Kinetic modelling of enzymatic starch hydrolysis

Karolina A. Bednarska

Thesis committee

Promotors

Prof. Dr M.A.J.S. van Boekel
Professor of Product Design and Quality Management
Wageningen University

Prof. Dr R.M. Boom
Professor of Food Process Engineering
Wageningen University

Co-promotor

Dr A.E.M. Janssen
Assistant professor, Food Process Engineering
Wageningen University

Other members

Prof. Dr Harry Gruppen, Wageningen University
Prof. Dr Mary Chan Bee Eng, Nanyang Technological University, Singapore
Dr Marco van den Berg, DSM
Prof. Dr Lubbert Dijkhuizen, University of Groningen

Kinetic modelling of enzymatic starch hydrolysis

Karolina A. Bednarska

Table of contents

CHAPTER

1

Introduction

1.1. Starch and the products of its hydrolysis

Starch, the main energy storage component in plants, is present in the form of granules. These granules are highly organized semi-spherical structures that, depending on the plant source, present different morphology and composition. The differences in the granules stem from the proportions and packaging of amylose and amylopectin - the two types of glucose polymers bound by glycosidic bonds. These two components of starch also differ in structure and their properties.

Amylose is a nearly linear polymer of glucose units joined with α-(1,4)-glycosidic bonds, with only 0.2-0.8% of branches formed by α-(1,6)-glycosidic bonds (Maningat et al. 2009). In wheat starch nearly 25% of the glucose units are present in the form of amylose (Swinkels 1985). The remaining 75% constitute the branched molecules of amylopectin. Around 5% of the glucose units in wheat amylopectin are bound by α(1,6)-glycosidic linkages (Maningat et al. 2009). Amylopectin branches have different lengths, characteristic of the source of starch and the position within the macromolecule. The chain length distributions are a form of characterizing amylopectin, by distinguishing three main types of chains (Figure 1.1). The A chains are the outer chains linked by α-(1,6)-glycosidic bonds to the inner chains of amylopectin. The B chains form the main part of the clusters and carry A chains and other B chains. The C chain is the sole carrier of a reducing glucose unit (Hizukuri and Maehara 1990). The exact arrangement of these types of chains in amylopectin is still being investigated, as the older cluster model (Hizukuri 1986) and the alternative building block model (Bertoft 2013) are still debated in literature.

Figure 1.1 A scheme representing the branched structure of amylopectin, following the classical cluster model. The A, B and C chains are indicated.

The most common modification of starch on industrial scale is the enzymatic hydrolysis. During hydrolysis, enzymes break the long chains of amylose and

amylopectin into shorter molecules. Depending on the enzyme's mode of action, the most common final products of starch hydrolysis can be maltodextrins, glucose, fructose or maltose.

The hydrolysis of starch is generally preceded by gelatinization. Gelatinization is a term that includes all events occurring as starch is heated in water: absorption of water by the starch granules, increase of the apparent viscosity of the solution, breakage of granules and exposure of amylose and amylopectin. The exposure of the substrate increases the rate of the enzymatic reaction by providing an easier access for the enzyme.

Conventionally, the hydrolysis of starch by α-amylase is referred to as liquefaction, because α-amylases lower the viscosity of the gelatinized starch (liquefy) by reducing the length of amylose and amylopectin. The products of starch liquefaction are called maltodextrins. Maltodextrin is a mixture of poly- and oligosaccharides with a broad molecular weight distribution. This mixture includes maltose, malto-oligosaccharides and linear and branched dextrins. The composition of maltodextrins depends on both the source of the enzyme and the source of starch. The final product composition will be influenced by both the intrinsic enzyme characteristics (e.g., activity and stability at different temperatures and pH values) and extrinsic characteristics, attributed to the substrate, including degree of gelatinization and amylose-amylopectin ratio (Marchal et al. (1999), Murthy et al. (2011)). Maltodextrins can be used as carriers, bulking agents, crystallization inhibitors, coatings, flavour carriers, fat replacers or texturisers (Kennedy et al. (1988), Marchal et al. (1999), White et al. (2003)).

If desired, maltodextrins can be further hydrolysed, e.g., by glucoamylases in a process called saccharification. The goal of saccharification is producing glucose syrups by decomposing starch to its monomer - glucose. Glucose syrups depending on their degree of hydrolysis, are used in food production (e.g., jams, soft drinks), frozen desserts, confectionery, brewing and fermentation (Kennedy et al. 1988).

Dextrose equivalent (DE) values are used to characterize starch hydrolysates. The dextrose equivalent expresses the extent of starch hydrolysis. It is a measure of the total reducing power of the sugars in relation to glucose and on a dry mass basis (van der Maarel et al. 2002). The higher the value of dextrose equivalent the greater the extent of hydrolysis and the lower the average molecular mass of the oligomeric

products. DE is an average value, and as such provides only an indication of the extent of hydrolysis. Mixtures with the same average DE value often contain different proportions of saccharides.

1.2.   Starch degrading enzymes and their mechanisms

Four main types of starch-converting enzymes are known: endoamylases, exoamylases, debranching enzymes and transferases (van der Maarel et al. 2002). The enzymes used in this thesis belong to the endo- and exoamylases and will be discussed in more detail.

1.2.1.   *Bacillus licheniformis α-amylase (BLA)*

α-Amylases [E.C. 3.2.1.1] are endo-hydrolases that randomly cleave α-(1,4)-glycosidic bonds in polysaccharides containing three or more α-(1,4)-linked glucose units (Dona et al. 2010). The term 'alpha' refers to the α-anomeric configuration of the liberated sugar and not to the configuration of the hydrolysed linkage. Endo-acting bacterial α-amylases produce larger oligosaccharides than their mammalian counterparts, mainly because hydrolysis of α-(1,4) linkages close to the non-reducing end of the substrate is difficult for these enzymes (MacGregor et al. 2001).

α-Amylases commonly require calcium ions to be present during the reaction in order to maintain their structural integrity (Machius et al. 1995). The three calcium ions bound by the *Bacillus licheniformis* α-amylase (BLA) molecule (Figure 1.2 A), along with the calcium ions added to the solution, facilitate the stability of the enzyme, even at elevated temperatures (Nazmi et al. 2008).

