq \mathrm{DP}^{\mathrm{free}}$$

the chain will be assigned the status $i = 3$. These chains are long and can be branched. They will be substrates for GS, GP and GBE. Finally if

$$\mathrm{DP}^{\mathrm{free}} = \mathrm{L}_{\mathrm{GDE}} \text{ and the chain is of type A}$$

, the chain will be assigned the status $i = 4$ These are unbranched chains having the specific length of the disconnection enzyme. They will be substrate for GDE, but can still be extended by GS. The concentration of the status i chain will be noted by C_i.

With this notation, the stochastic sub-system will be described by the following propensity:

$$\begin{aligned} a_{GS} &= k_{GS} \cdot (\mathcal{N}_a \cdot V) \cdot (C_2 + C_3 + C_4) \cdot [\mathrm{G6P}] \\ a_{GP} &= k_{GP} \cdot (\mathcal{N}_a \cdot V) \cdot (C_2 + C_3) \\ a_{GBE} &= k_{GBE} \cdot (\mathcal{N}_a \cdot V) \cdot C_3 \\ a_{GDE} &= k_{GDE} \cdot (\mathcal{N}_a \cdot V) \cdot C_4, \end{aligned}$$

Through C_i, we have the possibility to link our glycogen model to the Periodic-Coupling algorithm developed above. By default, we will use the branching parameters obtained in chapter 3. Empirically we set the volume of the system so that a $5,000$ glucose granule has a reasonable concentration ([0-10mM]). We ensure that the elongation and branching activities (GS and GBE) are higher than the breakdown activites (GP and GDE) to ensure that synthesis takes place. For generic parameters of the system, we obtain the two kinetics presented in figure 4.7.

On the top-left plot, the system is closed, no glucose equivalent can enter or leave the system ($v_{in} = 0$ and $v_{out} = 0$). One can see that glucose6-phosphate bind to the glycogen granule until the reserves in glucose units are exhausted (both glucose and glucose-6-phosphate). Once the equilibrium is reached, one can notice the presence of fluctuations due to the presence of residual degradation reactions (GP and GDE). The fact of having mass action kinetics makes the speed of synthesis exponential as the number of chains increases (zoom on the figure 4.7 top-left), before the units of glucoses eventualy become limiting. The figure on the right allows

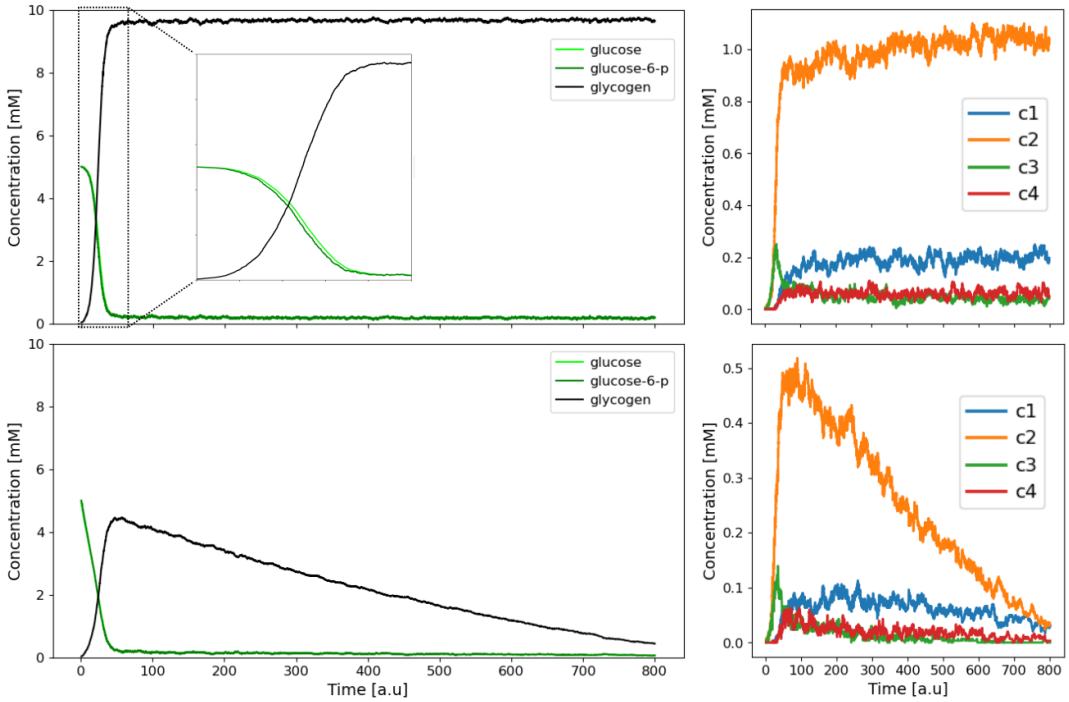

Figure 4.7: **Time courses of the system. Top-left:** Time course of the 3 metabolites for a closed system with no v_{in} nor v_{out}. The dotted square shows a zoom for the first 50 seconds of the simulation. **Top-right**: Corresponding evolution of the different chain types concentration. **Bottom-left:** Time course of the 3 metabolites for an open system with v_{out} but no v_{in}. When most of the glucose has been fixed onto glycogen, the molecule start being degraded. **Bottom-right**: Corresponding evolution of the different chain types concentration.

us to track the evolution of these chains as a function of time. During the synthesis phase (0-50sec), it can be seen that the elongation allows the formation of substrate for the branching (green peak). When equilibrium arises, there will not be enough elongation to create new chains for branching. Existing ones will react and branched or be reduced by GP, reducing their focus to almost zero. Therefore we logically see the appearance of shorter chains (chains of type 1 and 4). Indeed, even if the GP activity is residual, GS is limited by the quasi absence of glucose residue. Thus the relative activity of GP compared to GS increases which reduces the size of the chains. On the bottom part of the figure, we have simulated the same system with the possibility of glucose leaving the system. This leads to two differences, the maximum concentration of glycogen is logically lower because part of the glucose has already left the system. When all the glucose has either been exported or synthesized, the residual activity of GP and GDE becomes more important than that of GS and GDE triggering the degradation of the molecule. Concerning the concentrations of the different types of chains (Fig 4.7 Bottom-right), they decrease as the glycogen is degraded. Another important aspect of coupling the

glycogen structure model to our kinetic algorithm (Periodic-Coupling) is that we can access the structural properties of glycogen over time for, in principle, any kinetic model. The parameters used in these simulations are arbitrary and do not reflect physiological quantities. However, as a proof of concept we have looked at some of the structural properties of glycogen in the two simulations discussed above. The figure shows 4.8 the evolution of the A:B ratio (top-left), the radius of the molecule (top-right), the occupancy (middle-left), the average chain length (middle-right) and the branching degree (bottom-left).

Several interesting observations emerge from these plots. The A:B ratio seems to be dependent on the state of the system. Indeed, in chapter 3 we have seen that this ratio was 1 when $L_{\text{leftover}} = L_{\text{transferred}}$. This result was obtained during a pure synthesis phase. On the figure here, we see that this ratio also seems to be equal to 1 at the end of the synthesis phase. However, when the system reaches equilibrium or starts to degrade (black and ref curve respectively), the ratio seems to drop around 0.75. A possible explanation for this is a relative increase in debranching.

The evolution of the radius is also surprising. Indeed, while the granule is at equilibrium, the radius increases (black curve). A possible explanation is due to the equilibrium between branching and disconnection. Indeed, the branching reaction is subject to steric hindrance, and these reactions will therefore tend to take place more often towards the outside of the granule, where the density is lower. The debranching reaction, on the other hand, has no spatial constraint. Thus where GDE will degrade homogeneously over the whole granule, GBE will rebuild rather towards the outside of the granule, lowering its total density and increasing its radius. This observation probably does not occur in-vivo because GDE should also undergo steric constrains. However, if we imagine that GDE penetrates the molecule deeper than GBE, such a result could take place *in-vivo*. This previous observation can logically also be seen on occupancy plot (middle-left, black curve), which decreases while the glycogen content is constant. This is directly due to the increase of the radius. The average chain length is higher during the synthesis phase than during equilibrium or degradation. This is logically caused by the activity of GS during synthesis which allows the elongation of the chains. However, the gap between the peak and the more or less constant part could be reduced by increasing the branching activity. If during the synthesis, branching occurs instantaneously when the chain is sufficiently long, it would dilute the effect of having and higher activity in GS than GP. Finally, we observe that the degree of branching increases when we leave the synthesis phase. This is directly due to the decrease in chain length and therefore to the decrease in the number of bonds α-$1,4$.

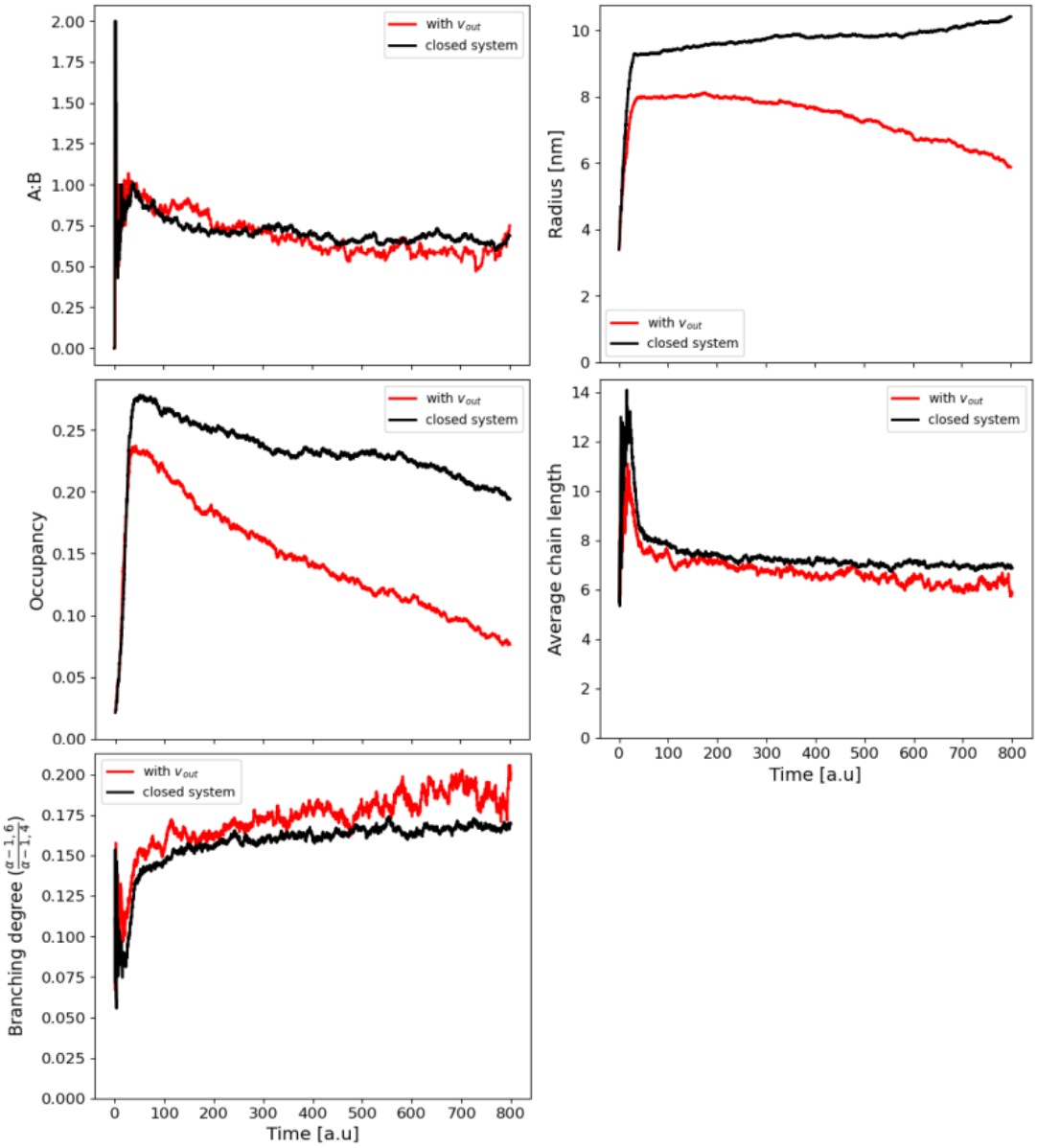

Figure 4.8: **Time dependency of various structural features. Top-left:** The A to B ratio decreases during the synthesis. **Top-right**: The radius of the granule. When the system is open, degradation is possible leading to a decrease in the radius. When the system is closed, a counter-intuitive increase of the radius appears, while the glycogen content is constant. **Middle-left:** The occupancy as a function of time. **Middle-right**: The average chain length is slightly lower with v_{out} due to the degradation. **Bottom-left**: The branching degree is higher with v_{out} because the average chain length is lower.

4.2 A reduced model for glycogen metabolism

Now that we are equipped with a set of methods and tools that allow us to couple the structure of glycogen to a kinetic model, we will orient the study towards something more quantitative, with the ultimate goal of being able to study certain glycogen-related diseases. To do this, we will build a kinetic model of glycogen metabolism from an existing one. Our basis will be the one presented in an article entitled "Quantifying the contribution of the liver to glucose homeostasis: a detailed kinetic model of human hepatic glucose metabolism" [König et al., 2012]. The authors present a ODEs model of human hepatic glucose metabolism, including glycolysis, gluconeogenesis, and glycogen metabolism. One major aspect of the model is that it includes hormonal control of these pathways by insulin, glucagon, and epinephrine, allowing investigations how liver cells respond to changes in plasma glucose levels. The results are in good agreement with experimental data on the contributions of each pathway to hepatic glucose production and utilization.

Exclusion of model components

In order to establish a kinetic model for glycogen metabolism, we excluding numerous reactions that were not suitable for our need. In the model presented by König et. al. it is not clear how some reactions are taken into account. For example, the reaction:

$$\mathrm{ATP} + \mathrm{AMP} \overset{\mathrm{AK}}{\longleftrightarrow} 2\mathrm{ADP}$$

is described in the model with an associated rate law, whereas the article states that the concentrations of AMP, ADP and ATP are all fixed (see https://doi.org/10.1371/journal.pcbi.1002577.s013). Moreover, the part considering the TCA cycle of their model is not very precise. On the one hand, many of the compounds in the mitochondrion are kept at constant concentration, resulting in reactions that are not effectively balanced. On the other hand, some reactions have their $v_{\max} = 0$, resulting in non-effective reactions. We tried to implement the complete model in two ways. A first one were we fixed the concentrations of the supposed constant metabolites. A second one were all metabolite concentrations were variables in the ODE system. In both cases the steady-state could not be reached, some compounds leaving the system through reaction v_{31}. In addition, the phosphate pool is not a conserved moiety. This results in an infinite accumulation of the compound. Despite trying different modifications we could not reproduce the results shown in the article. To build our reduced model, we therefore removed all the reactions in the mitochondrion compartment as well as all reactions from lactate to Fructose-6-phosphate. Another aspect which does not fit our needs is how glucose is considered. The external glucose, which regulates the hormonal signal, is also fixed in the model. The underlying hypothesis is that other parts of the body contribute to maintain glucose homeostasis. This is one of the aspects that interests us most. How does the external glucose level evolve according to the glycogenolysis and glycogenesis pathways? Fixing the external glucose level results in never changing hormone levels, whereas our interest

lies precisely in capturing the reciprocal effect between a dynamic hormonal response and the evolution of blood glucose level.

Selection and addition of model components

In spite of all the issues mentioned above, some aspects of the model are very interesting. The kinetic laws of chemical reactions, sometimes complex, and their respective parameters are very well detailed. They take into account two important regulatory mechanisms. The first one is of alosteric type, for which enzyme activities will be regulated, principally, by the glucose-6-phosphate level. The second one, is the modulation of activities in response to Insulin and Glucagon signal. Indeed, the corresponding inter-convertible enzymes, according to their state of phosphorylation γ, directly depending on hormones level, will be more or less active. This γ function is modeled with a logistic functions as follow and is depicted in figure 4.9-left:

$$\gamma = \frac{1}{1 + \exp \mathrm{Glc} - \mathrm{Glc}_0}, \tag{4.1}$$

where $\mathrm{Glc}_0 = 4.6$ is a glucose threshold value. The reactions involved will always be modeled through a flux having the following form [König et al., 2012]:

$$v = \gamma \cdot v_p + (1 - \gamma) \cdot v_{dp}. \tag{4.2}$$

The phosphorylation state, represented by γ, is directly dependent on the glucose level in the blood. When the blood glucose is high, γ will be low, resulting in a dephosphorylate state with v_{dp} dominating over v_p. Conversely, when the blood glucose is low, γ will be high, leading to a phosphorylate state where v_{dp} dominating over v_p. In our case, we adjusted a simplified sigmoid function to represent the phosphorylation behaviour depicted in the article. This γ function will be incorporated in our ODEs system, while in the article γ is kept constant with a value that depends on the regime of interest (glycogenolysis or glycogen synthesis).

In our model, the external glucose concentration is no longer constant and becomes a variable, which allows us to evaluate the value of $\gamma(\mathrm{Glc})$ at each moment of the simulation. To this we have added two reactions, v_{in} and v_{out} which allows to account for the import and export of glucose to other parts of the organism (Fig. 4.9-left and -middle respectively). For the import we used a constant influx to mimic periods of "feeding". To avoid numerical issues from abrupt changes in v_{in}, we used an approximation of an Heaviside function:

$$v_{in} = \frac{k_{in}}{1 + \exp^{a \cdot (t - t_{\mathrm{stop}})}}, \tag{4.3}$$

where t_{stop} is the time at which glucose import should stop, and a a factor that will act on the sharpness of the Heaviside function. This time dependant function will allows us to mimic some feeding/fasting cycles. To model the glucose export, one could think of several rate laws. In our study we make the following assumption: At rest, the demand in glucose from outside the liver is constant. One can see it as a maintenance glucose export, without considering any body extra activities. This is why we choose not to model this export as a mass action

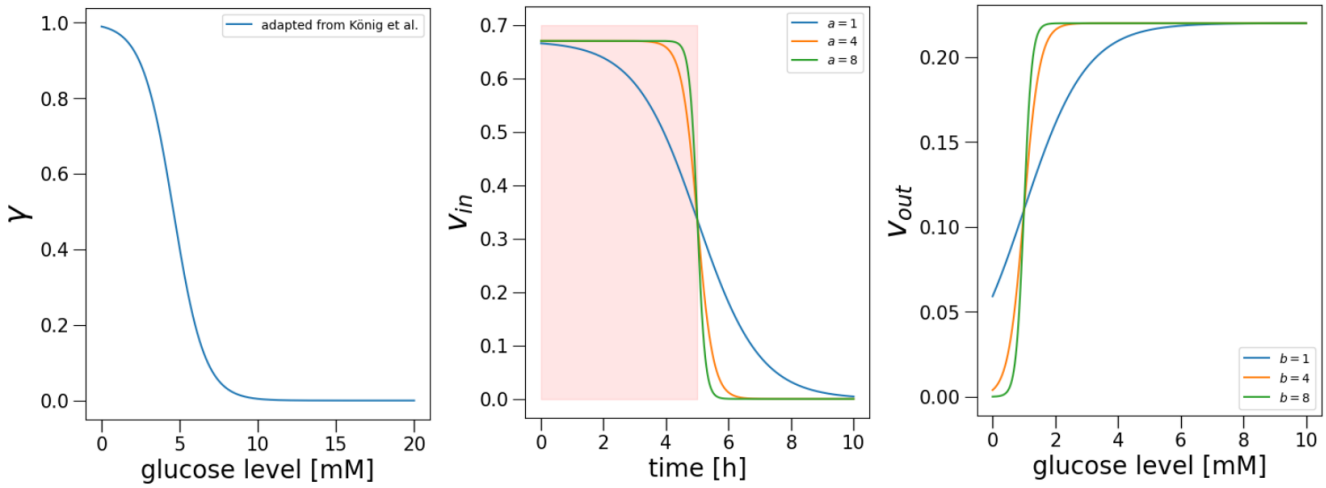

Figure 4.9: **Phosphorylation function γ, rates of glucose export and import. Left:** γ as a function of glucose level. **Middle:** Import of glucose as function of time. Rate is constant ($v_{in} = 0.67$) during a feeding period $t_{\text{stop}} = 5$h (red area), and start decreases after. The different values of a illustrate different sharpness coefficient for the Heaviside approximation. **Right:** Export of glucose as function of glucose level. Rate is constant ($k_{out} = 0.22$) above a critical value $\text{Glc}_c = 1mM$, and start decreases above this value. The different values of a illustrate different sharpness coefficient for the Heaviside approximation.

kinetics nor as a Michaelis-Menten like kinetics. We used an almost constant outflux, with a dependency on glucose at low value only. For this we also chose an approximation of an Heaviside function. We depicted the rate law behaviour in Figure 4.9-right:

$$v_{out} = \frac{k_{out}}{1 + \exp^{-b\cdot(\text{Glc}-\text{Glc}_c)}}, \tag{4.4}$$

where Glc_c is a critical level of glucose and b constant which act such as a.when blood glucose concentrations are very low, this fuelling can not be maintain and the export drop to $v_{out} = 0$. It is important to exercise caution when initializing the system, as a glucose concentration of 0 may result in negative concentration due to the fact that the outflux is never truly 0.

In our model, we re-introduced ATP, ADP, UTP and UDP as variable concentration in the reactions we kept. We identified that the reversible reaction of

$$\text{ATP} + \text{AMP} \xleftrightarrow{\text{AK}} 2\text{ADP},$$

catalyzed by Adenylat kinase (AK) in their model, is problematic. During synthesis ATP and UTP are highly demanded, however the ATP turnover is not efficient enough. It is not clear

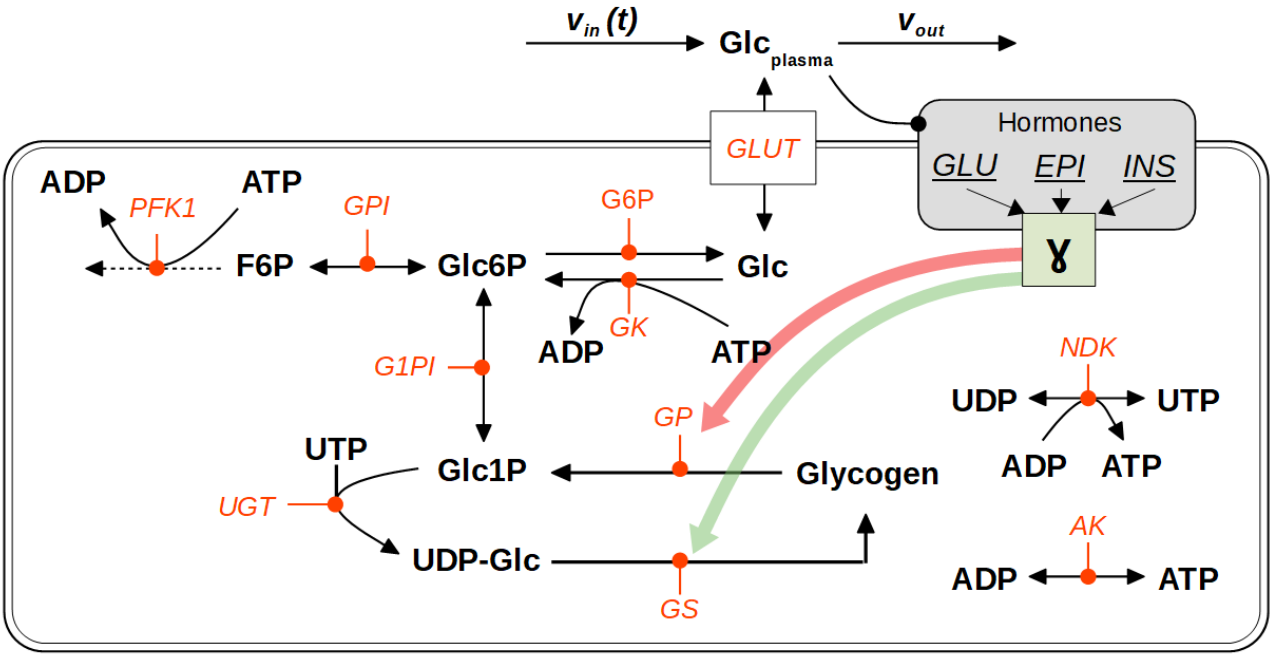

Figure 4.10: **Scheme of the reduced König Model.** Import and export of glucose are modeled through v_{in} and v_{out} outside the cytosol. This plasma glucose ($\mathrm{GLC}_{\mathrm{plasma}}$) impact the value of the phosphorylation state γ, mimicking the behaviour of the hormones. Depending on the value of γ, glycogen synthesis or glycogen breakdown will be enhanced.

to us the exact reason why the reverse reaction is effectively not working. We simplified it by removing AMP from the system with the following reaction:

$$\mathrm{ATP} \overset{\mathrm{AK}}{\longleftrightarrow} \mathrm{ADP},$$

but keeping the same kinetic rate. The AMP concentration in the rate-law becomes a fixed parameter. After this simplification, The system works properly in both the synthesis and the degradation phase. It comes at the cost of loosing regulatory effect from AMP.

Finally, we also added a reaction that export fructose-6-phosphate, which end the glycolytic pathway here. This export is modelled with a mass-action reaction rate:

$$v_{out}{}^{\mathrm{F6P}} = k_{out}{}^{\mathrm{F6P}} \cdot \mathrm{F6P}. \tag{4.5}$$

All the reactions (13 in total) in our reduced system are presented in Fig 4.10, together with the metabolites (11 in total). All the rate-laws (except v_{in}, v_{out} and $v_{out}{}^{\mathrm{F6P}}$) are taken from König et al., 2012 can be found at https://doi.org/10.1371/journal.pcbi.1002577.s014.

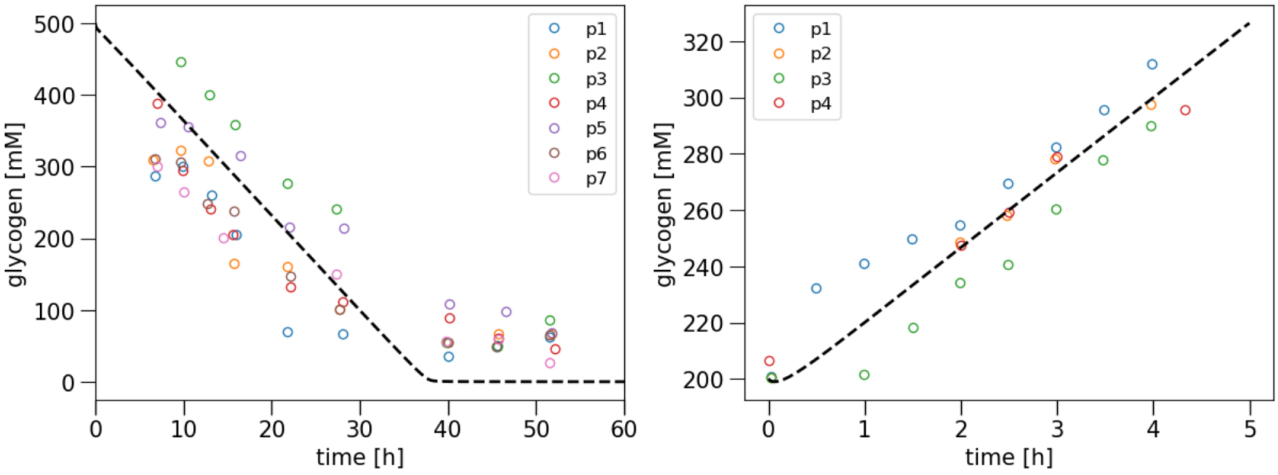

Figure 4.11: **Time courses of glycogen degradation and synthesis in the reduced model versus experimental data. Left:** Degradation of glycogen. Simulation (dotted line) starts with 500mM of glycogen. Each color point corresponds to individual patient. **Right:** Synthesis of glycogen. Simulation (dotted line) starts with 200mM of glycogen. Each color point corresponds to individual patient.

The only unknown kinetic parameters are those introduced through the glucose import and export reactions, as well as the fructose-1-Phosphate export. To have realistic ones, we adjusted k_{in} and k_{out} so that simulation of glycogen synthesis and degradation fit the data used in König et al., 2012 which originates from Ferrannini et al., 1985; Magnusson et al., 1992; Radziuk and Pye, 2001; Rothman et al., 1991; Taylor et al., 1996. We found $k_{in} = 0.67$ mM.min^{-1}.kg^{-1} and $k_{out} = 0.22$ mM.min^{-1}.kg^{-1}. ${k_{out}}^{\mathrm{F6P}}$ is set to 0.1 min^{-1}.kg^{-1}. All values below 0.5 min^{-1}.kg^{-1} have a very little to no impact on the evolution of other metabolites in the system. ${k_{out}}^{\mathrm{F6P}}$ can later be changed, together with Eq.4.5 if one wants to consider downstream changes in the glycolysis pathway. Time courses of glycogen degradation and synthesis together with the corresponding experimental data are presented in Fig 4.11.

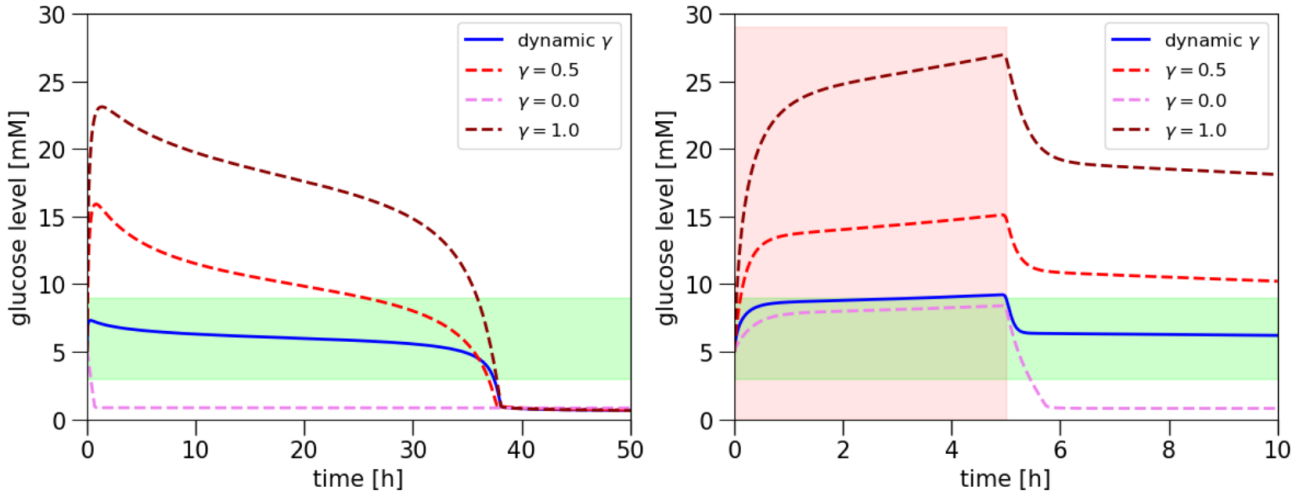

Figure 4.12: **Time courses of glucose level during fasting and feeding periods.** The green area indicates the glucose level range between hypoglycemia and hyperglycemia (3 mM and 9 mM respectively). The blue line shows the glucose level with a glucose dependent γ function. Dark red, red and pink curves show the glucose level with a constant γ function ($\gamma = 1$, $\gamma = 0.5$ and $\gamma = 0$ respectively). **Left:** Evolution of glucose during fasting. **Right:** Evolution of glucose during a feeding phase. The red area indicates the period of glucose intake.

Testing the model

In order to test how the γ function acts on the system, we looked at how glucose evolves during fasting and feeding phases (Fig. 4.12). γ have the expected regulatory effect. The blue curve shows us how the glucose level is kept constant by this hormonal regulation. We did the same simulations with constant $\gamma = 0.5$ (red curve). It corresponds more or less to the absence of hormonal regulation with Equation 4.2 which becomes:

$$v = \frac{v_{dp} + v_p}{2}.$$

We have also simulated the case of $\gamma = 0$ and $\gamma = 1$, which correspond to $v = v_{dp}$ and $v = v_p$ respectively. The green area indicates the range between 3 and 9mM. This defines the expected range of blood glucose level [Nuttall et al., 2008]. Fig. 4.12-left shows the evolution of glucose level during fasting ($v_{in} = 0$). Simulation starts with a full storage of glycogen (500 mM). One can observe for $\gamma = 0.5$ (red curve), a too important contribution of the glycogen degradation to the blood glucose level. The time interval for which the glucose level is in the right range is much shorter (26-38h) than when hormonal response is taken into account (0-38h). The case where $\gamma = 0$ would correspond to a completely dephosphorylated state, normally found when the glucose level is very high. In this situation the enzymes responsible for synthesis are hardly activated, resulting in the impossibility of using glycogen to supply glucose to the blood. On the other hand, when $\gamma = 1$, the degradation is too active compared to the need for glucose and this results in an excessively high glucose level.

We also performed the same study (Fig.4.12-right) during a feeding period. During glucose intake (red area), γ is close to zero. The blue curve therefore coincides with the pink one. For $\gamma = 0,5$ (red curve) and $\gamma = 1$ (dark-red curve), the degrading enzyme GP is too active as it was during the degradation phase, resulting in a high glucose level. When the intake stops, we find the same behavior as in (Fig.4.12-left). It is also interesting to note that the glucose level after a meal is estimated at 9 mM [Nuttall et al., 2008], which is approximately the value found here.

To further highlight the effect of the hormones, one can modified Eq 4.4 by adding a mass-action term:

$$v_{out} = \frac{k_{out}}{1 + \exp^{-b\cdot(\mathrm{Glc}-\mathrm{Glc}_c)}} + k \cdot \mathrm{Glc}, \tag{4.6}$$

The idea of this new term is to mimic the effect of refilling other parts of the body, together with the maintenance term introduced in Eq 4.4.

Figure 4.13 shows how glucose and glycogen evolves during fasting under these conditions. Simulation starts with a full storage of glycogen (500 mM). Since excess in glucose level is now used to fuel other part of the body through the mass-action kinetic term in Eq.4.6, glycogen is reduced too quickly when not regulated ($\gamma = 0,5$), resulting in both a shorter total supply duration (equivalent to glucose production duration) and a shorter duration of glucose level in the correct range [3 mM - 9 mM]. This is a striking case of the hormones effect on glycogen storage management.

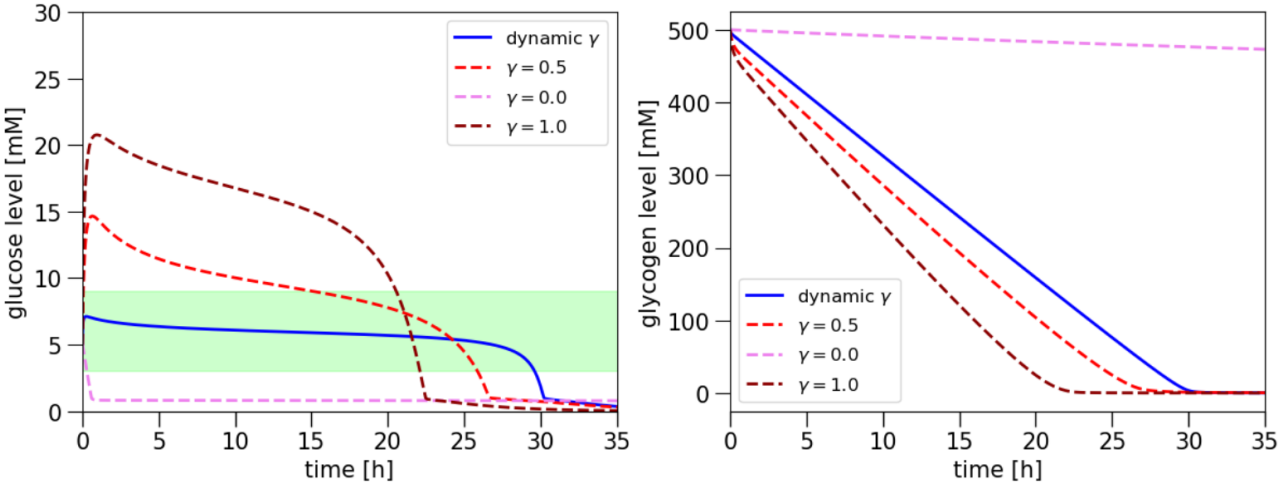

Figure 4.13: **Time courses of glucose and glycogen level during fasting in the reduced model with refilling** The green area indicates the glucose level range between hypoglycemia and hyperglycemia (3 mM and 9 mM respectively). The blue line shows the glucose and glycogen level with a glucose dependent γ function. Dark red, red and pink curves show the glucose and glycogen level with a constant γ function ($\gamma = 1$, $\gamma = 0.5$ and $\gamma = 0$ respectively). **Left:** Evolution of glucose during fasting. **Right:** Evolution of glycogen during fasting.

4.3 Connecting the glycogen module to the kinetic model

The subsection 4.1.3 presents a methodology for coupling a glycogen structural model with a simplified ordinary differential equation (ODE) system. The approach is extended to the reduced model of glycogen metabolism introduced in section 4.2, where only the stochastic reactions catalyzed by glycogen synthase, glycogen branching enzyme, glycogen phosphorylase, and glycogen debranching enzyme are considered, while all other reactions are modeled deterministically. We recall that this chapter aims to investigate the interplay between glycogen structure and metabolism, requiring the consideration of glycogen chains for the four stochastic reactions.

Verfication

Prior to outlining the suitable rate laws, it is essential to verify that utilizing identical reaction rates as outlined in 4.2 results in the same outcomes when connecting the glycogen granule model. In the ODE model, glycogen synthesis occurred solely through the catalytic activity of glycogen synthase (GS), whereas in the present study, both GS and glycogen branching enzyme (GBE) are involved. In this regard, we combine the branching and elongation processes, previously considered as distinct in König's model, into a single reaction:

$$\text{UDP}-\text{glc} \xrightarrow{\text{GS}} \text{UDP} + \text{glycogen}, \tag{4.7}$$

where the glycogen concentration refers, in fact, to the concentration of "glucose fixed." To enable a comparison between different reaction rates, we have included the branching reaction, which is catalyzed by GBE, within the elongation reaction catalyzed by glycogen GS. By doing so, we can utilize the same reaction rate as described in Eq.4.7. The combined reaction is expressed as follows:

$$\text{UDP}-\text{glc} + \text{glycogen}(\text{n}, \text{N}_b) \xrightarrow{\text{GS}} \text{UDP} + \text{glycogen}(n+1, N_b) \left(\xrightarrow{\text{GBE}} \text{glycogen}(n+1, N_b+1) \right), \tag{4.8}$$

where glycogen molecules are described through glycogen(n, N_b), a granule made of n glucose units arranged in N_b branches. With the same rate as in Eq.4.7, one glucose will be fixed to our glycogen granule producing glycogen(n+1, N_b). To account for branching, we incorporate a probability for GBE to act during the elongation reaction, which will sometime produce glycogen(n+1, N_b+1). Since the branching reaction does not contribute to the stoechimetries, it is a simple way to use the same rate-law in the two cases. The situation is a bit different for the reduction reaction catalyzed by GP in the ODE model:

where glycogen molecules are represented as glycogen(n, N_b), which is a granule consisting of n glucose units organized into N_b branches. Utilizing the same reaction rate as described in Eq.4.7, the addition of one glucose unit to the glycogen granule results in the formation of glycogen(n+1, N_b). To account for branching, we incorporate a probability factor for the action of GBE during the elongation reaction, resulting in the production of glycogen(n+1,

N_b+1) in certain instances. Since the branching reaction does not affect the stoichiometries, incorporating it in this manner allows us to utilize the same rate law for both elongation and branching reactions.

However, the situation differs for the reduction reaction catalyzed by glycogen phosphorylase (GP) in the ODE model:

$$\text{glycogen} + \text{P} \xrightarrow{\text{GP}} \text{Glc}-1\text{-P}, \tag{4.9}$$

while unlike glycogen branching enzyme (GBE), the debranching reaction (GDE) changes the number of fixed glucose units, denoted as n. Therefore, when accounting for the reaction rates, we cannot combine glycogen phosphorylase (GP) and GDE with a single reaction rate. Instead, the rate of GP will account for either GP or GDE individually, as follows:

$$\text{glycogen}(n, N_b) + \text{P} \xrightarrow{\text{GP}} + \text{Glc}-1\text{-P} + \text{glycogen}(n-1, N_b) \tag{4.10a}$$

OR

$$\text{glycogen}(n, N_b) + \text{P} \xrightarrow{\text{GP}} + \text{Glc}-1\text{-P} + \text{glycogen}(n-1, N_b-1) \tag{4.10b}$$

where Equation 4.10a represents the actual reaction catalyzed by GP, which reduces a glycogen chain by one unit and releases a glucose-1-phosphate molecule. On the other hand, Equation 4.10b represents the debranching reaction. It is important to note that the actual debranching reaction releases a glucose molecule instead of a glucose-1-phosphate, but for the sake of comparison, we consider the reaction to release a glucose-1-phosphate instead. The inclusion of these non-standard reactions allows for the use of the same rate-laws in both the ODE model and the hybrid model. This hybrid model, which combines ODE and stochastic reactions and employs identical rates, will be referred to as the "comparable model" henceforth, whereas the pure ODE model will be referred to as the "ODE model". The comparable model integrates the approach presented in subsection 4.1.2, which establishes the equivalence of the Periodic-Coupling algorithm to regular ODE solving, with the model introduced in subsection 4.1.3 that does not consider glycogen chains as substrates. The primary goal of constructing this model is to ensure the reproducibility of the pure ODE system when combining complex kinetics with our Periodic-Coupling algorithm and the granule structural model for glycogen.