The part of the protein where substrates, in this case glucose units of starch, interact with the enzyme and can undergo a chemical reaction is referred to as the active site of the enzyme. The active site of BLA is located in a cleft (Figure 1.2 B), at the point of junction between domain A and B (Figure 1.2 A; Nagano et al. (2001)). The subsite theory assumes that the active site of an enzyme consists of a number of subsites, each able to interact with a monomeric residue of the substrate. Even though each subsite interacts with only one glucose unit, a multipoint linkage at several binding sites is created with the whole oligosaccharide. This linkage allows for a correct arrangement of the long molecules of substrate in the active site (Synowiecki 2007). The subsites are formed by amino acid residues that interact (e.g., form hydrogen bonds, hydrophobic

interactions (Hiromi et al. 1973) with the glucose units (Figure 1.2 B). Theoretically the energies of the interactions between the amino acids and the monomers of the substrate (the binding energies) can be quantified.
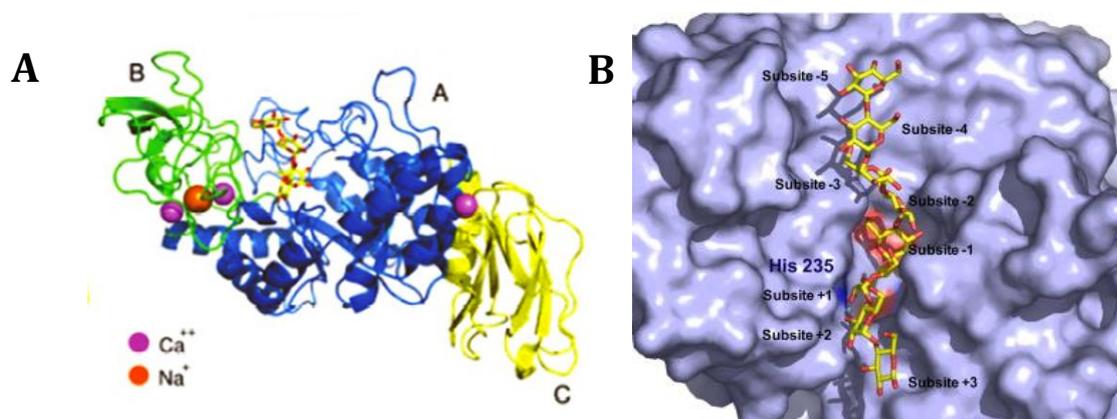

Figure 1.2 The model structure of *Bacillus licheniformis* α-amylase. (A) The three domains A, B and C of BLA are shown along with bound calcium and sodium ions (Muñoz et al. 2011) (B) Molecular model of the active site of BLA with a bound substrate (DP 8). The red colour indicates the residues of the active site (Tran et al. 2014a).

The binding energies of the subsites of an enzyme can be displayed in the form of a subsite map (Figure 1.3). In the subsite map, the cleavage site (the location in which the bond between two molecules will be hydrolysed) is located between subsite +1 and -1. Conventionally, subsites located to the right from the cleavage site (aglycone subsites) are assigned positive numbers and interact with the reducing end-side of the molecule. Subsites to the left (glycone subsites) are given negative numbers and are positioned towards the non-reducing end of the polymer.
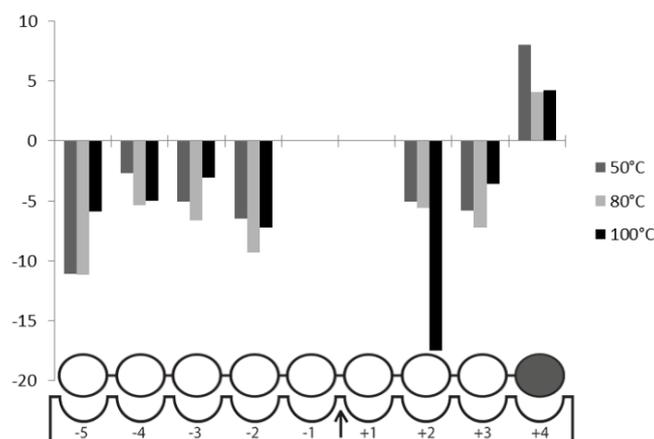

Figure 1.3 The binding energies (kJ·mol^{-1}) in subsite maps of BLA at different temperatures (adapted from Kandra et al. (2006)). The scheme underneath represents the subsite map interacting with an oligosaccharide (DP 9). The reducing end glucose unit is coloured. The arrow indicates where the bond cleavage occurs.

The binding energy values assigned to the subsites can be both negative and positive. The more negative the energy values, the stronger the attraction between the subsite and the monomeric residue. Positive binding energies signify a weaker interaction, implying repulsion. A positive binding energy has been assigned to a barrier site in subsite map of BLA (subsite +4). The barrier site enforces the often observed dual specificity of BLA, and without it a more equal distribution of products is observed (Kandra et al. 2002). Subsites with a high positive energy play a role in positioning the substrate in the active site and removing the products after hydrolysis.

The composition of products after starch hydrolysis depends on the number of subsites in the active site and the binding energy values assigned to these subsites (MacGregor et al. 2001). BLA contains nine subsites (Kandra et al. 2002), but new evidence from molecular modelling suggest that another subsite, subsite -6, might be present in BLA (Tran et al. 2014b). According to Kandra et al. (2006) the values of binding energies of BLA subsites depend on the temperature. The authors describe the binding of substrates at 80°C as more favourable and stronger than at 50°C.

Formation of the enzyme-substrate complex allows several possible binding modes, of which only those that overlap the catalytic site can become productive. The binding mode is productive when the glycosidic bond is hydrolysed. If the glucose units are bound by other subsites, but do not cover the catalytic site or cover the catalytic site but are not hydrolysed, the enzyme-substrate complexes are non-productive. The subsite maps of α-amylases are developed with experiments using short, linear substrates. However, enzymes like BLA, which can bind macromolecules, are able to bind small molecules in many different ways. There can be numerous non-productive bindings between the enzyme and the substrates that are not observed in the kinetics.

Bacillus licheniformis α-amylase (BLA) is commonly used in the industry as a liquefying enzyme. BLA is suitable for industrial use, because of its ability to remain active for several hours at temperatures over 90°C (Fitter et al. (2001), Declerck et al. (2003)). BLA is a stable enzyme. Not only is it active at high temperatures, but it remains stable at these temperatures for a long time (Fitter et al. 2001). The thermostability of enzymes is an evolutionary property – the enzymes must be active at their physiologically relevant temperature, but lower their catalytic efficiency when temperature drastically changes (Jaenicke (2000), Arnold et al. (2001)). The strategies

used by enzymes to achieve thermostability are not fully understood, but most likely they are not the direct result of the amino acid composition (Jaenicke 2000). All weak interactions that can contribute to the (thermo)stability of enzymes, e.g., hydrogen bonds, hydrophobic interactions, ion and metal binding, are affected by temperature to different degrees (Fontana et al. (1998), Jaenicke (2000), Arnold et al. (2001)).

1.2.2. *Glucoamylase from Aspergillus niger*

Glucoamylase (EC 3.2.1.3.) belongs to a family of exo-hydrolases. It hydrolyses the terminal α-1,4-glycosydic bonds of glucose from the non-reducing ends of amylose and amylopectin and releases glucose in beta-configuration. Unlike most α-amylases, glucoamylases can cleave the α-1,6-glycosydic bonds. The hydrolysis rate of the branch points is however 30-50 times slower than that of α-1,4-glycosydic bonds (Dona et al. 2010).

Aspergillus niger has been used for commercial production of glucoamylase (Pandey 1995). Fungal glucoamylases are usually glycoproteins consisting of more than one isoform, with the one from *A. niger* having two: GA I and GA II (James and Lee 1997). The GA I isoform consists of a catalytic site region, a highly glycosylated linker and a starch binding domain (Figure 1.4 A). A lower molecular weight due to the lack of the starch binding domain is what distinguishes the GA II isoform from GA I (Wong 1995). The active sites of the two enzyme forms do not differ in kinetic properties and subsite structure (Figure 1.4 B). Even their subsite maps are not significantly different (Ermer et al. 1993).