Figure 4.15 displays the time courses of blood glucose (top-left panel) and glycogen (top-right panel) for a synthesis phase (150 minutes, with glucose input rate $v_{in} = 0.7$ mM.min^{-1}) followed by a degradation phase (glucose input rate $v_{in} = 0$ mM.min^{-1}), with a constant small export of blood glucose ($v_{out} = 0.24$ mM.min^{-1}). The results of both the ODE model (black curves) and the comparable model (red curves) are depicted in the figure, and they exhibit identical behavior, as expected.

Appropriate rate-laws for GS, GBE, GP and GDE

Now that the full model (structural model, reduced glycogen metabolic model and Periodic-Coupling algorithm) has been verified to be working properly, appropriate rate-laws need to be found that depend on glycogen chains rather than glycogen glucose content. These kinetic laws should approximate the behavior observed in section 4.2 when the enzymes GS, GBE, GP, and GDE act normally (in contrast to the possible introduction of defects in their respective activities). A correspondence will be established between the flux $v_{GS}^{könig}$, which describes the total synthesis reaction without distinction between elongation and branching, and the individual reactions of elongation and branching. Similarly, a relation will be established between the total degrading reaction $v_{GP}^{könig}$ and the individual processes of reduction and debranching. It is worth noting that in our glycogen model, as discussed in subsection 4.1.3, chains are categorized according to their status, denoted as $i \in 1, 2, 3, 4$. The concentration of a chain with status i is denoted as c_i. Moreover, it should be mentioned that each of the four enzymes is capable of utilizing different combinations of chain status as their respective substrate:

$$c_{GS} = c_1 + c_2 + c_3 + c_4 \tag{4.11}$$

$$c_{GBE} = c_3 \tag{4.12}$$

$$c_{GP} = c_2 + c_3 \tag{4.13}$$

$$c_{GDE} = c_4 \tag{4.14}$$

Elongation (GS) The kinetic law in König et al., 2012, v_{GS} is written as follows:

$$v_{GS} = \gamma \cdot {v_p}^{GS} + (1 - \gamma) \cdot {v_{dp}}^{GS} \tag{4.15a}$$

$${v_p}^{GS} = v_{max} \cdot f_{max} \cdot \frac{[\text{UDP-glc}]}{[\text{UDP-glc}] + \left(\frac{{k_1}^p}{{k_2}^p + [\text{G6P}]}\right)} \tag{4.15b}$$

$${v_{dp}}^{GS} = v_{max} \cdot f_{max} \cdot \frac{[\text{UDP-glc}]}{[\text{UDP-glc}] + \left(\frac{{k_1}^{dp}}{{k_2}^{dp} + [\text{G6P}]}\right)} \tag{4.15c}$$

$$f_{max} = \left(1 + {k_1}^{max}\right) \cdot \frac{C - [\text{glycogen}]}{C - [\text{glycogen}] + {k_1}^{max} \cdot C} \tag{4.15d}$$

where C is a capacity term of glycogen (500 mM). One can see that there is no positive dependency on glycogen content, but only a product inhibition term ($C - [\text{glycogen}]$). The real process of elongation by GS requires two substrates: a glycogen molecules (regardless of it description, total glucose content or chain content) and a UDP-glucose molecules. However, when one look at $v_g s$, one can see that glycogen doesn't appear as substrate. Phenomenologically this means that the limiting substrate is UDP-glucose and not the glycogen content. To account for glycogen, Eq.4.7 can be cahnged into the raction depicted in Fig.4.14-elongation.

In the provided equation, C represents the capacity of glycogen (500 mM). One can see the absence of a positive dependence on glycogen content, and a term for product inhibition

(C − [glycogen]). The actual process of elongation by GS involves two substrates, namely a glycogen molecule (independent of its characterization in terms of total glucose content or chain content) and a UDP-glucose molecule. However, v_{GS} shows that glycogen does not feature as a substrate. This implies that the limiting substrate is UDP-glucose rather than glycogen content. To incorporate glycogen, Eq. 4.7 can be modified to the reaction illustrated in Fig. 4.14-elongation.

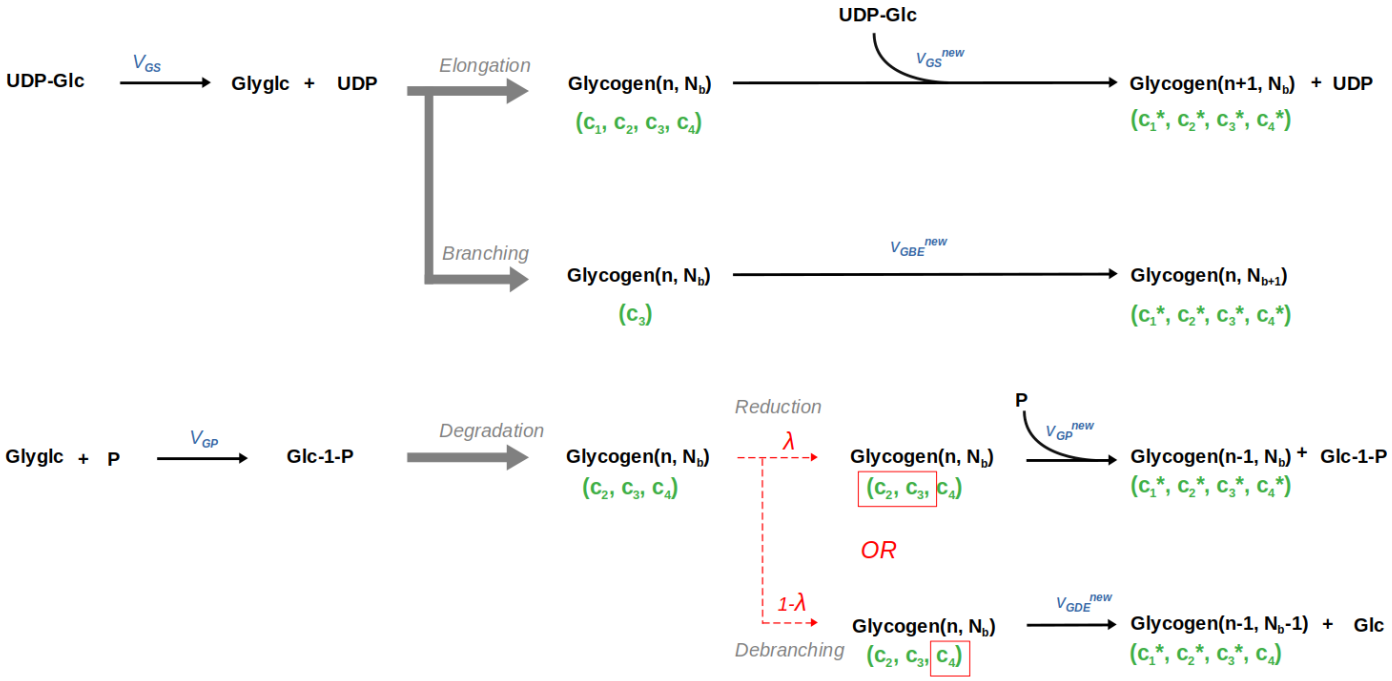

Figure 4.14: **New reactions considering glycogen chains as substrate.** The marker "*" highlight potential change in the status of the chains.

In these new reactions, the quantity of glucose fixed into glycogen, represented by *glyglc*, is no longer present. Instead, glycogen is considered as a metabolite consisting of n glucose molecules arranged in N_b branches. The GS enzyme now requires both glycogen and UDP-glucose as substrates to catalyze the reaction. Above the glycogen compound, the involved chains type are highlighted in green. The star markers in the products describe potential modifications of the chain type content. When GS takes place, every chains type can changed, which explain they are all marked.

The idea is the following: We will use the rates from König and multiply it by a factor going from 0 to 1, mimicking a Michaelis-Menten-like kinetics::

$$v_{GS}{}^{\text{new}} = \tilde{v} \cdot \frac{[\text{glycogen}]}{K_{GS} + [\text{glycogen}]}.$$

where $\tilde{v}$ can be seen as a varying $v_{\max}$ which is equal to ${v_{GS}}^{\text{König}}$:

$$ {v_{GS}}^{\text{new}} = {v_{GS}}^{\text{König}} \cdot \frac{[\text{glycogen}]}{K + [\text{glycogen}]}. \tag{4.16} $$

To work with a reasonable value for K_{GS}, we assume that during the synthesis phase presented in Fig.4.11-right, which start with an initial glycogen concentration of [glycogen] = 200 mM, the content in glycogen (as substrate) was sufficient for the enzyme to work close to $\tilde{v}$. For this concentration, we assert that ${v_{GS}}^{\text{new}}$ is equal to 99% of $\tilde{v}$. This allows us to calculate a corresponding K value:

$$ K_{GS} = \frac{\tilde{v} - 0.95 \cdot \tilde{v}}{0.99\tilde{v}} \cdot [\text{glycogen}] \Leftrightarrow K_{GS} = 2.02\ mM \tag{4.17} $$

We then convert this value in chains concentration. Because all of the chains are substrate for GS, we used the average chain length of 12.09 (table 1, page 12 of the article) found in chapter 3, for "healthy" glycogen. There is therefore a factor 12.09 between the concentration of chains c_{chain} and the concentration of glucose fixed into glycogen [glycogen]. It gives us the effective value of this Michaelis-Menten-like constant considering chain as substrate ${K_{GS}}^{\text{chain}} = 0.167\ mM$. The new rate-laws for GS reads:

$$ {v_{GS}}^{\text{new}} = \frac{c_{\text{chain}}}{c_{\text{chain}} + {K_{GS}}^{\text{chain}}} \cdot \left(\gamma \cdot {v_p}^{GS} + (1-\gamma) \cdot {v_{dp}}^{GS} \right) $$

which can be rewrite as a function of each chain type according to subsection 4.1.3:

$$ {v_{GS}}^{\text{new}} = \frac{c_1 + c_2 + c_3 + c_4}{c_1 + c_2 + c_3 + c_4 + {K_{GS}}^{\text{chain}}} \cdot \left(\gamma \cdot {v_p}^{GS} + (1-\gamma) \cdot {v_{dp}}^{GS} \right) \tag{4.18} $$

The same ${K_{GS}}^{\text{chain}}$ will be used independently for the GS and GBE. We will make the distinction between enzymes substrate specificities by discriminating the chains according to their status introduced in subsection 4.1.3. We are conscious it could not reflex exactly the precise kinetics at small glycogen concentrations, but allows in a simple way to account for chains concentrations. Similar to GSD type 0, the model does not contain reverse glycolysis or alternative pathways for glucose production, making glycogenolysis is the sole source of glucose production. The rate-laws governing this process are highly responsive to hormonal regulation, effectively mitigating any potential deviations in glycogen breakdown or synthesis fluxes, forcing the "correct" regulation of glucose.

Branching (GBE): In Chapter 3, our best-fit procedure shows that the ratio $\Gamma = v_{GS}/v_{GBE}$ should be equal to 0.6 in order to reproduce the chain length distribution of measured in Sullivan et al., 2014. We will use this result and write the rate-law so that it fulfill:

$$ v_{max}^{GBE} = \frac{v_{max}^{GS}}{\Gamma} $$

, together with the addition of the dependency in concentration of chains with status $i = 3$ leads to:

$$ v_{GBE}^{new} = \frac{c_3}{c_3 + {K_{GS}}^{\text{chain}}} \cdot \frac{1}{\Gamma} \cdot \left(\gamma \cdot {v_p}^{GS} + (1-\gamma) \cdot {v_{dp}}^{GS} \right) \tag{4.19} $$

Degradation by GP and GDE: We proceed a bit differently for the degradation. Because both the reduction by GP and the debranching by GDE release one glucose, we can not simply approximate the v_{GP}^{new} "max" as $v_{GP}^{\text{König}}$ as we have done for GS. Indeed, it would lead to an excess effective degradation

$$v_{GP}^{new} + v_{GDE}^{new} > v_{GP}^{\text{König}}$$

when chains are abundant. We instead consider in a first step the total flux resulting from the two degrading process as:

$$v_{\text{degradation}} = \frac{(c_2 + c_3) + c_4}{\big((c_2 + c_3) + c_4\big) + {K_{\text{deg}}}^{\text{chain}}} \cdot {v_{GP}}^{\text{König}} \tag{4.20}$$
$$ {v_{GP}}^{\text{König}} = \gamma \cdot {v_p}^{GP} + (1 - \gamma) \cdot {v_{dp}}^{GP}$$

We have introduced ${K_{\text{deg}}}^{\text{chain}}$, the Michaelis-Menten-like constant for degradation. Without prior knowledge about it we will use ${K_{\text{deg}}}^{\text{chain}} + {K_{GS}}^{\text{chain}}$ unless otherwise stated. When this reaction occurs, we then look at the ratio Λ between the substrates type for GP and for GP + GDE. We recall that the chains that are substrate for reduction are of types $i \in \{2, 3\}$ and the one for GDE of type $i = 4$. Therefore Λ writes:

$$\Lambda = \frac{c_2 + c_3}{c_2 + c_3 + c_4} \tag{4.21}$$

It allows to discriminate between the two reactions. GP will occur with a probability Λ and GDE with a probability $(1 - \Lambda)$ (Figure 4.14).

The total reaction rate of degradation can be seen has the sum of the two sub reaction rates:

$${v_{GP}}^{\text{new}} = \Lambda \cdot v_{\text{degradation}} \tag{4.22}$$
$${v_{GDE}}^{\text{new}} = (1 - \Lambda) \cdot v_{\text{degradation}} \tag{4.23}$$

Kinetic laws allowing for the consideration of different types of glycogen chains as substrates of the reactions have now been obtained. For this purpose, we have used the fact that the reactions present in König et al., 2012 for elongation and reduction intrinsically contain information on branching and debranching rates. Effective Michaelis-Menten constants for chain concentrations were evaluated using results obtained in Chapter 3, followed by expressing the flux v_{GBE} as a function of v_{GS} using the constant Γ. Finally, Λ was introduced to express reduction and debranching as a function of $v_{\text{degradation}}$.

We recall that Fig 4.15 shows the evolution of both blood glucose and glycogen (top-left panel and top-right panel respectively). The condition of the simulations is introduced at the end of subsection 4.3. The results from the final model introduced above is shown in blue, where the rate-laws of the 4 enzymes (GS, GBE, GP and GDE) were modified to account for glycogen chains as substrate. The overall behavior is respected, as compared with the ODE model and the comparable model. However, a few differences are revealed, along with interesting observations labeled with numbers from 1 to 6 in Fig 4.15.

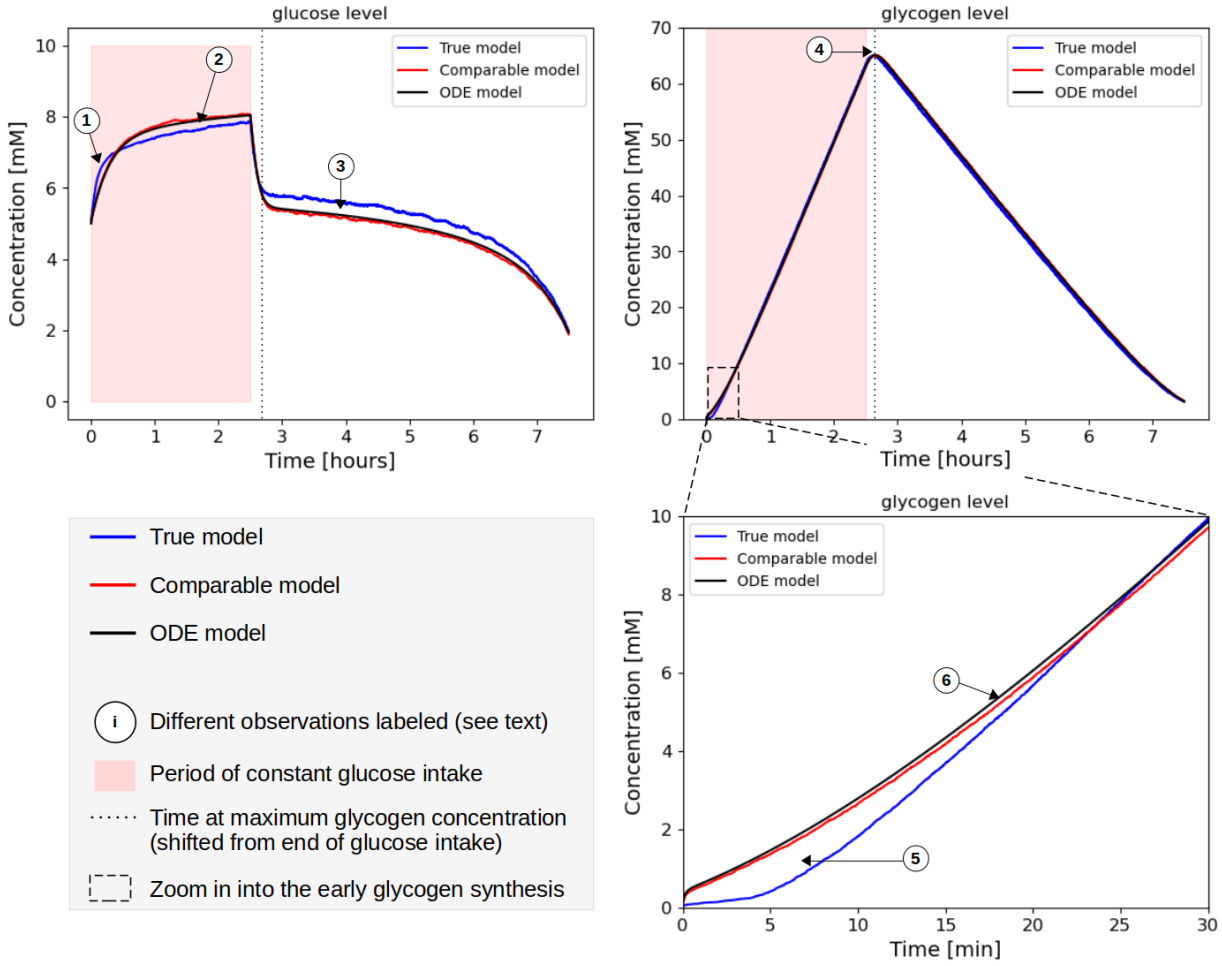

Figure 4.15: **Evolution of glucose and glycogen content during a short glucose intake period for different models. Top-left:** Evolution of glucose. The blood glucose increases until the glucose intake stops. It decrease drastically and stabilised using glycogen as a glucose source. **Top-right:** Evolution of glycogen. Glycogen fixes glucose during the glucose intake period. When the intake stops, glycogen start being degraded to provide glucose to the blood. **Bottom-right:** Zoom in the first 30 *min* of glycogen synthesis.

Observation 1 and 5 points at the differences in glucose and glycogen level (before $\approx$ 20 min) between the pure ODE model (also with the comparable model) and the final model. This is due to the introduction of a the K^{chain} constant that behave the same way as a Michaelis-Menten constant. Since now the chains are substrate for the enzymes, the start of the synthesis of glycogen is slower (the initial glycogen granule contains only 2 chains). It is now the chains that are limiting the reaction and not UDP-glucose. This early slow synthesis forces the system

to accumulates glucose residues.

Observation 2: at the opposite of observation **1**, from 30 *min* to 150 *min* the final model exhibit lower glucose level. One possible explanation could come from a reduced residual degrading activity. While the true synthesis rate is the difference between the rate of synthesis and the one of degradation, introducing chains concentration in the rate-laws might have lower v_{GP} and v_{GP} more than V_{GS} in this early phase.

Observation 3: During glycogen degradation the level of glucose is higher with the final model ($\approx 8\%$). This come from the fact that the debranching reaction is now releasing directly a glucose instead of a glucose-1-phosphate as before. This "short-cut" in the chemical reaction pathway leads to higher glucose level.

Observation 4: A temporal shift is observed between the end of the glucose intake and the peak of the glycogen synthesis. This is due to the hormone-response. To see an effective change in the activities of GS and GP, the phophorilase state function γ as to be lowered, and therefore the system has to wait a decrease in the glucose level before the degradation becomes effective.

Observation 6: There is small discrepancy between the ODE model and the comparable modeling the glycogen content. This is due to the Periodic-Coupling algorithm. During synthesis, UDP-glucose is very low and is substrate for the stochastic reaction catalyzed by GS. Even for consequent volume V_0, it appears that from time to time the integer value of [UDP-glucose]$\cdot V_0 \cdot \mathcal{N}_a$ drops below 1. In this situation, the stochastic algorithm can not proceed any GS reaction and will go to the next step. It introduces a small bias in the probabilities which reduces the difference between the flux of synthesis and the one of the degradation, leading to slower effective synthesis.

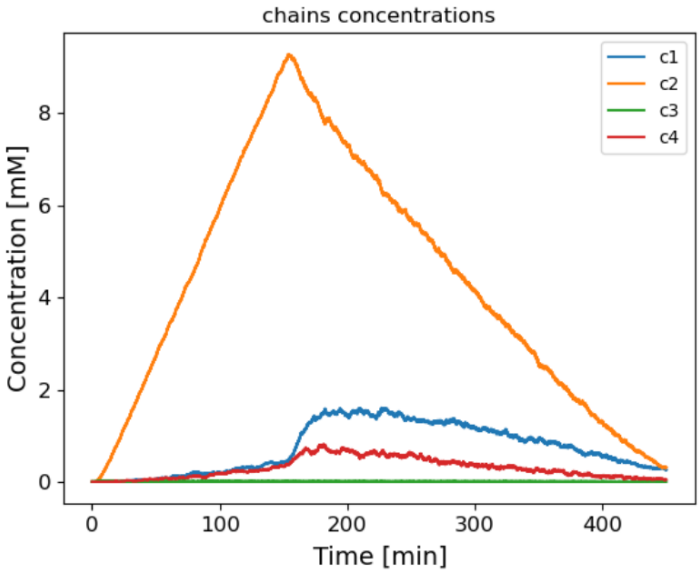

Figure 4.16: **Evolution of the substrate chains concentrations.**

Fig.4.16 shows the evolution of substrate chain type concentration for the final model, with the same initial condition as above. During the glucose intake period, most of the chains are of type $i = 2$ (intermediate length). Few type 1 and 4 are also forming. It is because as the synthesis occurs the number of branches increases. The newly formed branches are often of the size of $L_{\text{transferred}}$ since this simulation is obtained with high branching activity is high ($\Gamma = 0.6$). Since we have $L^{GDE} = L_{\text{transferred}} = 3$, the newly formed chain becomes directly substrate for debranching (status $i = 4$). Since $L_{\text{leftover}} = 3$ as well, only GS can react on the non-reducing end of the leftover residues, creating substrate of type $i = 1$. During the synthesis $c_1 = c_4$, precisely because $L_{\text{transferred}} = L_{\text{leftover}}$. Finally, during the synthesis phase, almost no substrate for branching is observed. It is because the branching reaction has the higher rate du to $\Gamma = 0.6$. Whenever a chain of type 3 is form by elongation, it react almost instantly. When the glucose intake stops, degradation starts and so the chain concentrations behaviour. Now $c_1 \neq c_4$, because debranching act on c_4 only, and c_1 chains have to wait that GDE removes all the daugther chains, so that c_1 becomes c_4 (most likely by transiting through type 2). No chain 4 are observed neither during degradation. There was almost none during synthesis, it is even more unlikely to see some during degradation.

Finally, the coupling of our glycogen structural model to a kinetic model also allows us to visualize the 3D evolution of the granule during its different phases. A animation of glycogen synthesis can be found at the following link: https://github.com/yvanrousset/hybrid_glycogen_models/tree/main/animation .

Specific K^{chain} for the degradation

When determining suitable rate-laws for stochastic reactions, the incorporation of ${K_{GS}}^{\text{chain}}$ and ${K_{\text{deg}}}^{\text{chain}}$ was deemed necessary to account for substrate chains dependency of the synthesis and degradation. To simplify the analysis of results, it was decided to investigate the effect of varying ${K_{\text{deg}}}^{\text{chain}}$ solely for degradation reactions. Simulations were conducted using the same initial conditions as those in Figure 4.15 for four different values of ${K_{\text{deg}}}^{\text{chain}}$: ${K_{GS}}^{\text{chain}}$ (as previously), 1, 5, and 10. The results indicate a noticeable effect on the evolution of blood glucose and glycogen at the end of the simulations, which range from 300-450 mM. As the substrate for degradation (c_2, c_3, c_4) becomes depleted, it becomes evident that degradation efficiency decreases with an increase in ${K_{\text{deg}}}^{\text{chain}}$.

Upon examining the experimental measurements of glycogen degradation depicted in Figure 4.11, it appears that there is a plateau towards the end of the degradation process, despite the different simulation conditions. This observation could suggest that the efficacy of the enzymes responsible for degradation in binding to glycogen may diminish, implying a high value for ${K_{GS}}^{\text{chain}}$. It is plausible to speculate that this could serve as a protective mechanism against excessive degradation, which would ultimately reduce the efficiency of the subsequent synthesis phase.

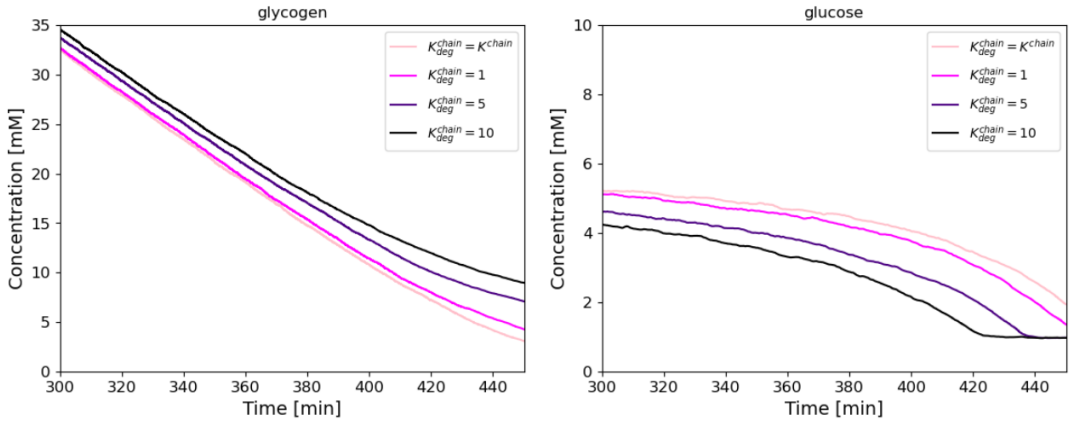

Figure 4.17: **Evolution of glycogen and glucose at the end of glycogen degradation for different values of $K_{\mathbf{deg}}{}^{\mathbf{chain}}$. Left:** Glycogen is almost fully degraded. Increasing $K_{\mathrm{deg}}{}^{\mathrm{chain}}$ diminish the efficiency of the degradation as substrate chains become rare. **Left:** Consequently glucose production is less effective leading to a lower blood glucose level.

4.4 Analysis of different glycogen related diseases

In this last part of this chapter, the final model is applied qualitatively to different cases of disease. There is no pretension here to explain the ins and outs of these diseases, but to provide a discussion, in the light of our model. For each diseases, different approaches will be provided. This section is intended to be exploratory and non exhaustive.

4.4.1 GSD type 0: impaired Glycogen Synthase

The enzyme involved in GSD0 is Glycogen Synthase (GS). 3 scenarios will be compared where the elongation reaction reaction rate is at 100%, 50% and 10% corresponding to what we define as healthy, mild and severe GSD0 models. It is done by introducing a constant α_{GSD} in the rate-law of Eq. 4.24:

$$v_{GS}{}^{\mathrm{GSD}} = \alpha_{GSD} \cdot \frac{c_1 + c_2 + c_3 + c_4}{c_1 + c_2 + c_3 + c_4 + K_{GS}{}^{\mathrm{chain}}} \cdot \left(\gamma \cdot v_p{}^{GS} + (1 - \gamma) \cdot v_{dp}{}^{GS} \right) \tag{4.24}$$

where α_{GSD} equal 1, 0.5 and 0.1. Remarkably, the outcomes attained from the mild model exhibit minimal deviation from the normal scenario, with almost identical levels of glycogen synthesis in both cases. While a slightly elevated glucose level was observed during the feeding phase, it remained within reasonable limits [Nuttall et al., 2008]. It might have been anticipated that the total glycogen synthesized would decrease by 50% given that $v_{GS}{}^{\mathrm{GSD}} = 0.5 \cdot v_{GS}{}^{\mathrm{new}}$. However, the hormonal response was dissimilar. The rise in blood glucose levels intensified the flux of synthesis, counteracting the effect of introducing $\alpha_{GSD} = 0.5$ However, even if this is

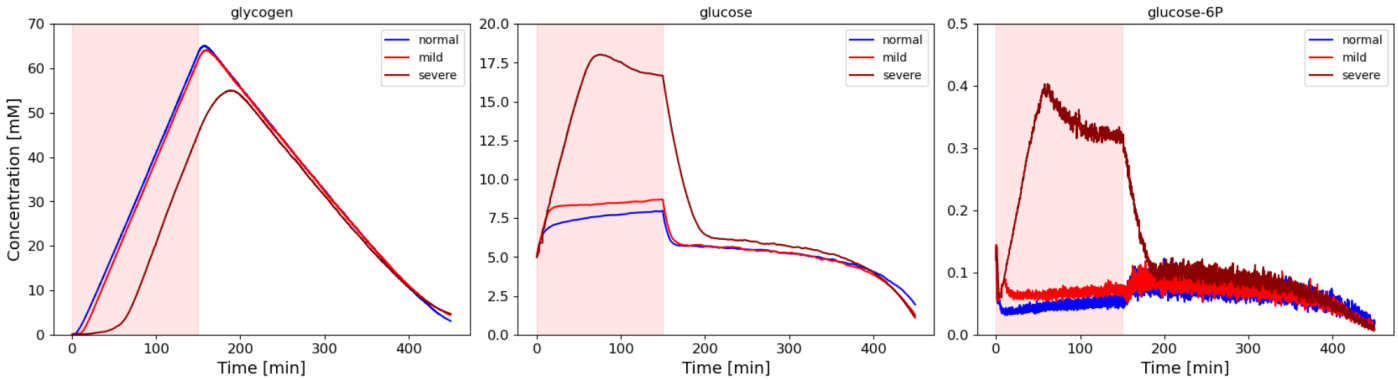

Figure 4.18: **Glycogen, glucose and glucose-6-phosphate levels for GSD0 models.** The pink area shows the duration of glucose intake. While the mild case exhibit the same behaviour as the normal case, the severe model shows a reduction of glycogen synthesis as well as abnormal glucose and glucose-6-phosphate level during feeding phases.

not modeled, the underlying physiological mechanism of our phosphorylation state function is that insulin is much more solicited in the present case to maintain glucose levels. Speculatively, this could potentially result in insulin resistance over time, as previously observed in patients [Tagliaferri et al., 2022]. In the severe model, distinct variations are observed from the normal and mild cases. At an early stage, glycogen synthesis is impaired, as indicated in Figure 4.18-left. Although there is a delay in its initiation, the rate of synthesis appears to be similar to that of the normal case once initiated. However, this shift in the initiation time results in a lower overall accumulation of glycogen. During the feeding phase, this also leads to a significantly higher blood glucose level, as depicted in Figure 4.18-middle. Nevertheless, no hypoglycemia were found while it is observed experimentally [Arko et al., 2020; Ozen, 2007b]. This is likely due to the relatively low outflow rates of glucose and fructose-6-phosphate from the system, which keeps the sum of all glucose entities (glc to glucose fixed into glycogen) almost constant. Consequently, the accumulated glucose in the blood serves as a glucose reservoir. This finding also explains the shift in the peak of glycogen synthesis between the severe and other cases, as described in Figure 4.18-left. The excess blood glucose is utilized for glycogen synthesis, which is subsequently degraded to maintain glucose homeostasis. An increase in the export of compounds from the system would likely lead to hypoglycemia, as glucose would be less likely to accumulate in the blood. To clarify why synthesis occurs at the same rate in the severe case despite differences from the normal and mild cases, it is necessary to examine the glucose-6-phosphate level. As previously stated, glucose-6-phosphate is an allosteric activator of GS (Eq. 4.15). During the feeding phase, glucose-6-phosphate accumulates (Figure 4.18-right), enhancing ${v_{GS}}^{\text{new}}$ and offsetting the effect of $\alpha_{GSD} = 0.1$. In contrast to the mild case, which only requires enhancement by γ to regulate synthesis, the severe case necessitates strong activation by glucose-6-phosphate.

4.4.2 GSD type 1: impaired Glycogen 6 Phosphatase

Two forms of Glycogen Storage Disease (GSD) type 1 have been identified, namely type 1a and type 1b. GSD type 1a is caused by a genetic mutation in the G6PC gene, resulting in impaired function of the glucose-6-phosphatase enzyme. This enzyme is responsible for converting glucose-6-phosphate into glucose. On the other hand, GSD type 1b results from a mutation in the G6PT1 gene, which encodes for the glucose-6-phosphate translocase 1 enzyme. This enzyme facilitates the transportation of glucose-6-phosphate from the cytosol to the endoplasmic reticulum, where it is transformed into glucose by glucose-6-phosphatase [Parikh and Ahlawat, 2023]. The current subsection will focus on GSD1a, which is also known as Cori disease. The investigation will examine three different scenarios in which the rate of G6PC is 100%, 50%, and 10% (normal, mild and sever respectively), by introducing the parameter α_{GSD1} into the rate-law of G6PC:

$$v_{G6PC} = \alpha_{GSD1} \cdot v_{\max} \cdot \frac{[\text{G6P}]}{K_M + [\text{G6P}]} \tag{4.25}$$

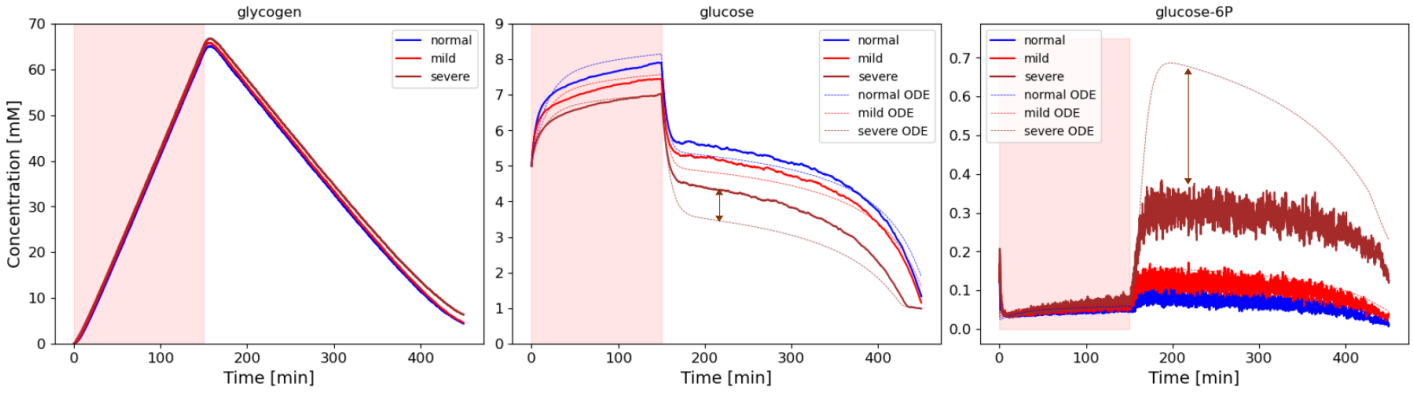

Figure 4.19: **Glycogen, glucose and glucose-6-phosphate levels for GSD1 models.** The pink shaded region represents the duration of glucose intake. Left: glycogen does not differ significantly across the three models. Middle: the glucose level is lower in the disease models, as the parameter α_{GSD1} decreases. The results of the pure ODE system are represented by the dotted line. The discrepancy between the stochastic and ODE models (represented by the brown arrow) is due to the release of one glucose molecule during the debranching reaction, emphasizing the importance of using a stochastic approach for this scenario. Right: When the degradation occurs, in the severe case, glucose-6-phosphate can not be efficiently transformed into glucose, resulting in a higher level of glucose-6-phosphate. However, the debranching process facilitates the reduction of the glucose-6-phosphate level, which would have been higher in the ODE model (represented by the brown arrow), where the debranching reaction is not included.

A simulation was performed to mimic a 150-minutes period of glucose intake followed by a 300-minutes period without intake in three scenarios. The changes in glycogen, glucose, and glucose-6-phosphate were analyzed and presented in Figure 4.19. No significant difference in the evolution of glycogen was observed among the models. However, patients with glycogen storage disease type 1a (GSD1a) typically exhibit glycogen accumulation [Kishnani et al., 2014; Ozen, 2007b], which was not observed in these simulations. Nevertheless, the last phase of glycogen degradation seems slower in the case of the disease, and repetitive glucose intakes may lead to excess glycogen content compared to the normal case.

Interestingly, the reduced activity of glucose-6-phosphatase (G6PC) led to lower glucose levels, as shown in Figure 4.19-middle, which is directly attributed to the enzyme's inability to convert glucose-6-phosphate to glucose. This was confirmed by the increase in glucose-6-phosphate during degradation, especially in the severe case, as observed in Figure 4.19-right. The duration of glucose levels in the range of 3-9 mM and the overall lower level in the disease case can be considered hypoglycemia, which has also been observed experimentally [Kishnani et al., 2014; Ozen, 2007b].

However, the differences observed between the models were less significant than if the system had been modeled with a pure ODE system. The incorporation of the glycogen structural model allowed for proper debranching reactions that released glucose directly, bypassing the regular pathway from glucose-1-phosphate to glucose, as in the reduction by GP. This resulted in higher glucose levels due to the debranching activity, which was also observed in Figure 4.19-right, where the glucose-6-phosphate levels were much lower than predicted by the ODE model in the severe case.

An increase in the export of fructose-6-phosphate into the glycolysis pathway was also observed in the GSD1 model. It could be speculated that an increase in lactate production, which has been observed in GSD1 [Kishnani et al., 2014; Oster et al., 2016], may result from this alteration.

Enhanced enzymatic activities

In order to expand the investigation, the activities of the four enzymes involved in glycogen synthesis and degradation (GS, GBE, GP, and GDE) were examined under the same simulation conditions as described previously. The time-dependent behavior of these activities is illustrated in Figure 4.20-a and 4.20-b. Notably, the reaction activities are expressed in terms of reactions per minute, with a conversion to $\mathrm{mM.min^{-1}}$ possibly achieved by dividing by $\mathcal{N}_a \cdot V$. For each enzyme, the number of single reactions was counted and averaged over one minute.

Figure 4.20-a indicates that reducing the efficiency of G6PC caused an increase in all four reaction activities. Given that glycogen concentration was unchanged across the three scenarios (Figure 4.19-left), it may be inferred that these increases were of the same proportion for all enzymes. Otherwise, a change in glycogen synthesis or degradation would have been observed. Figure 4.20-b shows the individual activities of the enzymes for each scenario.