The purpose of the starch binding domain is to anchor the enzyme to the chain of the substrate (Juge et al. 2002), even though the absence of this domain does not impair the activity of glucoamylase towards soluble starch. After the enzyme is attached to the substrate, the non-reducing end enters the catalytic domain and hydrolysis can occur. As glucose is removed from the catalytic site, glucoamylase can dissociate from the substrate (James and Lee 1997).
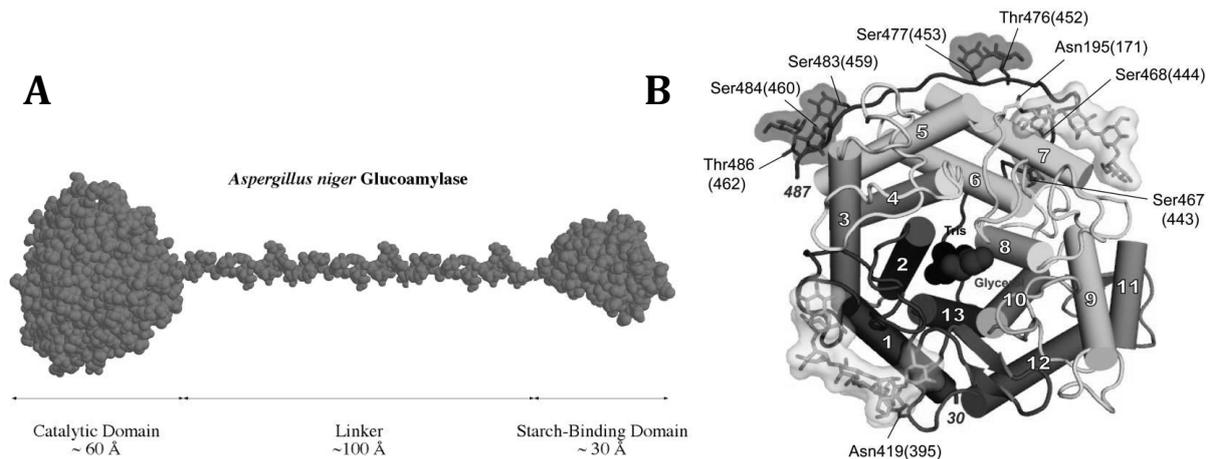
Figure 1.4 The schematic representation of *A. niger* glucoamylase (GA I). (A) A hypothetical arrangement of the domains of glucoamylase (Coutinho et al. 1997); (B) The structure of the catalytic domain (Lee and Paetzel 2011). The glycosylation of the protein by mannose (top) and NAG oligosaccharides (on the left) is shown.

The subsite theory for glucoamylases was described by Hiromi (1970; 1973). Different glucoamylases were assigned a common subsite arrangement: seven subsites present in the active site, with the catalytic site located between subsite -1 and +1. Glucoamylases contain only one glycone subsite, subsite -1 (Ermer et al. 1993). The affinity of the interaction between a glucose unit and a subsite corresponds to the decrease in the binding free energy (Ermer et al. 1993).

Glucoamylases are generally more stable at acidic pH and their optimum temperatures range from 40-60°C. The optimum experimental conditions of *A. niger* glucoamylase are pH 4.2-5 and temperature of 60°C (Crabb and Mitchinson (1997), James and Lee (1997)). The presence of metal ions is not required for the action of GA. Glucoamylases are primarily used to produce high glucose syrups, which can then be converted into crystalline dextrose, high fructose syrups or fermented in the production of ethanol, amino acids or organic acids (Crabb and Mitchinson 1997).

1.3. Models of starch hydrolysis

Empirical models are mathematical models that are fitted to the experimental data without the support of a chemical, physical or biological theory. The models that do use an underlying theory are referred to as mechanistic models. Models can be either deterministic, where as long as the input remains the same the output will not change, or stochastic, where the outcome differs with each simulation (van Boekel 2008).

The difficulties in modelling of starch hydrolysis stem from the complexity of defining the structure of the substrate, still debated and source dependent, and the action modes of enzymes. Predicting the outcomes of starch hydrolysis is complex, as two polymers of different molecular weight distributions and two types of linkages are hydrolysed at different rates. The hydrolysis of numerous linear and branched substrates occurs simultaneously and the action patterns of products additionally depend on the source and the type of used enzyme.

Murthy et al. (2011) describes three approaches to the modelling of starch hydrolysis:

- empirical modelling of the sugar concentrations by curve fitting to experimental data (Paolucci-Jeanjean et al. 2000),
- using differential equations (dynamic models) where reaction rate is described with expressions used for enzyme kinetics (Marc et al. (1983), Henderson and Teague (1988)),
- modelling of hydrolysis by describing the structure of starch and simulation of its hydrolysis (mechanistic models; Park and Rollings (1994), Marchal et al. (2003), Wojciechowski et al. (2001), Besselink et al. (2008)).

Many models have been developed to describe elements of starch hydrolysis reaction. They vary in both the approach used for modelling, and in the purpose of the model. Some models focus on finding the best set of reaction conditions (Åkerberg et al. (2000), Henderson and Teague (1988)), others try to predict the extent of hydrolysis (DE), the decrease of substrate concentration or the increase in product concentrations during the hydrolysis (Komolprasert and Ofoli (1991), Wojciechowski et al. (2001), Paolucci-Jeanjean et al. (2000), Åkerberg et al. (2000), Park and Rollings (1995), Park and Rollings (1994), Murthy et al. (2011), Besselink et al. (2008), Marchal et al. (2003), Bryjak et al. (2000)). There are models that define starch as a substrate (Murthy et al. 2011), and those that focus on either amylose or amylopectin (Marchal et al. 2003) or even smaller, well defined oligosaccharides. Many use isolated starch, some follow the reaction with less purified substrates, e.g., during mashing (Marc et al. (1983), Koljonen et al. (1995)). Overall models, just as the experimental conditions, differ in: pH, temperature, presence or lack of buffers, enzyme dosage, substrate concentration, enzyme type and source, starch source etc.

The three types of models described by Murthy et al. (2011) are used to follow elements of liquefaction and saccharification. In each of the categories models can be found that aim at: describing the inhibition, the inactivation of enzymes or the enzyme activity (Beschkov et al. (1984), Zanin and De Moraes (1996), González-Tello et al. (1996), Cepeda et al. (2001), Polakovič and Bryjak (2002), Apar and Ozbek (2004), Bryjak et al. (2004), Baks et al. (2006), Brandam et al. (2003)), the reversed reactions (Beschkov et al. (1984), Zanin and De Moraes (1996), Nikolov et al. (1989), Matsuno et al. (1978)), predicting the hydrolysis rate (González-Tello et al. (1996), Bryjak et al. (2000)), the DE (Henderson and Teague (1988)), the conversion of substrate (Rodriguez et al. 2006), the concentration of selected products (Rollings and Thompson (1984), Koljonen et al. (1995), Lee et al. (1992), MacGregor and MacGregor (1985), Nagy et al. (1992), (Vidal et al. 2009), Morales et al. (2008), Bryjak et al. (2000), Åkerberg et al. (2000), Paolucci-Jeanjean et al. (2000), Polakovič and Bryjak (2004), Marchal et al. (2003), Wojciechowski et al. (2001), van der Veen et al. (2005), Besselink et al. (2008), Murthy et al. (2011), Nakatani (1996)), or kinetic parameters (Allen and Thoma (1976), Kandra et al. (2002), Gyémánt et al. (2002), Hiromi et al. (1983)).

It has been hypothesized that the assumptions of the subsite theory can be used to theoretically predict product distribution during the time course of hydrolysis (Hiromi et al. 1983; Torgerson et al. 1979). Subsite mapping is a method to calculate the values of apparent binding energies of a substrate to the subsites of enzymes. Subsite mapping can be performed by analysing the bond-cleavage frequencies (BCF) of maltooligosaccharides of known degree of polymerization (DP) and high purity, often labelled with a chromophore. Bond cleavage frequencies are defined as relative rates of formation of every hydrolysis product (Kandra et al. 2002). To minimize transglycosylation and condensation that can complicate the estimations, the substrate concentration has to be low and the reaction products need to be analysed at the initial reaction time (<10% conversion; Kandra et al. (2006)). The effect of the label attached to the substrate is uncertain and even the authors of the papers investigating the action patterns confirm that the labels might interact with the enzymes' subsites and influence the binding modes (Ermer et al. 1993).

The subsite maps of enzymes are established, so that they can be used to predict the composition of hydrolysis products during the hydrolysis of a particular substrate by

a particular enzyme. This has been done by Marchal et al. (2003) and Besselink et al. (2008) for amylases of *Bacillus amyloliquefaciens* on potato amylopectin, and *Bacillus amyloliquefaciens* and *Bacillus licheniformis* on wheat starch. The authors did not consider how temperature might affect the product composition and focused only on small carbohydrates and low concentrations of the substrates. However, Marchal et al. (2003) expected that extending the model towards larger carbohydrates will not be an issue, as the mechanistic model allows for predictions outside the experimentally evaluated range of data.

1.4.   Intensifications of starch hydrolysis - low water conditions

In the industry the gelatinization and liquefaction processes take place at 30-35% solids (van der Maarel et al. 2002). In a 35 w/w% starch-water mixture, 5% of the initial water concentrations is used for chemical gain (van der Veen et al. 2005). That said, starch processing in the industry uses around 5 times the minimal required water content (Figure 1.5). The excess of water is necessary for easier gelatinization and mixing (van der Veen et al. 2006).
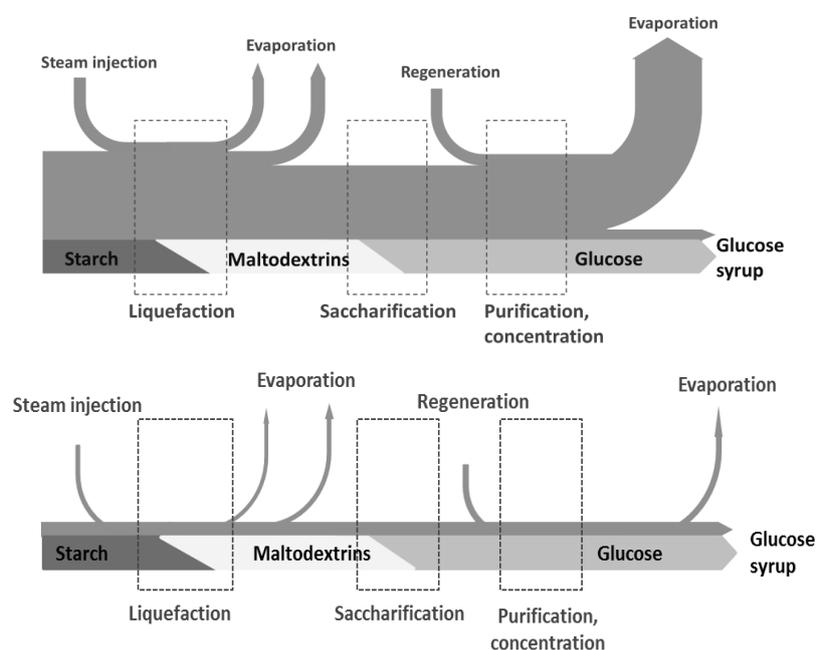

Figure 1.5 A scheme of the mass flow of water and dry matter in the current process of starch hydrolysis (top) and in a theoretical process at high dry matter content (bottom). Scheme adapted from (van der Veen 2005).

The benefits of increasing the dry matter contents during starch processing are connected with better use of equipment, lower water consumption and waste

production (Figure 1.5), and even energy savings. Condensation and evaporation steps can be shortened or even omitted when working in a concentrated system.

However, intensifying starch processing is not an easy task. When increasing the concentration of dry matter, the difficulties begin already at the stage of gelatinization. The degree of gelatinization depends on temperature, pressure, starch concentration and treatment time (Baks et al. 2008). The temperature required for complete gelatinization increases with increasing concentration of starch (Baks et al. 2008), as limited availability of water prevents the granules from swelling. If no mechanical force is applied during gelatinization at low moisture content, complete gelatinization cannot be reached at temperatures below 100°C (Baks et al. (2007), van der Veen et al. (2006)).

The gelatinization and liquefaction process should be separated at low moisture contents (van der Veen et al. 2006). Even thermostable enzymes, which can withstand the high processing temperatures, will undergo inactivation when high shear is applied. Additionally, the presence of enzymes hinders the gelatinization process (van der Veen et al. 2006).

At elevated substrate concentrations the enzyme dosage will always remain the limiting factor during hydrolysis. A high enzyme to substrate ratio will decrease the time required for hydrolysis, but using high concentrations of enzymes is not economically feasible. One of the advantages of lowering the moisture content is the stabilizing effect the substrate has on the enzyme (BLA).

A major issue during the saccharification process is the reversed reaction. At high product concentrations glucoamylase starts producing maltose and isomaltose at the expense of glucose, which decreases the conversion degree. The right enzyme dosage, temperature and reaction time are crucial to avoid the unwanted reaction products (Crabb and Mitchinson 1997). As the concentration of glucose increases in the process the risk of glucose crystallization also increases, solidifying the contents of the reactor and hampering mixing. This effect will become more pronounced as the moisture content is reduced.

1.5.   Aim of the thesis

The goal of our research was to extend our insight in the enzymatic liquefaction and saccharification of starch by establishing mechanistic models that allow for prediction of all the products present in the reaction mixtures during hydrolysis. Since α-amylase from *Bacillus licheniformis* and glucoamylase from *Aspergillus niger* are industrially relevant enzymes in starch processing, we choose these enzyme to collect experimental data and validate the predictions of the model. For this reason, we explored the mechanisms of both enzymes, the factors affecting the final product composition and the differences between the enzymatic hydrolysis process conducted at low and at high dry solid contents and different temperatures.