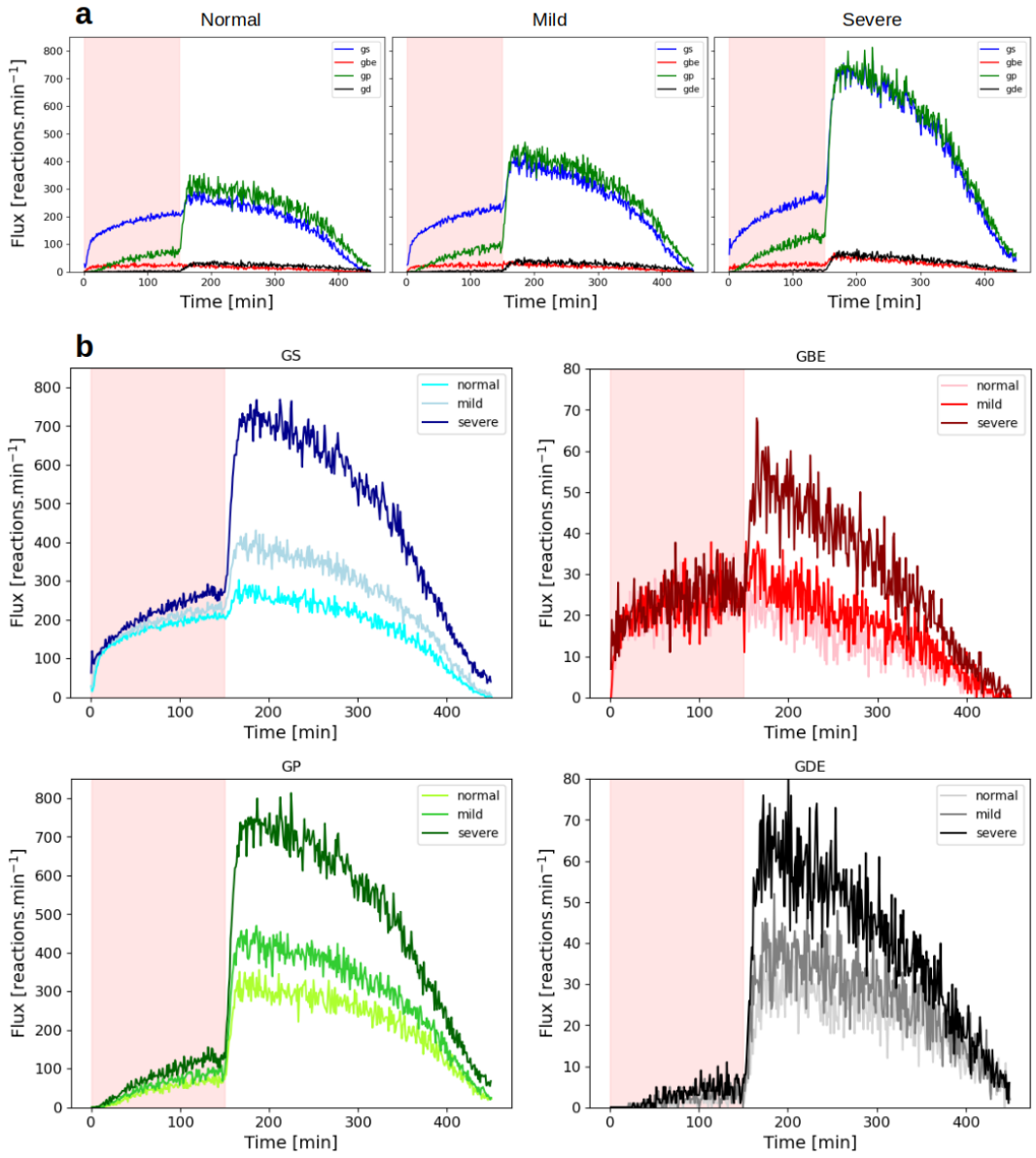

Figure 4.20: **Activities of GS, GBE, GP and GDE in glycogen storage disease 1.** The pink shaded area represents the period of glucose intake. The enzyme activities are expressed in units of reactions per minute. Panel **(a)** displays the activities of four enzymes (GS, GBE, GP, GDE) under normal, mild, and severe conditions (α_{GSD1}=1.0, 0.5, and 0.1, respectively). Decreased activity of G6PC results in increased activities of these four enzymes. Panel **(b)** illustrates individual enzyme activities in the three scenarios. Although no significant differences were observed during synthesis, the degradation phase revealed significant variations in enzyme activities.

While no significant differences were observed during the synthesis phase, the degradation phase revealed significant variations in enzyme activities. During synthesis, glucose is primarily transformed into glucose-6-phosphate, rendering G6PC less active. Thus, there were no major differences in activities observed. However, when degradation started, the effects of impaired G6PC became apparent as the reaction became central to glycogen breakdown.

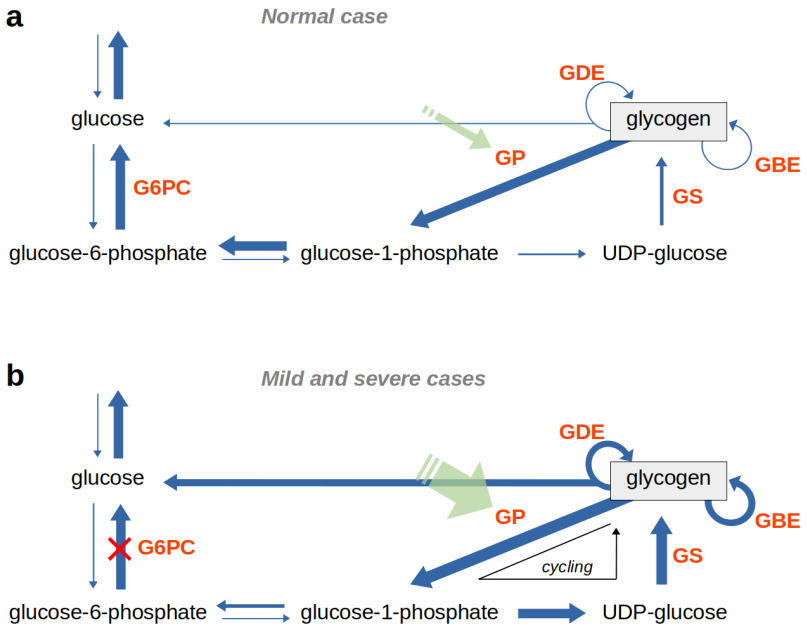

Figure 4.21: **Enzyme enhancements in glycogen storage disease 1.**

The justification for observing higher activity levels in all enzymes in the case of the disease is depicted in Figure 4.21. This schematic outlines the main reactions involved during glycogen breakdown in the normal case and with G6PC impairment (mild, severe). The arrows are intended to give a qualitative idea of the fluxes. In cases of reduced G6PC activity, glucose levels in the blood decrease, leading to a higher response by hormones that increase GP activities. Subsequently, the reduction in chains generates more substrates for the debranching reaction, thereby enhancing GDE activity. Since glucose-1-phosphate produced by GP cannot efficiently follow the normal pathway to form glucose, it accumulates, creating more substrate for elongation and enhancing GS activity (through it conversion to UDP-glucose). Ultimately, this also increases the formation of substrate for branching. Glucose equivalents are cycling between glucose fixed to glycogen, glucose-1-phosphate, and UDP-glucose, allowing for a tight interplay between all enzymes and forming substrate for debranching (Figure 4.21-b). This prevents the accumulation of glucose equivalents. Therefore, the rate of degradation remains unchanged from the normal case (Figure 4.19-left), but the contribution of debranching becomes increasingly important as α_{GSD1} decreases.

In the end, the cyclical process of glucose equivalent is expected to result in an elevation of energy consumption due to the fact that the conversion of glucose-1-phosphate to UDP-glucose requires UTP and releases two phosphate groups.

Glycogen structure in GSD1

With regards to the structure of the glycogen molecule, there were no discernible differences observed between the three scenarios examined with respect to the A to B ratio, Branching degree, and Average chain length (Figure 4.22). However, a small effect on the A:B ratio may be observed in the case of the disease, but this would require additional simulations to be statistically appreciated. However this would not be surprising since GDE is more active in the case of the disease model, and is the only enzyme that specifically work on A chains, resulting in a net dicrease of the A:B ratio. The effect would be further dilluted by the effect of GBE (see discussion "Probalistic approach to the A:B ratio" in the article from Chapter 3).

This observation would not be unexpected given that GDE is more active in the disease model and is the sole enzyme that specifically targets A chains, ultimately leading to a net reduction in the A:B ratio. The impact of GBE would reduces this effect since when $L_{\text{leftover}} = L_{\text{transferred}}$ it GBE act toward and A:B ratio of 1 (see discussion on "Probabilistic approach to the A:B ratio" in Chapter 3).

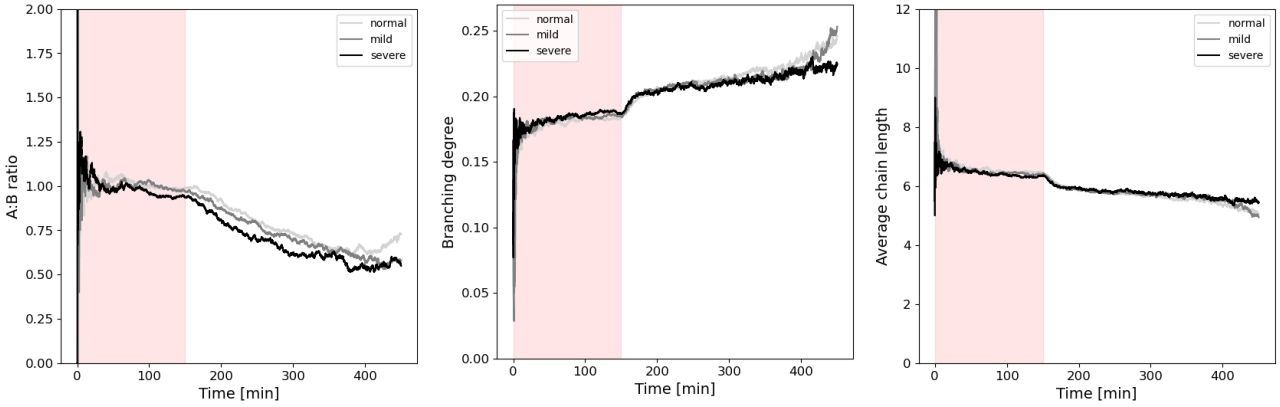

Figure 4.22: **Structural features of glycogen in the GSD1 models.** The A:B ratio (left panel) seems to exhibit small variation when reducing G6PC activity. The branching degree (middle panel) and the average chain length do not exhibit significant differences.

4.4.3 GSD type 2 and 3: impaired Glycogen Debranching Enzyme

In this section, we will discuss how our model can be applied to analyze GSD type 2 (Pompe disease) and GSD type 3 (Cori disease) in the context of reduced activity and impaired hydrolysis of the branching enzyme. Although these diseases have different genes involved, they both code for a branching enzyme. For simplicity, we will combine the discussion of both diseases since our model does not distinguish between the lysosome and cytosol locations.

Reduced activity

To analyze the effect of reduced activity of the branching enzyme, we first tested the case where we simply decreased the activity of GDE. This was achieved by introducing $\alpha_{\mathrm{GSD2,3}}$ in the determination of $\Lambda_{\mathrm{GSD2,3}}$ (see eq. 4.21). It now reads:

$$\Lambda_{\mathrm{GSD2,3}} = \frac{c_2 + c_3 + \alpha_{\mathrm{GSD2,3}} \cdot (1 - c_4)}{c_2 + c_3 + c_4} \tag{4.26}$$

When $\alpha_{\mathrm{GSD2,3}} = 1$ we recovered the same behaviour as before where the probability $\Lambda_{\mathrm{GSD2,3}}$ is just the ratio between substrate for GP to substrate for degradation (GP + GDE). However when $\alpha_{\mathrm{GSD2,3}} = 0$, the probability $\Lambda_{\mathrm{GSD2,3}}$ becomes 1, and the model can only reduced chains without debranching. We tested 3 different scenarios where $\alpha_{\mathrm{GSD2,3}}$ was set to 1, 0.5, and 0.1. We found no significant differences in the rate of glycogen synthesis and degradation or in the concentrations of other metabolites.

Impaired hydrolysis

The debranching reaction catalyzed by GDE involves two enzymatic processes. The first is a transferase activity where GDE cleaves the glycogen chain up to one glucose and transfers it to another chain. The second process is the hydrolysis of the remaining $\alpha - 1,6$ glucosidic bond releasing the glucose unit. In this paragraph, we focus on modeling the impairment of the hydrolysis activity, which would be the case in a genetic mutation affecting only the corresponding catalytic site of the enzyme.

In such a scenario, a chain would be left with a single glucose residue linked in $\alpha - 1,6$. This remaining glucose blocks the structure from further degradation since GP can only reduce a chain that is longer than DP4. To model this, we introduced a probability p_{fail} that during a debranching reaction, the hydrolysis part fails. We simulated the cases where p_{fail} was set to 0.1 and 0.5. Figure 4.23 shows the concentration of glycogen and glucose in these two scenarios, as well as in the normal case.

The granule reaches a blocked state (known as limit dextrin in *in vitro* experiment), in both cases of hydrolysis impairment (Fig. 4.23-top-left), where all the A chains have been missed-unbranched and the remaining structure consists of only B chains, and A chains in the form of a single glucose residue that cannot enter any other reactions. At this point, degradation stops, and the granule is frozen, resulting in an inability to produce glucose. This frozen state leads

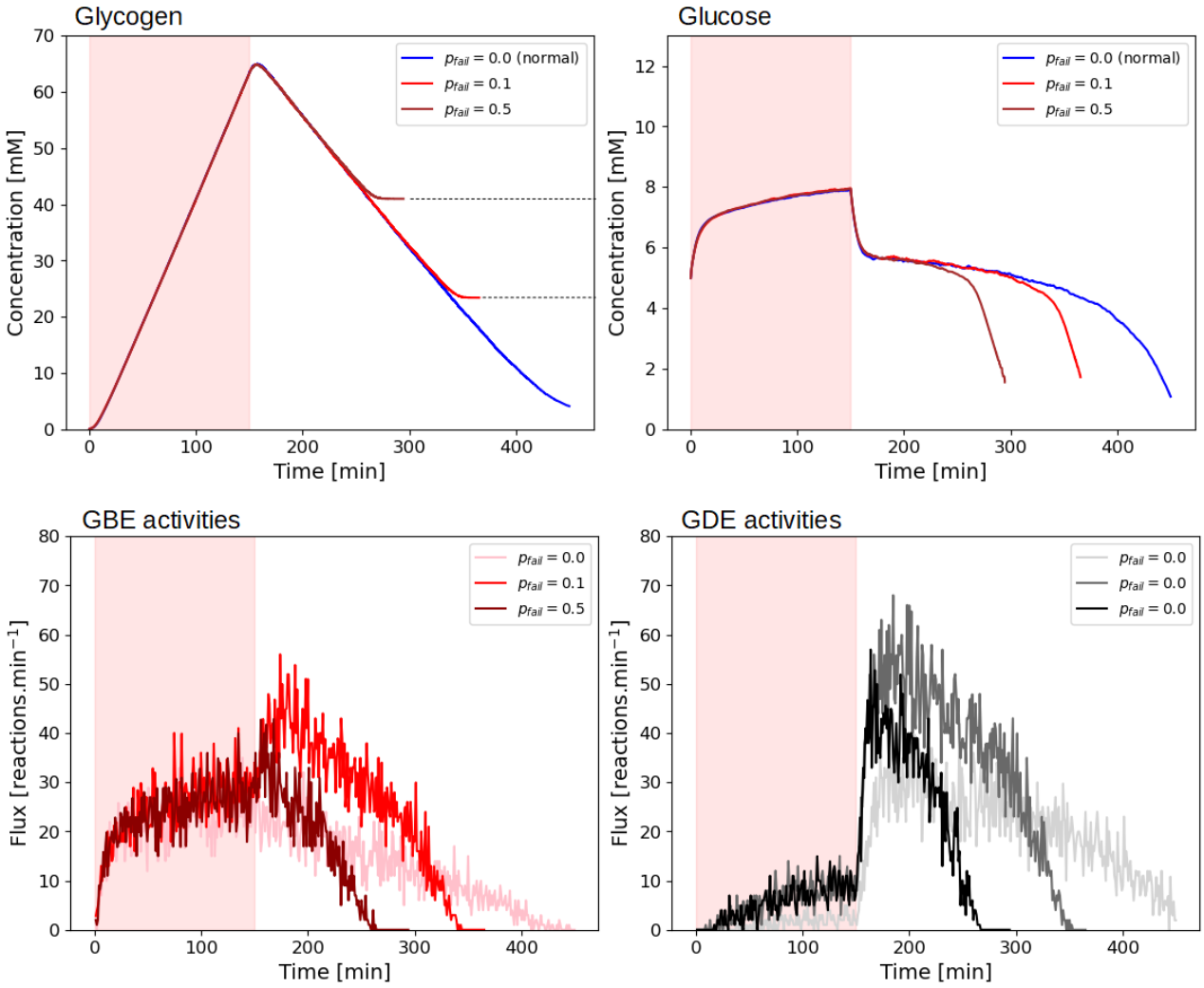

Figure 4.23: **Concentration of glycogen and glucose with branching and debranching fluxes with GDE hydrolysis activity impaired.**

to a rapid decrease in blood glucose when reached (Fig. 4.23-top-right), which is consistent with the observed glucose accumulation in GSD type 2 and 3.

Interestingly, both the branching and debranching activities appear to increase with the introduction of this hydrolysis failure. The reason for this phenomenon is not entirely clear and requires further investigation. However, one possible explanation is that the impairment initially decreases the substrate for glycogen phosphorylase by blocking some chains. To compensate, the system attempts to unbranch even more. This unbranching, even when hydrolysis fails, elongates some chains and provides more substrate for branching, which increases GBE activity.

To further observe glycogen accumulation in this model, we simulate three periods of 50 minutes of glucose influx, each followed by 100min of without. The results are presented in.

This small naive model without hydrolysis failure does not allows observation of glycogen accumulation

4.4.4 GSD type 4 and ADPB: impaired Glycogen Branching Enzyme

Reduced activities

In an initial attempt to replicate GSD type 4, changes in Γ were introduced. The values of Γ used for comparison were 0.6 (normal case), which was derived from the best fit as described in Chapter 3, as well as 5.0, 10.0, and 30.0. As previously done for other GSD, a simulation was conducted for a period of 450 minutes, with glucose intake occurring during the first 150 minutes. The resulting concentrations for glycogen, glucose and glucose-6-phosphate are illustrated in Figure 4.24.

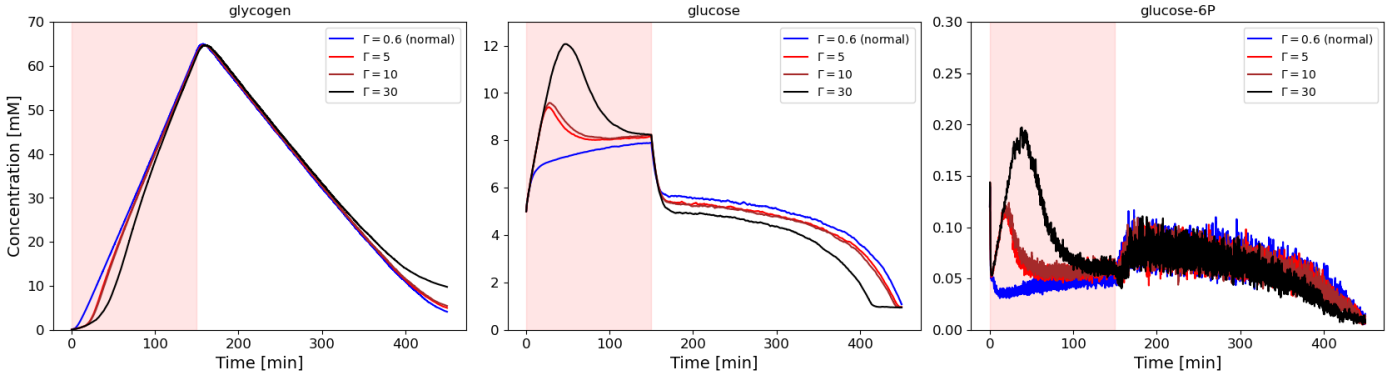

Figure 4.24: **Glycogen, glucose, and glucose-6-phosphate levels in GSD4 model. Left:** Glycogen early synthesis appears slower when Γ increases, as GS has less substrate due to GBE low activity.. **Middle:** Early glucose levels rise when Γ increases, as glycogen synthesis is slowed down. **Right:** Early glucose-6-phosphate levels increase, enhancing the synthesis of glycogen.

There is a noticeable decrease in the rate of initiation of glycogen synthesis as Γ increases, particularly when $\Gamma \geq 5$. This is due to the fact that the reduced branching caused by lower values of Γ leads to a deceleration in the interplay between glycogen synthase (GS) and glycogen branching enzyme (GBE), which in subsequently reduces the formation of new glycogen chains. This trend is further observed towards the end of the simulation period, where degradation becomes more challenging because of the reduced number of chains within the glycogen granule, leading to slower activity of glycogen phosphorylase (GP) and glycogen debranching enzyme (GDE). However, there is a point at which the number of glycogen chains

is sufficient to support proper synthesis and degradation. This is illustrated in the black curves (corresponding to $\Gamma = 30$) which eventually catch up to the synthesis rate observed in the normal case. Due to the aforementioned lack of chains in the glycogen granule and the subsequent impaired synthesis results in the early accumulation of glucose and glucose-6-phosphate (Figure 4.24-middle and -right), which quickly vanish when chains number become sufficient.