1.6.   Outline of the thesis

This thesis describes the enzymatic hydrolysis and kinetic modelling of liquefaction and saccharification of wheat starch. In **chapter 2** we describe a model predicting the outcome of wheat starch liquefaction by α-amylase from *Bacillus licheniformis* at 50°C. We demonstrate the ability of the model to predict starch hydrolysis products considerably larger than the oligosaccharides considered in the existing models. The model in its extended version follows all the products of wheat starch hydrolysis by BLA separately and despite the quantitative differences, the qualitative predictions are satisfactory. We also show that the difference between the experimental and computed data might stem from the inaccuracy of the subsite map.

In **chapter 3** and **4** the model from chapter 2 is adapted to find a better description of the hydrolysis data at two temperatures (50°C and 80°C), by varying the energy values of the subsite map and evaluating the inhibition. We hypothesize that a subsite map that is based on the cleavage patterns of linear, short molecules does not account for the complexity of hydrolysis of amylopectin. The branched structure of amylopectin molecules influences the composition of the hydrolysis products by restricting the access to some of the bonds. The presence of branches creates steric obstacles for the enzyme. The used α-amylase has difficulties hydrolysing and accommodating α-(1,6)-glycosidic bonds, which imposes on the hydrolysis of the α-(1,4)-glycosidic bonds located in its proximity. On this basis, we analyse the subsite maps in detail and suggest which of the subsites are crucial when making predictions

about the product composition of starch hydrolysates. On top of that we propose new subsite maps that allow a quantitative description of the experimental data.

In **chapter 5** we increase the dry matter content during wheat starch hydrolysis. We follow both the liquefaction by BLA and the saccharification process by glucoamylase from *Aspergillus niger* at low moisture content. The same liquefaction model as in chapters 2, 3 and 4 is used to predict all of the products of wheat starch hydrolysis at higher dry matter contents (30-60 w/w%). The liquefaction model also creates the substrate matrices representing maltodextrins to be used in the saccharification model. The saccharification of liquefacts to glucose is followed with a new mechanistic model, also using the assumptions of the subsite theory. The saccharification model predicts all of the reaction products using the subsite maps of glucoamylase available in literature.

Finally, **chapter 6** contains the general discussion. We start the chapter by demonstrating how the parameters of the liquefaction model at low moisture contents were chosen. The outcomes of the model are also compared with the experimental data at 30-60 w/w%. Next, we test our liquefaction model with starch hydrolysis data at 5 and 60 w/w% taken from literature, to verify both the approach we used and the validity of the parameters we obtained in chapters 3 and 4. The method used to improve the subsite maps is also tested on another enzyme, *Bacillus amyloliquefaciens* α-amylase. After discussing the factors that influence saccharification at high dry matter contents, we conclude the chapter with describing the potential of stochastic modelling and its practical use.

1.7.   References

Åkerberg C, Zacchi G, Torto N, Gorton L. 2000. A kinetic model for enzymatic wheat starch saccharification. Journal of Chemical Technology & Biotechnology 75(4):306-314.

Allen JD, Thoma JA. 1976. Subsite mapping of enzymes - depolymerase computer modeling. Biochemical Journal 159(1):105-120.

Apar DK, Ozbek B. 2004. Alpha-amylase inactivation by temperature during starch hydrolysis. Process biochemistry 39:1137-1144.

Arnold FH, Wintrode PL, Miyazaki K, Gershenson A. 2001. How enzymes adapt: lessons from directed evolution. Trends in Biochemical Sciences 26(2):100-106.

Baks T, Bruins ME, Janssen AEM, Boom RM. 2007. Effect of pressure and temperature on the gelatinization of starch at various starch concentrations. Biomacromolecules 9(1):296-304.

Baks T, Janssen AEM, Boom RM. 2006. A kinetic model to explain the maximum in α-amylase activity measurements in the presence of small carbohydrates. Biotechnology and Bioengineering 94(3):431-440.

Baks T, Kappen FHJ, Janssen AEM, Boom RM. 2008. Towards an optimal process for gelatinisation and hydrolysis of highly concentrated starch–water mixtures with alpha-amylase from B. licheniformis. Journal of Cereal Science 47(2):214-225.

Bertoft E. 2013. On the building block and backbone concepts of amylopectin structure. Cereal Chemistry Journal 90(4):294-311.

Beschkov V, Marc A, Engasser JM. 1984. A kinetic model for the hydrolysis and synthesis of maltose, isomaltose, and maltotriose by glucoamylase. Biotechnology and Bioengineering 26(1):22-26.

Besselink T, Baks T, Janssen A, Boom R. 2008. A stochastic model for predicting dextrose equivalent and saccharide composition during hydrolysis of starch by alpha-amylase. Biotechnology and Bioengineering 100:684 - 697.

Brandam C, Meyer XM, Proth J, Strehaiano P, Pingaud H. 2003. An original kinetic model for the enzymatic hydrolysis of starch during mashing. Biochemical Engineering Journal 13(1):43-52.

Bryjak J, Ciesielski K, Zbiciński I. 2004. Modelling of glucoamylase thermal inactivation in the presence of starch by artificial neural network. Journal of Biotechnology 114(1–2):177-185.

Bryjak J, Murlikiewicz K, Zbicinski I, Stawczyk J. 2000. Application of artificial neural networks to modelling of starch hydrolysis by glucoamylase. Bioprocess Engineering 23(4):351-357.

Cepeda E, Hermosa M, Ballesteros A. 2001. Optimization of maltodextrin hydrolysis by glucoamylase in a batch reactor. Biotechnology and Bioengineering 76(1):70-76.

Coutinho PM, Dowd MK, Reilly PJ. 1997. Automated docking of isomaltose analogues in the glucoamylase active site. Carbohydrate research 297(4):309-324.

Crabb WD, Mitchinson C. 1997. Enzymes involved in the processing of starch to sugars. Trends in Biotechnology 15(9):349-352.

Declerck N, Machius M, Joyet P, Wiegand G, Huber R, Gaillardin C. 2003. Hyperthermostabilization of Bacillus licheniformis -amylase and modulation of its stability over a 50 C temperature range. Protein Engineering Design and Selection 16(4):287-293.

Dona AC, Pages G, Gilbert RG, Kuchel PW. 2010. Digestion of starch: In vivo and in vitro kinetic models used to characterise oligosaccharide or glucose release. Carbohydrate Polymers 80(3):599-617.

Ermer J, Rose K, Hubner G, Schellenberger A. 1993. Subsite affinities of Aspergillus-niger glucoamylase-II determined with P-nitrophenylmaltooligosaccharides. Biological Chemistry Hoppe-Seyler 374(2):123-128.

Fitter J, Herrmann R, Dencher NA, Blume A, Hauss T. 2001. Activity and stability of a thermostable alpha-amylase compared to its mesophilic homologue: Mechanisms of thermal adaptation. Biochemistry 40(35):10723-10731.

Fontana A, Filippis VD, Laureto PPd, Scaramella E, Zambonin M. 1998. Rigidity of thermophilic enzymes. Progress in Biotechnology: Elsevier. p 277-294.

González-Tello P, Camacho F, Jurado E, Guadix EM. 1996. A simple method for obtaining kinetic equations to describe the enzymatic hydrolysis of biopolymers. Journal of Chemical Technology & Biotechnology 67(3):286-290.