We also looked at some structural features during these simulations. The results show that

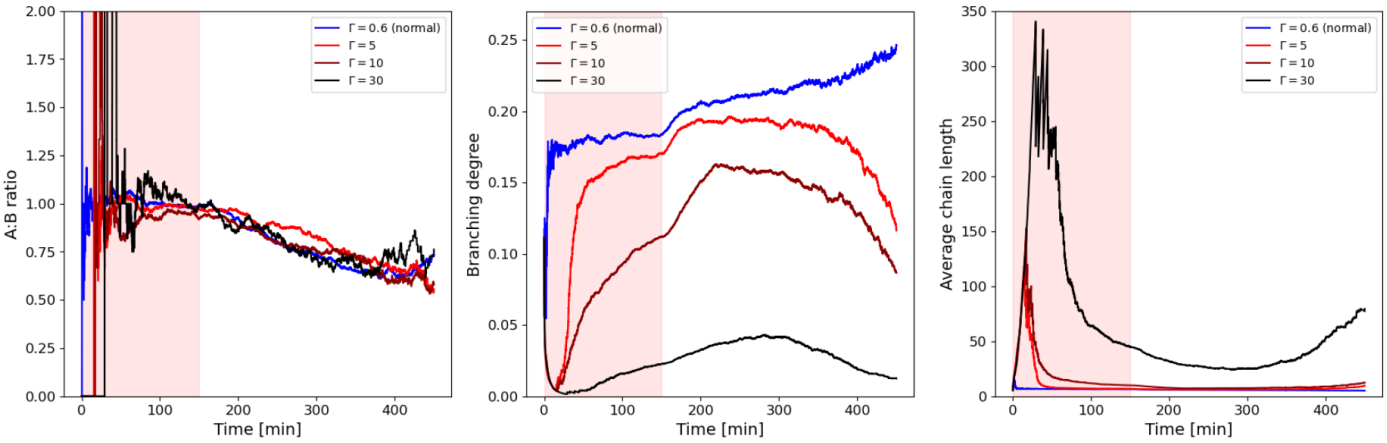

Figure 4.25: **Structural features of glycogen in the naive GSD4 models ($\Gamma = 0.6, 5, 10, 30$).** The A:B ratio (left panel) does show significant differences in all four scenarios. The branching degree (middle panel) and the average chain length (right panel) however, differs consequently.

the A to B ratio remains unchanged in all three cases, but there are significant differences in branching degree and average chain length (Fig. 4.25). As expected, when Γ increases, the elongation process is much more efficient than branching, resulting in longer chains and lower branching degrees. During degradation, the average chain length decreases as GP reduces the chains in all scenarios. However, after 300 minutes, the chain length seems to increase again, leading to a decrease in the branching degree.

One possible interpretation is due to the fact that the initial average chain length is higher at the beginning of the synthesis phase (around 45 minutes) than at the end of the synthesis phase (150 minutes), indicating that the first generation of chains might be longer. These chains can still grow through their non-reducing end but are prevented from reduction until the debranching reaction removes their $\alpha - 1,6$ bonds, forming the granule backbone. The increases observed at the end of the degradation (300-450 minutes) in Fig. 4.25-right are likely the contribution of these first-generation branches to the average chain length, as smaller ones have been removed.

The branching degree found here and the average chain length is more something one can observe in amylopectin.

Double helices formation

Long linear chains have a tendency to form double helices with neighboring chains, as observed in starch [Hejazi et al., 2008]. It is important to consider this mechanism, especially given that individuals with GSD4 exhibit insoluble glycogen [Sullivan et al., 2019]. Indeed, the formation of double helices can potentially lead to a local crystalline structure, making glycogen less soluble. To model this, we introduced the possibility for our chain description to become frozen when it becomes too long. The probability distribution for this frozen state follows logistic growth as follows:

$$\mathcal{P}(DP) = \frac{1}{1 + \exp^{-(DP-20)}} \tag{4.27}$$

Thus, the probability of a chain becoming frozen increases with the degree of polymerization (DP) of the chain and becomes greater than 0.5 for DP greater than 20.

When a chain is frozen, it prevents all enzymes from using it as a substrate. The introduction of this freezing state considerably changes the previous results for glycogen synthesis (Figure 4.26). The glycogen synthesis is significantly more impaired with the introduction of this frozen state, and therefore the level of glucose increases considerably compared to the approach without it. In both scenarios, the granule eventually reaches a completely frozen state (all chains are blocked), preventing further reactions from occurring. However, counterintuitively, the simulation with the higher Γ exhibits a higher glycogen content. This is potentially a bias in the timing of updating the status of the chains (frozen or not frozen), which depends on the activities of the reaction, and would require further investigation.

4.4.5 Outlooks

One recurring observation when applying the reduced model of glycogen metabolism to various genetic conditions is that the process is excessively regulated by the hormones-response function, denoted by γ. When defects are introduced in the enzymatic process that would intuitively reduce glycogen synthesis or degradation, this regulatory mechanism compensates. This can be partially explained by the fact that glycogen is the sole source of glucose production or utilization in the model. As a result, blood glucose levels, and thus glucose intake and output, indirectly regulate the rate of glycogen synthesis and degradation, instead of the other way around. This possibility is not necessarily incorrect, and it echoes a question posed in a previous work [Palm et al., 2012] on whether glycogen synthase regulates glycogen synthesis or if it's the glucose intake that does so.

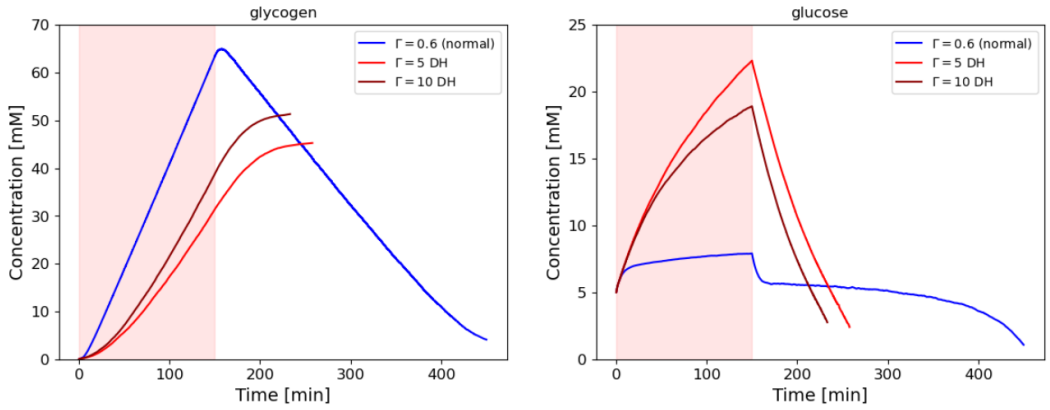

Figure 4.26: **Glycogen and glucose concentration in a GSD4 model with and without double helix formation.** The blue lines represent previous results obtained under normal conditions. The red line shows the values obtained when $\Gamma = 5$, and the dark red line shows the concentration when $\Gamma = 10$.

In the following section, we will briefly examine mechanisms that could alter glycogen synthesis and degradation, potentially weakening the significance of hormonal regulation.

Insulin resistance

Insulin resistance (IR) is a condition where cells become less sensitive to insulin, resulting in decreased activation of enzymes involved in glycogen synthesis. This leads to lower glycogen synthesis and higher blood glucose levels. To model insulin resistance in our study, we revised the γ function (Eq. 4.1) as follows:

$$\gamma_{\mathrm{IR}} = \max\left(\frac{1}{1+\exp\left(\mathrm{Glc}-\mathrm{Glc}_0\right)}\ ,\ 0.5\right). \tag{4.28}$$

For high glucose levels, γ_{IR} cannot decrease below 0.5, which impairs the activation of the enzymes responsible for glycogen synthesis. We conducted simulations using this function for both normal and GSD0 models (Fig. 4.27).

Figure 4.27-left column show the changes in glycogen, glucose, and glucose-6-phosphate concentrations during the synthesis and degradation phases in the non-GSD0 model ($\alpha_{\mathrm{GSD0}} = 1$). The pink line represents the results under insulin resistance, and the black line represents the normal case without insulin resistance. As seen, the rates of synthesis and degradation have already changed, resulting in hyperglycemia and an increase in glucose-6-phosphate levels.

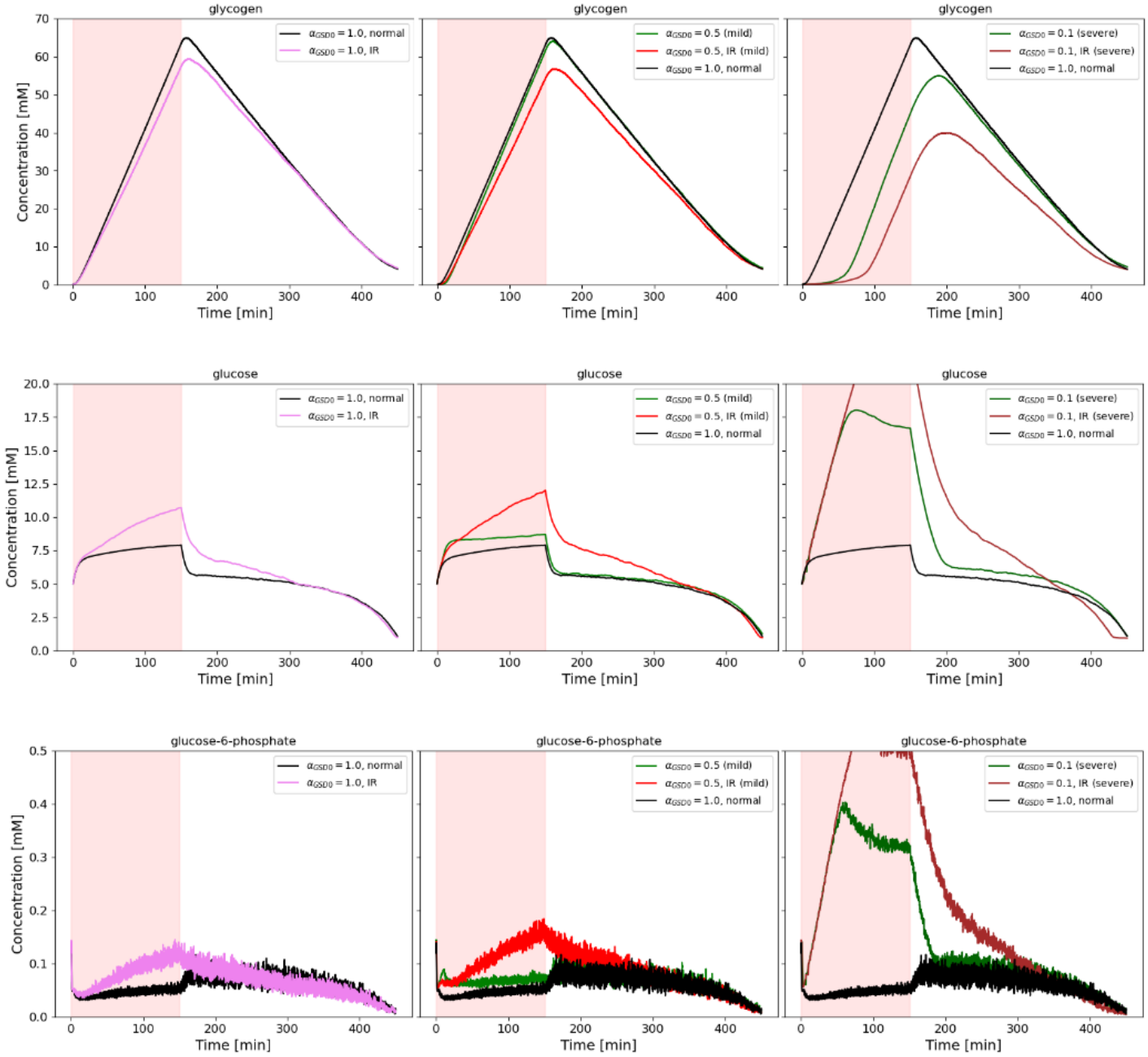

Figure 4.27: **Glycogen, glucose and glucose-6-phosphate levels for GSD0 models with and without insulin resistance.** The pink, red and dark red curves are GSD0 model with insulin resistance (IR). The black geen and dark green show the same model without insulin resitance. **Left to right:** increasing the severity of the disease. **Top to bottom:** Glycogen, glucose and glucose-6-phosphate levels.

Insulin resistance in the GSD0 model ($\alpha_{\text{GSD0}} = 0.5$ and 0.1) had an even more pronounced effect than that observed in subsection 4.4.1. The red and dark red lines show the effects of insulin resistance, while the green and dark green lines represent the normal case (Fig 4.27-midle and -right columns). In this scenario, hormones are unable to compensate for the impairment of glycogen synthase activity.

Toward addition of glycolysis

To accurately assess the true impact of glycogen breakdown on glucose production, it's essential to take into account alternative sources of glucose. To achieve this, it's necessary to augment the current model with reactions that encompass a reversible glycolysis pathway. In pathological situations, incorporating these reactions will prevent the system from solely relying on glycogen synthesis and breakdown to maintain glucose levels, and instead prioritize other pathways.

Chapter 5

Mini-chapter: About β and α glycogen granule

5.1 Opinion: The fractal view of glycogen

In the introduction, it was mentioned that models of glycogen structure were originally developed to explain the observed structure resulting from glycogen digestion in enzyme cocktail experiments, with particular emphasis on the A:B ratio. Whelan proposed a suitable structure to explain these observations as well as other enzymatic observations. However, Whelan was cautious when introducing the model [Gunja-Smith et al., 1970b], stating that:

> " In explanation of the combined observations we offer the structure shown in fig. 4. This is intended only to express certain concepts and is not to be regarded as precisely defining glycogen structure. In this model half the B chains carry an average of 2 A chains, while the other B chains each carry 2 B chains. The model has been drawn in its most symmetrical form but we would envisage variations from this symmetry in the actual polymer, both in regard to chain length and average degree of substitution of the chains."

and that the actual structure is more complex and less regular than the model suggests, anticipating potential misinterpretations.

Despite this caution, the model gained popularity and was taken as exact in its pure symmetric form by Goldsmith in 1981 [Goldsmith et al., 1982], who introduced the notion of regular tiers. It was supported apparently by [Thornell, 1974] observation of glycogen particle being less dense at the center. Later, Melendez explored the idea of glycogen potentially being the first known real biological fractal structure. While the paper introduced clever optimality concepts, it had the effect of further popularizing the fractal view of glycogen.

However, numerous articles have since presented arguments or results that contradict this fractal view. Despite this, the belief in the fractal view of glycogen persists and continues to be spread in many glycogen-related papers today. While the word "fractal" has diminished in usage, many papers still use the term "tiers". This word is just a derived way of describing glycogen as a fractal structure. In the article by Goldsmith, the fractal pattern leads to different

layers of density (tiers), which finally reach the last tier (12th) with a limit density where the space is completely filled.

The limit density and corresponding last tier have also provided support for the idea that glycogen has a limited size and cannot grow indefinitely, as this mechanism prevents enzymes from further synthesizing the granule due to the steric hindrance effect. This mechanism has been shown in some dendrimeric structures exhibiting limited growth do to steric hindrance [Fréchet, 2003].

The persistence of the fractal view within the glycogen community has the potential to overshadow other interesting mechanisms or research questions related to glycogen structure.

In the following, different arguments against this fractal view of glycogen are presented.

The hypothetical symmetry of the glycogen granule would be highly sensitive to even small variations in the position of its branches, which can lead to deviation from a fractal structure. As the granule grows, these defects are spread to the next chains, causing any regular pattern to disappear. It is important to note that this model is depicted in two dimensions, so small variations from a perfectly symmetrical structure may have even more impact in three dimensions.

The average radius of glycogen granules typically falls between 10 and 15 nm [Drochmans, 1962; Sullivan et al., 2014]. When a granule with a radius of 13 nm is filled with a fractal pattern, it would have 6.8 tiers and a total of 111 chains, containing 2800 glucose units, resulting in a molecular mass of $M = 5 \cdot 10^5$. This value is 20 times lower than what is reported for larger granules. Despite being a medium-sized granule, a glycogen granule with only 2800 glucose units, with half of them in the last tier, can be considered almost empty. Gunja-Smith pointed out that the hypothetical 13th tier would not be possible as it would exceed the density of closely packed single helices. The volume between R_{11} and R_{12} can be calculated as follows:

$$\delta V = V_{12} - V_{11} = \frac{4}{3}\pi({R_{12}}^3 - {R_{11}}^3)$$

and the number of residues in the layer between R_{11} and R_{12} is given by

$$\delta N = 2^{t-1}$$

When packed in helices, one residue occupies 0.318 cubic nanometers. Therefore, the volume occupancy in the 12th tier can be expressed as:

$$\rho^* = 0.318 \left(\frac{\delta N}{\delta V}\right) \Rightarrow \rho^* = 0.90 \ .$$

Performing the same calculation for a glycogen with a radius of 13 nm, the volume occupancy is found to be $\rho^* = 0.10$. This means that for a typical glycogen molecule, assuming this tiered model, the occupancy at the surface is only about 10% of the volume.

Another argument can be made by questioning the presence of the dense hypothetical outer layer from the fractal view. If it is assumed to exist, it raises the question of why not a single fluctuation would have resulted in chains that are slightly longer than the others, extending

beyond the layer, and becoming a seed for another synthesis process, similar to the initial chains initiated by glycogenin.

Besford et al., 2015 found that small angle X-ray scattering experiments on β glycogen are inconsistent with regular branching has proposed in the fractal model introduced by Goldsmith et al., 1982 and further extended, as the tiered model by Meléndez-Hevia et al., 1993. In a recent study, Kim and Duhamel, 2023 addressed the radial density of glycogen. They found that the density is higher at the center and decreases only at the outermost part of the granule. This finding invalidates the fractal view of glycogen.

It would be very welcome if research in glycogen refrains from using terms such as "tiered structure," "highly regular," or "fractal" when referring to the structure of glycogen. The use of such terms may create the false impression that glycogen degradation occurs layer by layer and that glycogen granules cannot exceed a certain size limit, typically a radius of 21 nm. This view may obscure potentially important scientific questions related to glycogen structure and metabolism.

With a non fractal glycogen, the question of what stops β glycogen granules to grow beyond observed size remains.

5.2 A toy model system for β and α interactions

In this section we present a simple model to discuss β and α granule interaction. We recall that beta granule are single glycogen granule that we have seen all along this thesis. α granule, on the opposite are aggregate of β granule and are found in the liver.

It is believed that α granules [Besford et al., 2015] are a way to regulate glycogen synthesis and degradation by reducing the surface to volume ratio. When β granules aggregate, the molecular weight of the aggregates is the same as the sum of all individual weights of the beta granules that constitute it. However the surface is not the sum of all individuals.

In what follows, we approximate α and β granules as spheres and that we get the surface of the aggregates from its conserved volumes. If two β granules, 1 and 2, aggregates into an α granule, one can write:

$$V_{\beta,1} + V_{\beta,2} = V_\alpha \tag{5.1}$$

where

$$S_\alpha = (4\pi)^{1/3} \cdot \big(3 \cdot (V_{\beta,1} + V_{\beta,2})\big)^{2/3}, \tag{5.2}$$

with:

$$S_\alpha < S_{\beta,1} + S_{\beta,2}. \tag{5.3}$$

In a naive approach we will assume that the rate of the synthesis and degradation follows a mass-action kinetic, and that the enzyme works at the glycogen surface. We further assume

that the synthesis and degradation processes are described by the rates of glycogen synthase and glycogen phosphorylase respectively.

Consider a system in its initial state with N_{glc} glucose molecules, N_β β granules, and $N_\alpha = 0$ α aggregates. The β granules in the initial state are composed of only 10 glucose units and can be viewed as the Glycogenin cores of a granule (refer to Chapter 2 for more information). As time passes, the β granules will react with another β granule to form α aggregates. Subsequently, these α aggregates can release a β granule.

With these assumptions, the corresponding reactions and system can be described as follows, where v represents the reaction rate:

$$v_{GS}(t) = k_{GS} \cdot \Big(\sum_{i=1}^{N_\beta(t)} S_i(t) + \sum_{i=1}^{N_\alpha(t)} S_i(t) \Big) \tag{5.4}$$

$$v_{GP}(t) = k_{GP} \cdot \Big(\sum_{i=1}^{N_\beta(t)} S_i(t) + \sum_{i=1}^{N_\alpha(t)} S_i(t) \Big) \tag{5.5}$$

$$v_{\beta\to\alpha}(t) = k_{\beta\to\alpha} \cdot N_\beta(t) \cdot \Big(N_\beta(t) + N_\alpha(t) \Big) \tag{5.6}$$

$$v_{\alpha\to\beta}(t) = k_{\alpha\to\beta} \cdot N_\alpha(t) \tag{5.7}$$

The elongation reaction catalyzed by GS occurs at the surface of granules, and its rate is proportional to the kinetic rate constant, k_{GS}, multiplied by the sum of all accessible surfaces of both beta and alpha granules. Reduction by GP follows the same kinetics. To form an alpha aggregate, a beta granule must react with another beta granule or directly with an already formed alpha aggregate. The release of a beta granule from a cluster is proportional to the number of clusters. The system is solved stochastically, and each granule has a unique identifier that allows for a detailed description of the compositions of different clusters.

How the surface affects the kinetics

In Figure 5.1, we observe the total glucose fixed into glycogen under different scenarios. When there is no aggregation (i.e., $k_{\beta\to\alpha} = 0$), an increase in the number of granules maximizes the surface area, resulting in faster synthesis, as shown in the top-left panel.

However, when $k_{\beta\to\alpha} > 0$, the system's total surface area decreases, which is equivalent to a decrease in the number of beta granules, leading to slower synthesis (as described by equation 5.3). The time courses are mixed due to variability in the number of clusters formed.

The average radius at the end of the simulation differs, as the same amount of glucose is incorporated into all scenarios, distributed among the different granules. A system with more granules will have smaller granules.

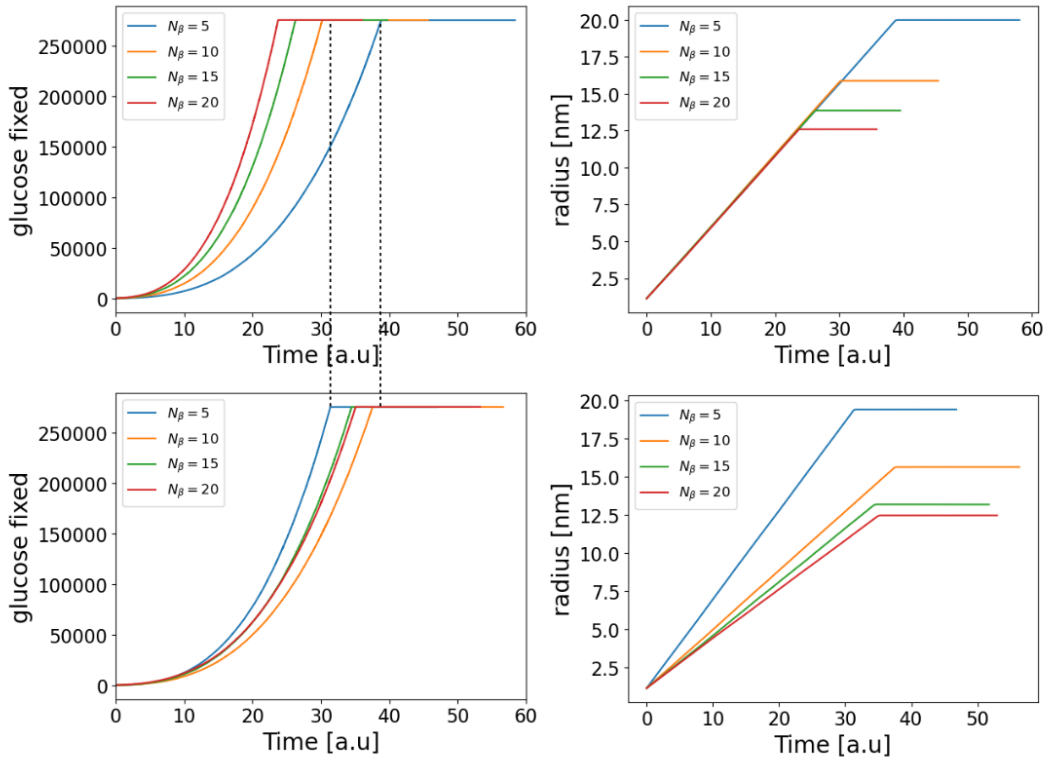

Figure 5.1: **Total glucose fixed and average radius of β granules with and without aggregation. Top panels:** without aggregation. **Bottom panels:** with aggregation. When α granules are formed, it reduces the rate of synthesis.

Interestingly, the granule radius increases linearly with time. If the rate of the reaction were to be constant, one could have expect the volume V to increase linearly, as each glucose molecules fixed on the system contribute to a small δv. Therefore one have:

$$\frac{dV}{dt} = \text{cst},$$

which gives:

$$r^2 \cdot \frac{dr}{dt} = \text{cst}$$

and thus,

$$\frac{dr}{dt} \propto r^{-2}. \tag{5.8}$$

But here the rate v_{GS} is proportional to the surface S, therefore one can write:

$$r^2 \frac{dr}{dt} \propto S \quad \Rightarrow \quad \frac{dr}{dt} \propto r^2 \cdot r^{-2}$$

and thus:

$$\Rightarrow \frac{dr}{dt} = \text{cst} \tag{5.9}$$

It could have been suggested that the limitation in observed glycogen could be due to equation (5.8), requiring a prolonged feeding phase to observe a significant increase in the radius profile distribution. However, with the rate-laws depicted above, which lead to a radius increasing proportionally with time, this is not a possible explanation.

In Chapter 4, section 4.2, the experimental data on glycogen synthesis in humans showed a linear behavior in time, which differs from the behavior described in the previous paragraph. The relation between surface and synthesis rate of a granule is observed on a single or small number of glycogen granules, as opposed to the larger number of β granules in a cell. The aggregation of a large number of β granules may reduce the slope of glycogen accumulation, resulting in a more linear experimental observation.

Further investigation of this system could lead to interesting results. For instance, it could be interesting to characterize the effect of glycogen synthesis and degradation in diabetic individuals who maintain less easily glycogen in the form of α aggregates [Besford et al., 2015; Sullivan et al., 2014]. Increasing $v_{\alpha\to\beta}(t)$ could represent this effect. It is also believed that hormones regulate α and β granule interaction, where low phosphorylation results in aggregation [Besford et al., 2015]. The model could potentially be incorporated into the metabolic model of glycogen introduced in Chapter 4, which contains the phosphorylation state function γ.

Chapter 6

Conclusion

6.1 Main Results

A three-dimensional realistic model of glycogen

Since the 1950s, experimental investigations have provided valuable insight into the structure of glycogen. However, few numerical and modeling approaches have been used to complement these experiments. While in vitro and in vivo experiments offer data on glycogen structure, interpretation can be challenging. In contrast, numerical procedures, despite being based on assumptions, can unravel this complexity and offer explanations, discussions, and predictions.

This work presents a detailed 3D model of glycogen structure. The geometry used to describe the chains is inspired by self-avoiding polymer models and is parametrized using experimental measurements of how glucose assembles in helical structures within glycogen linear chains. In this model, glucose units are represented as hard spheres that fill the 3D space, preventing newly formed chains from occupying the same position. We introduce rules for the enzymes involved in the synthesis and degradation of glycogen, with particular focus on the branching enzyme mechanism, which requires more complex rules due to its nature.

We also employ a stochastic algorithm to track single reaction events, enabling us to make corresponding structural changes. In a first qualitative study, we investigate the effect of the elongation to branching rate ratio (Γ) and observe how the granule varies depending on this ratio. This highlights the potential of such a model in tracking various structural properties of glycogen, such as chain length distribution, branching degree, A to B ratio, radius, and density.

Furthermore, we observe how sensitive the structure of glycogen is to the branching mechanism and, more precisely, to the chain lengths involved in the branching process. We show that the branching mechanism, along with the elongation to branching rate, are the main contributors to the shape of the chain length distribution in glycogen. One particularly striking result is the formation of multi-modal chain length distributions when the branching mechanism favors spacing between two consecutive sister chains, and more generally when the minimal lengths involved in the mechanism are long.

We later used experimentally measured chain length distributions as a signature of the branching mechanism. By testing various branching scenarios, we concluded that the branching process is likely more flexible than previously reported in the literature, in terms of substrate specificity constraints. This method can be used to characterize differences in glycogen branching enzyme among different organisms, as different chain length distributions are observed.

We used the parameters of the model that were found to reproduce the chain length distribution in mice data to predict other structural features of glycogen. We found that all other structural properties were in good agreement with experimental data, despite the model being parameterized only for the branching mechanism.

One of the properties we studied was the A to B ratio, which characterizes the proportion of chains without a daughter (A chains, only branched through their reducing end) to the chains that have at least one daughter. We provided a probabilistic explanation why this ratio tends to one when the branching reaction transfers, on average, the same chain length as it lets on the mother chain. When this symmetry is broken, different A:B ratios can be obtained, providing an elegant explanation for the differences in this ratio observed in starch.

In addition, our model also accurately predicts granule size and glycogen radial density profile. Our model is in good agreement with a recent study on glycogen density, which contradicts the fractal view of glycogen, in which the density exponentially increases with radius.

The interplay between glycogen structure and glycogen metabolism

The primary objective behind developing the glycogen structural model was to investigate the mutual relationship between glycogen structure and metabolism. Specifically, the aim was to understand how glycogen structure impacts glycogen metabolism, and vice versa. The 3D structural model involves only four stochastic reactions (GS, GBE, GP and GDE), which could allows a proper investigation of a small system composed by these four enzymes, glycogen and glucose equivalent. However such a model would become rapidly limited as it would not catch the complexity and potential regulatory mechanisms involved in glycogen metabolism. The ideal goal being to be able to integrate the our structural model into a realistic glycogen metabolic model.

As the number of reaction increases, a pure stochastic approache becomes rapidly computationally expensive, where solving ordinary differential equations are way more efficient at handling even intermediate size system. Therefore, we aimed to incorporate our stochastic 3D structural model into an ODE system. Some existing stochastic approaches (e.g Tau-leaping and Slow-scal stochastic simulation algorithm) have been introduced to reduce the computational time of the classical Gillespie algorithm, but are not applicable in our case, in their actual forms. The main reason is that, in our system, certain reactions must be modeled purely stochastically throughout the simulation, regardless of their rate.

Therefore, we developed our own in-house stochastic algorithm, called the Periodic-coupling algorithm, to incorporate our stochastic structural glycogen model into an ordinary differen-

tial system. This algorithm consists of two modules, one stochastic and one deterministic, which communicate and synchronize their respective changes. We showed that the algorithm converged when the communication time is small enough. The value of this communication time is determined empirically by comparing the result values obtained to the results in a pure ODE system, for equivalents system. We could have also approached the problem by modeling all reactions in a classical stochastic manner, but we show that our model surpasses a classical Gillespie approach, especially for large systems. However, we still need to investigate a theoretical approach to determine the maximum communication time that guarantees the convergence. With our newly developed algorithm, we were able to incorporate our stochastic structural glycogen model into an ordinary differential system.

Our reduced glycogen metabolism kinetic model draws significant inspiration from König et al's work. Although we retained several reactions and their corresponding rate-laws, we made numerous modifications to suit our needs. The resulting model comprises 13 reactions and 11 metabolites, with dynamic hormonal regulation based on blood glucose levels, as well as allosteric regulation of glucose-6-phosphate. The model can simulate periods of glucose intake and constant export. During the intake phase, we observed tightly controlled glycogen synthesis, mediated by hormones to maintain stable blood glucose levels. Similarly, when glucose intake ceases, hormonal functions fine-tune the rate of degradation to ensure glucose homeostasis. Furthermore, we investigated how the system behaves in the absence of hormonal regulation, as well as in cases of insulin or glucagon resistance. These scenarios result in either extremely high or low glucose levels, respectively.

Next, we integrated our reduced glycogen metabolism model with our stochastic structural glycogen model using our periodic-coupling algorithm. However, to achieve this, we needed to revisit some of the rate laws adapted from König's model to make them dependent on the glycogen chains. Each enzyme involved in glycogen synthesis and degradation has now a unique combination of substrate chain types. The resulting complete model allows us to explore the interplay between glycogen structure and metabolism in detail, as well as catching the effect of the debranching reaction on the kinetic model. One interesting observation is that the debranching reaction contributes to the blood glucose level ($\approx 8\%$).

We presented an exploratory discussion on glycogen storage diseases, outlining simple models to understand them. In GSD0, we noticed impairment in glycogen synthesis, resulting in high glucose levels during the synthesis phase, consistent with the observed physiological phenotypes in this disease. In GSD1, we examined the role of the debranching reaction in regulating blood glucose levels. As the disease severity increases, this contribution becomes more significant. Examining the activities of the four enzymes involved in glycogen breakdown and synthesis (GS, GBE, GP, and GDE), we found a marked increase in their activities during the degradation phase. This increase is necessary for maintaining glucose homeostasis, resulting in glucose cycling around glycogen, with glucose being fixed and released. We also observed that in GSD1, the A to B ratio is likely higher during degradation.

In GSD2 and GSD3, we simulated hydrolysis failure, leading to glycogen accumulation and an inability to maintain glucose homeostasis for an extended period. Surprisingly, we found

that GBE and GDE are more active in these cases. In GSD4, we observed a shift in the early phase of glycogen synthesis, with inefficient branching enzyme slowing down the process. As a result, blood glucose levels increase shortly after synthesis begins and can not be maintained for as long as in the normal case. Structurally, GSD4 appears to have the same A:B ratio as in the normal case, but with longer branches and lower branching degree, which is a logical consequence of the slower synthesis.

Our model is highly regulated by hormones and glucose-6-phosphate levels. During the fasting phase, glycogen is the sole source of glucose, and any deficiencies are compensated for through positive regulation to maintain glucose homeostasis. To address this, we can either reduce the impact of hormones on the rate-law or introduce another glucose source such as reverse glycolysis. In order to test this, we allowed for the import of fructose-6-phosphate in a simple manner. Our observations indicated that when glycogen is no longer the sole source of glucose, the impact of each disease on glycogen synthesis is much more significant. Finally, we examined a case of insulin resistance in GSD0, and found that glycogen synthesis was barely able to occur.

β and α granules

We discussed the concept of a fractal glycogen and its prevalence in literature despite experimental evidence against it. While Whelan's original model was later interpreted in a more symmetrical form and widely propagated, there is now evidence and arguments against this view. Nevertheless, remains of this idea persist in the literature and may misguide research on glycogen. We reviewed some arguments contradicting this view and provided some others.

The fractal hypothesis elegantly explains why glycogen beta granules should not grow indefinitely and instead reach a maximum density at a 21 nm radius, which inhibits enzyme elongation of glycogen chains. However, given that glycogen is likely not fractal and that density is probably uniform, we sought an alternative explanation for why larger beta granules aren't observed in nature. To this end, we developed a toy model in which beta granules aggregate into alpha granules, with synthesis rate proportional to glycogen surface area. Our model shows how aggregation can slow down synthesis by minimizing the surface-to-volume ratio. However, although glucose addition contribute less importantly to the radius for larger radii, the radius exhibit a linear dependency with time. A potential explanation in which one would require a prolonged period of glycogen synthesis to observe beta granules with radii greater than 30 nm is not observed it the model we proposed.

6.2 Future Work

Possible improvement on the glycogen structural model

Our current coarse-grained structural model for glycogen efficiently reproduces its structural properties, but it does not adequately consider the effects of steric hindrance on the dynamics of the reactions involved. We only account for steric hindrance by rejecting reactions that try to place glucose monomers in already occupied spaces, but this does not accurately reflect the volume exclusion that enzymes would experience in reality. In our model, enzymes tend to favor less dense regions, which are usually located on the external part of the granule, solely because reactions in denser regions are rejected a posteriori. To improve the accuracy of our model, we could prevent the enzymes from accessing chains in regions with too high density, according to their size. However, if we consider the volume exclusion of the enzyme, it would be impossible for them to react if a single glucose overlaps with their volume. To address this issue, we could incorporate a maximal density for which the enzyme is prevented from reacting, which would decrease the number of accessible chains and potentially alter the dynamic behavior of the system.

Additionally, our glycogen model assumes that branches are arranged in a helical structure. While this is the most stable configuration for single chains, interactions between chains could lead to variations in the dihedral angles forming $\alpha - 1,4$ bonds. To improve the accuracy of our model, we could randomly draw dihedral angles from a Ramachandran energy map, which provides the energy profiles of the two angles for a pair of glucose molecules. We could extend such a model to different types of polysaccharides and consider not only $\alpha - 1,4$ and $\alpha - 1,6$ linkages.

We could also investigate the generalization of the model to a generic dendrimer structure and study some physical properties of such a model. Furthermore, we suspect that the entropy of configuration may play an important role in the chain length distribution of the structure. To investigate this, a simpler version of the model where branching is fully random could help us determine the contribution of the combinatory effect compared to the enzyme mechanistic contribution.

Investigation glycogen synthesis under other branching scenarios

Despite an already detailed analysis of the branching enzyme mechanism, it could actually be more complex. Recent study have found that GBE is potentially able to hydrolyse a chain to form preferred substrate, releasing a small chain in the system. Modelling this process could introduce variation in the chain length distribution analysis.

A theoretical approach to the coupling algorithm

Our periodic-coupling algorithm has been effective and more efficient than a full stochastic method. However, there is significant scope for improvement in achieving greater optimality. One example of sub-optimality in our algorithm is due to the constant frequency at which the stochastic and deterministic module communicates. This means that during steady-state of a system, the communication time remains the same as during a transient regimes where the system variations are higher. To address this issue, we could extend the communication time during periods of low activity to improve efficiency, and reduce it when strong variations appear. To achieve this goal, a more theoretical approach is required, and one potential strategy would be to incorporate a communication time that is dependent on the system's variation.

Glycogen metabolism

Our current reduced glycogen model has shown promising results in capturing complex regulatory mechanisms and exhibiting interesting behaviors under different conditions. However, it has some limitations, as it only considers glycogen degradation as the source of glucose production. Therefore, incorporating a more detailed model that includes both glycolysis and reverse glycolysis pathways would enable us to have a second source of glucose production, allowing for a better characterization of disease phenotype. Additionally, the inclusion of the TCA cycle could provide further insight into how energy is managed in disease cases.

Although the glucose influx and outflux rates have been adjusted to match observed rates of glycogen synthesis and degradation, it is essential to incorporate data on glucose intakes and the relative contributions of glycolysis and glycogenolysis to glucose homeostasis into the model. Furthermore, other regulatory mechanisms, such as the role of carbohydrate-responsive element-binding protein (ChREBP), a transcription factor that plays a crucial role in insulin sensitivity, should also be incorporated.

Bibliography

Aalst, Marvin van, Oliver Ebenhöh, and Anna Matuszyńska (2020). "Constructing and analysing dynamic models with modelbase v1.2.3 - a software update". In: *bioRxiv*, pp. 1–15. ISSN: 26928205. DOI: 10.1101/2020.09.30.321380. URL: https://doi.org/10.1186/s12859-021-04122-7 (cit. on p. 25).

Andersson, Christian, Claus Führer, and Johan Åkesson (2015). "Assimulo: A unified framework for {ODE} solvers". In: *Mathematics and Computers in Simulation* 116.0, pp. 26–43. ISSN: 0378-4754. DOI: http://dx.doi.org/10.1016/j.matcom.2015.04.007 (cit. on p. 25).

Arko, Janez Jan, Marusa Debeljak, Mojca Zerjav Tansek, Tadej Battelino, and Urh Groselj (Aug. 2020). "A patient with glycogen storage disease type 0 and a novel sequence variant in GYS2: a case report and literature review". en. In: *J Int Med Res* 48.8, p. 300060520936857 (cit. on p. 100).

Bali, Deeksha S, Areeg El-Gharbawy, Stephanie Austin, Surekha Pendyal, and Priya S Kishnani (1993). "Glycogen Storage Disease Type I". en. In: *GeneReviews(®)*. Seattle (WA): University of Washington, Seattle (cit. on p. 8).

Berg, Jeremy M., John L. Tymoczko, Gregory J. Gatto jr., and Lubert Stryer (2018). *Stryer Biochemie*. Springer Spektrum Berlin, Heidelberg. DOI: https://doi.org/10.1007/978-3-662-54620-8 (cit. on pp. 1, 3, 16).

Besford, Quinn Alexander, Xiao-Yi Zeng, Ji-Ming Ye, and Angus Gray-Weale (Oct. 2015). "Liver glycogen in type 2 diabetic mice is randomly branched as enlarged aggregates with blunted glucose release". en. In: *Glycoconj J* 33.1, pp. 41–51 (cit. on pp. 117, 120).