Gyémánt G, Hovánszki G, Kandra L. 2002. Subsite mapping of the binding region of α-amylases with a computer program. European Journal of Biochemistry 269(21):5157-5162.

Henderson WE, Teague WM. 1988. A kinetic model of Bacillus stearothermophilus α-amylase under process conditions. Starch - Stärke 40(11):412-418.

Hiromi K. 1970. Interpretation of dependency of rate parameters on the degree of polymerization of substrate in enzyme-catalyzed reactions. Evaluation of subsite affinities of exo-enzyme. Biochemical and Biophysical Research Communications 40(1):1-6.

Hiromi K, Nitta Y, Numata C, Ono S. 1973. Subsite affinities of glucoamylase: Examination of the validity of the subsite theory. Biochimica et Biophysica Acta (BBA) - Enzymology 302(2):362-375.

Hiromi K, Ohnishi M, Tanaka A. 1983. Subsite structure and ligand binding mechanism of glucoamylase. Molecular and Cellular Biochemistry 51(1):79-95.

Hizukuri S. 1986. Polymodal distribution of the chain lengths of amylopectins, and its significance. Carbohydrate Research 147(2):342-347.

Hizukuri S, Maehara Y. 1990. Fine structure of wheat amylopectin: the mode of A to B chain binding. Carbohydrate Research 206(1):145-159.

Jaenicke R. 2000. Do ultrastable proteins from hyperthermophiles have high or low conformational rigidity? Proceedings of the National Academy of Sciences of the United States of America 97(7):2962-2964.

James JA, Lee BH. 1997. Glucoamylases: Microbial sources, industrial applications and molecular biology - A review. Journal of Food Biochemistry 21(1):1-52.

Juge N, Le Gal-Coeffet MF, Furniss CSM, Gunning AP, Kramhoft B, Morris VJ, Williamson G, Svensson B. 2002. The starch binding domain of glucoamylase from Aspergillus niger: overview of its structure, function, and role in raw-starch hydrolysis. Biologia 57:239-245.

Kandra L, Gyemant G, Remenyik J, Hovanszki G, Liptak A. 2002. Action pattern and subsite mapping of Bacillus licheniformis alpha-amylase (BLA) with modified maltooligosaccharide substrates. Febs Letters 518(1-3):79-82.

Kandra L, Remenyik J, Gyémánt G, Lipták A. 2006. Effect of temperature on subsite map of Bacillus licheniformis α-amylase. Acta Biologica Hungarica 57(3):367-375.

Kennedy JF, Cabalda VM, White CA. 1988. Enzymic starch utilization and genetic engineering. Trends in Biotechnology 6(8):184-189.

Koljonen T, Hamalainen JJ, Sjoholm K, Pietila K. 1995. A model for the prediction of fermentable sugar concentrations during mashing. Journal of Food Engineering 26(3):329-350.

Komolprasert V, Ofoli RY. 1991. A dispersion model for predicting the extent of starch liquefaction by Bacillus licheniformis alpha-amylase during reactive extrusion. Biotechnology and Bioengineering 37(7):681-690.

Lee CG, Kim CH, Rhee SK. 1992. A kinetic model and simulation of starch saccharification and simultaneous ethanol fermentation by amyloglucosidase and Zymomonas mobilis. Bioprocess Engineering 7(8):335-341.

Lee J, Paetzel M. 2011. Structure of the catalytic domain of glucoamylase from Aspergillus niger. Acta Crystallographica Section F: Structural Biology and Crystallization Communications 67(Pt 2):188-192.

MacGregor EA, Janeček Š, Svensson B. 2001. Relationship of sequence and structure to specificity in the α-amylase family of enzymes. Biochimica et Biophysica Acta (BBA) - Protein Structure and Molecular Enzymology 1546(1):1-20.

MacGregor EA, MacGregor AW. 1985. A model for the action of cereal alpha amylases on amylose. Carbohydrate Research 142(2):223-236.

Machius M, Wiegand G, Huber R. 1995. Crystal structure of calcium-depleted Bacillus licheniformis alpha-amylase at 2.2-angstrom resolution. Journal of Molecular Biology 246(4):545-559.

Maningat CC, Seib P, Bassi SD, Woo KS, Lasater GD. 2009. Wheat starch: production, properties, modification and uses. In: BeMiller JN, Whistler RL, editors. Starch: Chemistry and Technology: Elsevier Science.

Marc A, Engasser JM, Moll M, Flayeux R. 1983. A kinetic model of starch hydrolysis by α- and β-amylase during mashing. Biotechnology and Bioengineering 25(2):481-496.

Marchal L, Ulijn R, Gooijer C, Franke G, Tramper J. 2003. Monte Carlo simulation of the a-amylolysis of amylopectin potato starch. 2. a-amylolysis of amylopectin. Bioprocess and Biosystems Engineering 26:123 - 132.

Marchal LM, Beeftink HH, Tramper J. 1999. Towards a rational design of commercial maltodextrins. Trends in Food Science & Technology 10(11):345-355.

Matsuno R, Suganuma T, Fujimori H, Nakanishi K, Hiromi K, Kamikubo T. 1978. Rate equation for amylase-catalyzed hydrolysis, transglycosylation and condensation of linear oligosaccharides and amylose. The Journal of Biochemistry 83:385-394.

Morales S, Alvarez H, Sanchez C. 2008. Dynamic models for the production of glucose syrups from cassava starch. Food and Bioproducts Processing 86:25-30.

Muñoz J, Quintero M, Gutierrez PA. 2011. Characterization of the alpha-amylase gene from Bacillus sp. BBM1. Vitae 18:363-369.

Murthy G, Johnston D, Rausch K, Tumbleson M, Singh V. 2011. Starch hydrolysis modeling: application to fuel ethanol production. Bioprocess and Biosystems Engineering 34:879 - 890.

Nagano N, Porter CT, Thornton JM. 2001. The (βα)8 glycosidases: sequence and structure analyses suggest distant evolutionary relationships. Protein Engineering 14(11):845-855.

Nagy E, Belafibako K, Szabo L. 1992. A kinetic study of the hydrolysis of maltodextrin by soluble glucoamylase. Starch-Starke 44(4):145-149.

Nakatani H. 1996. Monte Carlo simulation of multiple attack mechanism of alpha-amylase. Biopolymers 39(5):665-669.

Nazmi AR, Reinisch T, Hinz HJ. 2008. Calorimetric studies on renaturation by CaCl2 addition of metal-free α-amylase from Bacillus Licheniformis (BLA). Journal of Thermal Analysis and Calorimetry 91(1):141-149.

Nikolov ZL, Meagher MM, Reilly PJ. 1989. Kinetics, equilibria, and modeling of the formation of oligosaccharides from D-glucose with Aspergillus niger glucoamylases I and II. Biotechnology and Bioengineering 34(5):694-704.

Pandey A. 1995. Glucoamylase Research: An Overview. Starch - Stärke 47(11):439-445.

Paolucci-Jeanjean D, Belleville MP, Zakhia N, Rios GM. 2000. Kinetics of cassava starch hydrolysis with Termamyl (R) enzyme. Biotechnology and Bioengineering 68(1):71-77.