Cao, Yang, Daniel T. Gillespie, and Linda R. Petzold (2005). "The slow-scale stochastic simulation algorithm". In: *The Journal of Chemical Physics* 122.1, p. 014116. DOI: 10.1063/1.1824902. eprint: https://doi.org/10.1063/1.1824902. URL: https://doi.org/10.1063/1.1824902 (cit. on p. 71).

Cryer, P. E. (1993). "Glucose counterregulation: prevention and correction of hypoglycemia in humans". In: *American Journal of Physiology-Endocrinology and Metabolism* 264.2. PMID: 8447379, E149–E155. DOI: 10.1152/ajpendo.1993.264.2.E149. eprint: https://doi.org/10.1152/ajpendo.1993.264.2.E149. URL: https://doi.org/10.1152/ajpendo.1993.264.2.E149 (cit. on p. 2).

Dasouki, Majed, Omar Jawdat, Osama Almadhoun, Mamatha Pasnoor, April L McVey, Ahmad Abuzinadah, Laura Herbelin, Richard J Barohn, and Mazen M Dimachkie (Aug. 2014). "Pompe disease: literature review and case series". en. In: *Neurol Clin* 32.3, pp. 751–76, ix (cit. on p. 8).

Drochmans, P. (1962). "Morphologie du glycogène: Etude au microscope électronique de colorations négatives du glycogène particulaire". In: *Journal of Ultrastructure Research* 6.2, pp. 141–163. ISSN: 0022-5320. DOI: https://doi.org/10.1016/S0022-5320(62)90050-3 (cit. on pp. 2, 3, 6, 116).

Federation, International Diabetes (2011). *IDF diabetes atlas*. International Diabetes Federation, Executive Office (cit. on p. 9).

Ferrannini, Eleuterio, Ola Bjorkman, George A Reichard Jr, Alessandro Pilo, Maggie Olsson, John Wahren, and Ralph A DeFronzo (1985). "The disposal of an oral glucose load in healthy subjects: a quantitative study". In: *Diabetes* 34.6, pp. 580–588 (cit. on p. 86).

Fréchet, Jean M. J. (2003). "Dendrimers and other dendritic macromolecules: From building blocks to functional assemblies in nanoscience and nanotechnology". In: *Journal of Polymer Science Part A: Polymer Chemistry* 41.23, pp. 3713–3725. DOI: https://doi.org/10.1002/pola.10952. eprint: https://onlinelibrary.wiley.com/doi/pdf/10.1002/pola.10952. URL: https://onlinelibrary.wiley.com/doi/abs/10.1002/pola.10952 (cit. on pp. 6, 116).

Gennes, P. G. de (1979). *Scaling concepts in polymer physics / Pierre-Gilles de Gennes*. English. Cornell University Press Ithaca, N.Y, 324 p. : ISBN: 080141203 (cit. on p. 15).

Gerich, John E. (1993). "Control of glycaemia". In: *Baillière's Clinical Endocrinology and Metabolism* 7.3. Hypoglycaemia, pp. 551–586. ISSN: 0950-351X. DOI: https://doi.org/10.1016/S0950-351X(05)80207-1. URL: https://www.sciencedirect.com/science/article/pii/S0950351X05802071 (cit. on p. 2).

Ggunja-Smith, Zeenat, J.J. Marshall, and E.E. Smith (1971). "Enzymatic determination of the unit chain length of glycogen and related polysaccharides". In: *FEBS Letters* 13.5, pp. 309–311. DOI: 10.1016/0014-5793(71)80248-X (cit. on pp. 2, 4).

Gillespie, Daniel T (1976). "A general method for numerically simulating the stochastic time evolution of coupled chemical reactions". In: *Journal of Computational Physics* 22.4, pp. 403–434. ISSN: 0021-9991. DOI: https://doi.org/10.1016/0021-9991(76)90041-3. URL: https://www.sciencedirect.com/science/article/pii/0021999176900413 (cit. on pp. 19, 23, 70).

– (1977a). "Exact stochastic simulation of coupled chemical reactions". In: *The Journal of Physical Chemistry* 81.25, pp. 2340–2361. DOI: 10.1021/j100540a008. eprint: https://doi.org/10.1021/j100540a008. URL: https://doi.org/10.1021/j100540a008 (cit. on pp. 19, 70).

– (1977b). "Exact stochastic simulation of coupled chemical reactions". In: *The Journal of Physical Chemistry* 81.25, pp. 2340–2361. DOI: 10.1021/j100540a008. eprint: https://doi.org/10.1021/j100540a008. URL: https://doi.org/10.1021/j100540a008 (cit. on p. 70).

– (2001). "Approximate accelerated stochastic simulation of chemically reacting systems". In: *The Journal of Chemical Physics* 115.4, pp. 1716–1733. DOI: 10.1063/1.1378322. eprint: https://doi.org/10.1063/1.1378322. URL: https://doi.org/10.1063/1.1378322 (cit. on p. 70).

Goldsmith, E, S Sprang, and R Fletterick (1982). "Structure of maltoheptaose by difference Fourier methods and a model for glycogen". In: *J Mol Biol.* 156.2, pp. 411–427. DOI: 10.1016/0022-2836(82)90336-9 (cit. on pp. 5, 6, 12, 115, 117).

Gunja-Smith, Zeenat, J.J. Marshall, Christiane Mercier, E.E. Smith, and W.J. Whelan (1970a). "A revision of the Meyer-Bernfeld model of glycogen and amylopectin". In: *FEBS Letters* 12.2, pp. 101–104. DOI: 10.1016/0014-5793(70)80573-7 (cit. on pp. 4, 5).

– (1970b). "A revision of the Meyer-Bernfeld model of glycogen and amylopectin". In: *FEBS Letters* 12.2, pp. 101–104. DOI: https://doi.org/10.1016/0014-5793(70)80573-7. eprint: https://febs.onlinelibrary.wiley.com/doi/pdf/10.1016/0014-5793%2870%2980573-7. URL: https://febs.onlinelibrary.wiley.com/doi/abs/10.1016/0014-5793%2870%2980573-7 (cit. on pp. 4, 5, 115).

Gunja-Smith, Zeenat, J.J. Marshall, E.E. Smith, and W.J. Whelan (1970c). "A glycogen-debranching enzyme from Cytophaga". In: *FEBS Letters* 12.2, pp. 96–100. DOI: 10.1016/0014-5793(70)80572-5 (cit. on p. 4).

Haseltine, Eric L. and James B. Rawlings (2002). "Approximate simulation of coupled fast and slow reactions for stochastic chemical kinetics". In: *The Journal of Chemical Physics* 117.15, pp. 6959–6969. DOI: 10.1063/1.1505860. eprint: https://doi.org/10.1063/1.1505860. URL: https://doi.org/10.1063/1.1505860 (cit. on p. 71).

Hejazi, Mahdi, Joerg Fettke, Sophie Haebel, Christoph Edner, Oskar Paris, Claus Frohberg, Martin Steup, and Gerhard Ritte (July 2008). "Glucan, water dikinase phosphorylates crystalline maltodextrins and thereby initiates solubilization". en. In: *Plant J* 55.2, pp. 323–334 (cit. on p. 110).

Hicks, John, Eric Wartchow, and Gary Mierau (2011). "Glycogen Storage Diseases: A Brief Review and Update on Clinical Features, Genetic Abnormalities, Pathologic Features, and Treatment". In: *Ultrastructural Pathology* 35.5. PMID: 21910565, pp. 183–196. DOI: 10.3109/01913123.2011.601404. eprint: https://doi.org/10.3109/01913123.2011.601404. URL: https://doi.org/10.3109/01913123.2011.601404 (cit. on pp. 7, 8).

Kim, Damin and Jean Duhamel (2023). "Interior of glycogen probed by pyrene excimer fluorescence". In: *Carbohydrate Polymers* 299, p. 120205. ISSN: 0144-8617. DOI: https://doi.org/10.1016/j.carbpol.2022.120205. URL: https://www.sciencedirect.com/science/article/pii/S0144861722011109 (cit. on pp. 6, 117).

Kishnani, Priya S., Stephanie L. Austin, Jose E. Abdenur, Pamela Arn, Deeksha S. Bali, Anne Boney, Wendy K. Chung, Aditi I. Dagli, David Dale, Dwight Koeberl, Michael J. Somers,

Stephanie Burns Wechsler, David A. Weinstein, Joseph I. Wolfsdorf, and Michael S. Watson (2014). "Diagnosis and management of glycogen storage disease type I: a practice guideline of the American College of Medical Genetics and Genomics". In: *Genetics in Medicine* 16.11, e1–e29. ISSN: 1098-3600. DOI: https://doi.org/10.1038/gim.2014.128. URL: https://www.sciencedirect.com/science/article/pii/S1098360021026514 (cit. on p. 102).

König, Matthias, Sascha Bulik, and Hermann-Georg Holzhütter (June 2012). "Quantifying the contribution of the liver to glucose homeostasis: a detailed kinetic model of human hepatic glucose metabolism". en. In: *PLoS Comput Biol* 8.6, e1002577 (cit. on pp. 82, 83, 85, 86, 92, 95).

Lehninger, Albert L, David L Nelson, and Michael M Cox (2017). *Lehninger Principles of Biochemistry*. W. H. Freeman (cit. on p. 2).

Li, Cheng and Zhenxia Hu (2020). "Is liver glycogen fragility a possible drug target for diabetes?" In: *The FASEB Journal* 34.1, pp. 3–15. DOI: https://doi.org/10.1096/fj.201901463RR. eprint: https://faseb.onlinelibrary.wiley.com/doi/pdf/10.1096/fj.201901463RR. URL: https://faseb.onlinelibrary.wiley.com/doi/abs/10.1096/fj.201901463RR (cit. on p. 9).

Lodish, Harvey, Arnold Berk, Chris A Kaiser, Monty Krieger, Matthew P Scott, Anthony Bretscher, Hidde Ploegh, and Paul T Matsudaira (Aug. 2007). *Molecular Cell Biology*. 6th ed. New York, NY: W.H. Freeman (cit. on p. 2).

Magnusson, I, DL Rothman, LD Katz, RG Shulman, GI Shulman, et al. (1992). "Increased rate of gluconeogenesis in type II diabetes mellitus. A 13C nuclear magnetic resonance study." In: *The Journal of clinical investigation* 90.4, pp. 1323–1327 (cit. on p. 86).

Marshall, J.J. and W.J. Whelan (1974). "Multiple branching in glycogen and amylopectin". In: *Archives of Biochemistry and Biophysics* 161.1, pp. 234–238. ISSN: 0003-9861. DOI: https://doi.org/10.1016/0003-9861(74)90256-2 (cit. on p. 4).

Meléndez, Ruth, Enrique Meléndez-Hevia, and Enric I. Canela (1999). "The Fractal Structure of Glycogen: A Clever Solution to Optimize Cell Metabolism". In: *Biophysical Journal* 77.3, pp. 1327–1332. ISSN: 0006-3495. DOI: https://doi.org/10.1016/S0006-3495(99)76982-1. URL: https://www.sciencedirect.com/science/article/pii/S0006349599769821 (cit. on p. 6).

Meléndez-Hevia, E, T G Waddell, and E D Shelton (Oct. 1993). "Optimization of molecular design in the evolution of metabolism: the glycogen molecule". en. In: *Biochem J* 295 (Pt 2).Pt 2, pp. 477–483 (cit. on pp. 3, 6, 117).

Meyer, Kurt H. and P. Bernfeld (1940). "Recherches sur l'amidon V. L'amylopectine". In: *Helvetica Chimica Acta* 23.1, pp. 875–885. DOI: https://doi.org/10.1002/hlca.194002301112 (cit. on p. 4).

Nuttall, Frank Q., Angela Ngo, and Mary C. Gannon (2008). "Regulation of hepatic glucose production and the role of gluconeogenesis in humans: is the rate of gluconeogenesis constant?" In: *Diabetes/Metabolism Research and Reviews* 24.6, pp. 438–458. DOI: https://doi.org/10.1002/dmrr.863. eprint: https://onlinelibrary.wiley.com/doi/pdf/10.1002/dmrr.863. URL: https://onlinelibrary.wiley.com/doi/abs/10.1002/dmrr.863 (cit. on pp. 88, 99).

Orho, M, N U Bosshard, N R Buist, R Gitzelmann, A Aynsley-Green, P Blümel, M C Gannon, F Q Nuttall, and L C Groop (Aug. 1998). "Mutations in the liver glycogen synthase gene in children with hypoglycemia due to glycogen storage disease type 0". en. In: *J Clin Invest* 102.3, pp. 507–515 (cit. on p. 8).

Oster, Yonatan, Isaiah D Wexler, Samuel N Heyman, and Elchanan Fried (Nov. 2016). "Recoverable, Record-High Lactic Acidosis in a Patient with Glycogen Storage Disease Type 1: A Mixed Type A and Type B Lactate Disorder". en. In: *Case Rep Med* 2016, p. 4362743 (cit. on p. 102).

Ozen, Hasan (May 2007a). "Glycogen storage diseases: new perspectives". en. In: *World J Gastroenterol* 13.18, pp. 2541–2553 (cit. on p. 7).

– (May 2007b). "Glycogen storage diseases: new perspectives". en. In: *World J Gastroenterol* 13.18, pp. 2541–2553 (cit. on pp. 100, 102).

Palm, Daniel C, Johann M Rohwer, and Jan-Hendrik S Hofmeyr (Dec. 2012). "Regulation of glycogen synthase from mammalian skeletal muscle–a unifying view of allosteric and covalent regulation". en. In: *FEBS J* 280.1, pp. 2–27 (cit. on p. 110).

Parikh, Nirzar S and Rajni Ahlawat (Jan. 2023). "Glycogen Storage Disease Type I". en. In: *StatPearls*. Treasure Island (FL): StatPearls Publishing (cit. on p. 101).

Petzold, L.R. (Jan. 2011). "Legitimacy of the stochastic Michaelis–Menten approximation". English. In: *IET Systems Biology* 5 (1), 58–69(11). ISSN: 1751-8849. URL: https://digital-library.theiet.org/content/journals/10.1049/iet-syb.2009.0057 (cit. on p. 23).

Radziuk, J_ and S Pye (2001). "Hepatic glucose uptake, gluconeogenesis and the regulation of glycogen synthesis". In: *Diabetes/metabolism research and reviews* 17.4, pp. 250–272 (cit. on p. 86).

Roche, Thomas E (2002). "Initiation of glycogen synthesis". In: *Biochemical Society Transactions* 30.2, pp. 279–283 (cit. on p. 2).

Rothman, Douglas L, Inger Magnusson, Lee D Katz, Robert G Shulman, and Gerald I Shulman (1991). "Quantitation of hepatic glycogenolysis and gluconeogenesis in fasting humans with 13C NMR". In: *Science* 254.5031, pp. 573–576 (cit. on p. 86).

Stick, Robert V. (2008). *Carbohydrates: The Essential Molecules of Life, Second Edition*. 2nd. Elsevier. ISBN: 978-0-444-52936-3 (cit. on p. 3).

Sullivan, Mitchell A., Samuel T. N. Aroney, Shihan Li, Frederick J. Warren, Jin Suk Joo, Ka Sin Mak, David I. Stapleton, Kim S. Bell-Anderson, and Robert G. Gilbert (Feb. 2014). "Changes in Glycogen Structure over Feeding Cycle Sheds New Light on Blood-Glucose Control". In: *Biomacromolecules* 15.2, pp. 660–665. ISSN: 1525-7797. DOI: 10.1021/bm401714v. URL: https://doi.org/10.1021/bm401714v (cit. on pp. 3, 94, 116, 120).

Sullivan, Mitchell A., Silvia Nitschke, Evan P. Skwara, Peixiang Wang, Xiaochu Zhao, Xiao S. Pan, Erin E. Chown, Travis Wang, Ami M. Perri, Jennifer P.Y. Lee, Francisco Vilaplana, Berge A. Minassian, and Felix Nitschke (2019). "Skeletal Muscle Glycogen Chain Length Correlates with Insolubility in Mouse Models of Polyglucosan-Associated Neurodegenerative Diseases". In: *Cell Reports* 27.5, 1334–1344.e6. ISSN: 2211-1247. DOI: https://doi.org/10.1016/j.celrep.2019.04.017. URL: https://www.sciencedirect.com/science/article/pii/S2211124719304826 (cit. on p. 110).

Sullivan, Mitchell A., Francisco Vilaplana, Richard A. Cave, David Stapleton, Angus A. Gray-Weale, and Robert G. Gilbert (Apr. 2010). "Nature of α and β Particles in Glycogen Using Molecular Size Distributions". In: *Biomacromolecules* 11.4, pp. 1094–1100. ISSN: 1525-7797. DOI: 10.1021/bm100074p. URL: https://doi.org/10.1021/bm100074p (cit. on p. 4).

Suzuki, K, E David, and B Kutschman (July 1971). "Presenile dementia with "Lafora-like" intraneuronal inclusions". en. In: *Arch Neurol* 25.1, pp. 69–80 (cit. on p. 8).

Tagliabracci, Vincent S, Jean Marie Girard, Dyann Segvich, Catalina Meyer, Julie Turnbull, Xiaochu Zhao, Berge A Minassian, Anna A DePaoli-Roach, and Peter J Roach (2008). "Abnormal metabolism of glycogen phosphate as a cause for Lafora disease". In: *Journal of Biological Chemistry* 283.49, pp. 33816–33825 (cit. on p. 8).

Tagliaferri, Francesco, Miriam Massese, Luisa Russo, Anna Commone, Serena Gasperini, Roberta Pretese, Carlo Dionisi-Vici, and Arianna Maiorana (July 2022). "Hepatic glycogen storage diseases type 0, VI and IX: description of an italian cohort". In: *Orphanet Journal of Rare Diseases* 17.1, p. 285. ISSN: 1750-1172. DOI: 10.1186/s13023-022-02431-5. URL: https://doi.org/10.1186/s13023-022-02431-5 (cit. on p. 100).

Taylor, Martha, Eric Simon, Jean Dickey, Kelly Hogan, and Jane Reece (May 2021). *Campbell biology: Concepts & connections, global edition*. 10th ed. London, England: Pearson Education (cit. on p. 1).

Taylor, Roy, Inger Magnusson, Douglas L Rothman, Gary W Cline, Andrea Caumo, Claudio Cobelli, Gerald I Shulman, et al. (1996). "Direct assessment of liver glycogen storage by 13C nuclear magnetic resonance spectroscopy and regulation of glucose homeostasis after a mixed meal in normal subjects." In: *The Journal of clinical investigation* 97.1, pp. 126–132 (cit. on p. 86).

Thornell, L.-E. (1974). "The fine structure of Purkinje fiber glycogen. A comparative study of negatively stained and cytochemically stained particles". In: *Journal of Ultrastructure Research* 49.2, pp. 157–166. ISSN: 0022-5320. DOI: https://doi.org/10.1016/S0022-

5320(74)80029-8. URL: https://www.sciencedirect.com/science/article/pii/S0022532074800298 (cit. on p. 115).

Wu, Alex Chi and Robert G. Gilbert (Dec. 2010). "Molecular Weight Distributions of Starch Branches Reveal Genetic Constraints on Biosynthesis". In: *Biomacromolecules* 11.12, pp. 3539–3547. ISSN: 1525-7797. DOI: 10.1021/bm1010189. URL: https://doi.org/10.1021/bm1010189 (cit. on p. 9).

Young, F G (June 1957). "Claude Bernard and the discovery of glycogen; a century of retrospect". en. In: *Br Med J* 1.5033, pp. 1431–1437 (cit. on p. 2).

Zhang, Peng, Sharif S. Nada, Xinle Tan, Bin Deng, Mitchell A. Sullivan, and Robert G. Gilbert (2018). "Exploring glycogen biosynthesis through Monte Carlo simulation". In: *International Journal of Biological Macromolecules* 116, pp. 264–271. ISSN: 0141-8130. DOI: https://doi.org/10.1016/j.ijbiomac.2018.05.027. URL: https://www.sciencedirect.com/science/article/pii/S0141813018312923 (cit. on p. 6).

List of Figures

List of Tables

Eidesstattliche Erklärung
laut §5 der Promotionsordnung vom 15.06.2018

Ich versichere an Eides Statt, dass die Dissertation von mir selbständig und ohne unzulässige fremde Hilfe unter Beachtung der „Grundsätze zur Sicherung guter wissenschaftlicher Praxis an der Heinrich-Heine-Universität Düsseldorf“ erstellt worden ist.

Ort, Datum

Yvan Rousset