Park JT, Rollings JE. 1994. Effects of substrate branching characteristics on kinetics of enzymatic depolymerizaion of mixed linear and branched polysaccharides: I. Amylose/amylopectin α-amylolysis. Biotechnology and Bioengineering 44(7):792-800.

Park JT, Rollings JE. 1995. Effects of substrate branching characteristics on kinetics of enzymatic depolymerization of mixed linear and branched polysaccharides: II. Amylose/glycogen α-amylolysis. Biotechnology and Bioengineering 46(1):36-42.

Polakovič M, Bryjak J. 2002. Modelling of the kinetics of thermal inactivation of glucoamylase from Aspergillus niger. Journal of Molecular Catalysis B: Enzymatic 19-20(0):443-450.

Polakovič M, Bryjak J. 2004. Modelling of potato starch saccharification by an Aspergillus niger glucoamylase. Biochemical Engineering Journal 18(1):57-63.

Rodriguez VB, Alameda EJ, Gallegos JFM, Requena AR, Lopez AIG. 2006. Enzymatic hydrolysis of soluble starch with an alpha-amylase from Bacillus licheniformis. Biotechnology Progress 22(3):718-722.

Rollings JE, Thompson RW. 1984. Kinetics of enzymatic starch liquefaction: Simulation of the high-molecular-weight product distribution. Biotechnology and Bioengineering 26(12):1475-1484.

Swinkels JJM. 1985. Composition and properties of commercial native starches. Starch - Stärke 37(1):1-5.

Synowiecki J. 2007. The use of starch processing enzymes in the food industry. In: Polaina J, MacCabe A, editors. Industrial Enzymes: Springer Netherlands. p 19-34.

Torgerson EM, Brewer LC, Thoma JA. 1979. Subsite mapping of enzymes - Use of subsite map to simulate complete time course of hydrolysis of a polymeric substrate. Archives of Biochemistry and Biophysics 196(1):13-22.

Tran PL, Cha H-J, Lee J-S, Park S-H, Woo E-J, Park K-H. 2014a. Introducing transglycosylation activity in Bacillus licheniformis α-amylase by replacement of His235 with Glu. Biochemical and Biophysical Research Communications 451(4):541-547.

Tran PL, Lee JS, Park KH. 2014b. Experimental evidence for a 9-binding subsite of Bacillus licheniformis thermostable α-amylase. FEBS Letters 588(4):620-624.

van Boekel MAJS. 2008. Kinetic Modeling of Reactions In Foods: CRC Press.

van der Maarel MJEC, van der Veen B, Uitdehaag JCM, Leemhuis H, Dijkhuizen L. 2002. Properties and applications of starch-converting enzymes of the α-amylase family. Journal of Biotechnology 94(2):137-155.

van der Veen ME. 2005. Towards intensification of starch processing. PhD thesis. [PhD thesis]. Wageningen: Veen, M.E van der. 104 p.

van der Veen ME, van der Goot AJ, Boom RM. 2005. Production of glucose syrups in highly concentrated systems. Biotechnology Progress 21(2):598-602.

van der Veen ME, Veelaert S, Van der Goot AJ, Boom RM. 2006. Starch hydrolysis under low water conditions: A conceptual process design. Journal of Food Engineering 75(2):178-186.

Vidal BC, Rausch KD, Tumbleson ME, Singh V. 2009. Kinetics of granular starch hydrolysis in corn dry-grind process. Starch - Stärke 61(8):448-456.

White DR, Hudson P, Adamson JT. 2003. Dextrin characterization by high-performance anion-exchange chromatography–pulsed amperometric detection and size-exclusion chromatography–multi-angle light scattering–refractive index detection. Journal of Chromatography A 997(1-2):79-85.

Wojciechowski P, Koziol A, Noworyta A. 2001. Iteration model of starch hydrolysis by amylolytic enzymes. Biotechnology and Bioengineering 75:530 - 539.

Wong DWS. 1995. Food enzymes: structure and mechanism: Springer Science & Business Media.

Zanin GM, De Moraes FF. 1996. Modeling cassava starch saccharification with amyloglucosidase. Applied Biochemistry and Biotechnology 57-58(1):617-625.

Describing the full extent of wheat starch hydrolysis by *Bacillus licheniformis* α-amylase - Studies on a stochastic model

Abstract

A stochastic model of starch hydrolysis was extended to predict the dextrose equivalent (DE) and the complete composition of carbohydrates produced by *Bacillus licheniformis* α-amylase (BLA) during wheat starch hydrolysis. The model can predict all the products of the hydrolysis reaction provided a subsite map of the enzyme is available. The model's predictions were compared with experimental data (50°C) analysed by using HPLC-SEC. The gradual hydrolysis of starch revealed that BLA does not hydrolyse starch in a random manner.

The absolute values predicted by the model initially are different from the actual concentrations of the carbohydrates in the experimental data. Changing the value of one of the subsites improves the fit for the smaller oligosaccharides, but is not enough to account for all the differences. However, by being able to make these changes, we show that our model can be used not only for predicting carbohydrate composition, but also for verifying the accuracy of experimentally obtained subsite maps.

2.1. Introduction

Starch is one of the most abundant polysaccharides in the world. It can be easily extracted from plants and used in its native form, but it can also be transformed allowing numerous other applications. One of the methods of transforming starch is its hydrolysis. Starch is hydrolysed, e.g., in breweries or alcohol distilleries, where the produced glucose is fermented to ethanol. Since starch hydrolysis is so commonly used on industrial scale, it has been studied extensively and in a variety of manners.

Several models of starch hydrolysis have already been discussed in the literature. Most of them focus on predicting the concentrations of small carbohydrates (with the degree of polymerization (DP) of up to 10 glucose units) or the dextrose equivalent (DE). Those models, although useful, are unable to characterize the complete hydrolysis process.

The stochastic model of Besselink et al. (2008) describes starch hydrolysis by predicting both the average dextrose equivalent and the concentrations of small carbohydrates (DP < 8) during hydrolysis of wheat starch. This model can be further developed to predict the concentrations of all hydrolysis products. The model of Besselink et al. (2008) is based on the subsite theory for depolymerizing enzymes (Allen and Thoma 1976). The subsite theory provides the information about the affinities of the individual 'docking stations' in the active center of the enzyme towards the substrate. In this model the subsite affinities are combined with information on the structure and ratio of amylose to amylopectin in wheat starch. Monte Carlo simulation is then implemented in the synthesis and the hydrolysis of starch. Besselink et al. (2008) fitted their model to the experimental data obtained from starch hydrolysis conducted with α-amylase derived from *Bacillus licheniformis* (BLA).

In this chapter we report the results of wheat starch hydrolysis by *Bacillus licheniformis* α-amylase at 50°C and describe the trends that all the hydrolysis products follow over time. The data we collected are then compared with the extended hydrolysis model to predict over time the concentrations of all carbohydrates present during wheat starch hydrolysis.

2.2. Materials

Unmodified wheat starch (S5127) was purchased from Sigma-Aldrich (Steinheim, Germany). Starch samples contained 13.7 ± 0.2% of moisture, which was taken into account while preparing solutions. Bacterial thermostable α-amylase (EC 3.2.1.1) Type XII-A from *Bacillus licheniformis* was purchased from Sigma (Thermamyl 120®, a product of Novozyme Corp.). The enzyme concentration was expressed as grams of enzyme stock solution per 100 grams of total reaction mixture (w/w%). Sodium hydroxide (Merck, Germany) and calcium chloride di-hydrate (Merck, Germany) were at least analytical grade. MilliQ water was used for preparation of all the solutions. Carbohydrate standards for HPLC calibration (glucose, maltose, maltotriose, maltotetraose, maltopentaose, maltohexaose and maltoheptaose, all minimum 90% purity) and dextran analytical standards for gel permeation chromatography (5, 12, 25 and 50 kDa) were purchased from Sigma-Aldrich, Germany.

2.3. Methods

2.3.1. Starch hydrolysis

The method used for gelatinization and enzymatic starch hydrolysis was based on the method described by Besselink et al. (2008). All gelatinization and hydrolysis experiments were conducted in a temperature-controlled glass batch reactor (200 ml volume). The reactor was filled with a suspension of wheat starch in demineralized water (10 w/w%) with 5 mM $CaCl_2 \cdot 2H_2O$. This suspension was heated to 90°C at the start of the starch gelatinization process. To ensure that the solution was properly mixed, a stainless steel anchor stirrer was used (300 rpm). The pH was not adjusted (pH 5.6, measured at 50°C), but remained unchanged during the course of the reaction.

After one hour, the temperature of the gelatinized mixture was lowered to 50°C (± 1°C) before the hydrolysis reaction. When the solution reached the desired temperature, 0.01 w/w% of α-amylase from Bacillus licheniformis was added and 6 hours were allowed for hydrolysis. During the reaction hydrolysed starch was pipetted into Eppendorf tubes and directly frozen in liquid nitrogen to stop the hydrolysis. Samples were stored in a freezer (-80°C) until further analysis.

2.3.2. Determination of carbohydrate composition

The crushed frozen samples (0.3 g) were transferred into a new tube. To increase the pH value and to stop the enzyme, 90 µl of 2 M NaOH and 1,110 µl of demineralized water were added to each sample. All samples were centrifuged for 10 minutes at 2,400 × g and 4°C to separate the undissolved remains. The remaining supernatant was filtered using syringes and Minisart single use filters into glass HPLC bottles. The carbohydrate composition was measured by using size exclusion chromatography (HPLC-SEC). The Dionex Ultimate 3000 HPLC system was equipped with a Shodex Sugar KS-803 column with KS-G guard and a RI-detector. The column was operated at 80°C and used MilliQ water as eluent at a flow-rate of 0.3 ml·min^{-1}.

Table 2.1 Groups of carbohydrates and the retention times used to divide the chromatograms of hydrolysed starch samples.

Group content	Average M_w	Retention time
Glucose (DP 1)	180	32.3 - 33.5
Maltose (DP 2)	342	31.1 - 32.3
Maltotriose (DP 3)	504	30.2 - 31.1
Maltotetraose (DP 4)	666	29.7 - 30.2
DP 5 - 8	1071	28.5 - 29.7
DP 9 - 15	1962	27.5 - 28.5
DP 16 - 35	4149	26.0 - 27.5
DP 36 - 80	9414	24.5 - 26.0
DP 81 - 180	21159	23.0 - 24.5
DP 181 - 350	43029	21.8 - 23.0
DP 351 - 450	64818	21.3 - 21.8
DP 451 - 700	93168	20.5 - 21.3

Every chromatogram obtained from the analysed hydrolysis samples was divided into 12 groups. Each group contained either a single carbohydrate (if the peak was distinguishable) or a range of carbohydrates (Table 2.1). To calculate the retention times at which these groups of sugars were eluting from the column, we calibrated the HPLC system. The calibration plot was prepared based on the elution times of 11 standards of known molecular weight (M_w). Using the equation from the calibration curve and the average molecular weight of each group, we estimated elution times. The

chromatograms were then divided accordingly and the mass fractions (X_w) of carbohydrates within each interval were calculated:

$$X_w(\%) = \frac{C_{DPi}}{\sum_{j=1}^{700} C_{m\,DPj} \cdot M_{w,w} + C_0} \cdot 100 \tag{1}$$

where C_{DPi} [g·l^{-1}] is the mass-based concentration of carbohydrates with degree of polymerization i (DP_i), C_{mDPj} [mol·l^{-1}] is the molar concentration of carbohydrates with the degree of polymerization j (DP_j), C_0 [g·l^{-1}] is the initial carbohydrate concentration in the reaction mixture and $M_{w,w}$ is the molar mass of water (18.02 g·mol^{-1}).

2.3.3. Model

For each model simulation a new substrate was built. The substrate consisted of around 100,000 glucose units, distributed between amylose (26%) and amylopectin (74%). The use of Monte Carlo method provided a variation in the starch structure. The positions of the branches along the amylopectin molecule were randomized in the model and based on probability, preventing molecules from being identical in every simulation. The subsite map of Kandra et al. (2006) for BLA was used unchanged, just as the inhibition constants (b_I = 0.1; b_{II} = 0.2; b_{III} = 0.4) proposed by Marchal et al. (2003). Matlab 2012a (7.14) was used to perform all of the simulations.

A number of modifications was made compared to the model of Besselink et al. (2008). First, the model was extended to quantify all carbohydrates present in the hydrolysed matrix; the carbohydrates with the same numbers of glucose units, both branched and linear, were added together using their overall mass fractions (X_i):

$$X_i = \frac{n_i \cdot (i \cdot M_{w,g} + M_{w,w})}{m_{tot}} \tag{2}$$

Where i is the degree of polymerization of the molecule for which the X_i is calculated, n_i is the number of moles of i in the matrix, $M_{w,g}$ is the molar weight of a glucose unit (162,14 g·mol^{-1}), $M_{w,w}$ is the molar weight of water (18,02 g·mol^{-1}) and m_{tot} is the total mass of the starch matrix in grams, including the mass of water added after each hydrolysis event. The carbohydrates were then grouped based on their degree of polymerization into 12 groups corresponding to the groups of the collected experimental data.

The second modification enhanced the time resolution (timespan of the model). The model by Besselink et al. (2008) was limited to only 8 time steps of the set timespan in order to save computation time and data storage space. We used a larger number of time steps: 1001 time steps were stored out of around 3 million discrete hydrolysis events. At each of the time steps, all molecules present in the hydrolysed substrate matrix were quantified and stored. This higher resolution ensured a better comparison of the model with the experimental results.

Finally, we changed the fitting procedure in the model. Comparing the model to the data required translating the model time (i.e. discreet hydrolysis events) into physical time values. In this model the physical time was calculated by fitting the modelled DE values to the experimental DE. Besselink et al. (2008) used only the linear part of the curve of their experimental DE data to fit the model DE to the corresponding experimental values. We decided that it was more accurate to use all the data points from the experiments for the fitting. Each model simulation was performed five times and then averaged, before the fitting procedure took place. We always used the full set of the experimental DE values for the non-linear fitting of the model DE. The experimental DE values were calculated from the mass fractions of the oligosaccharides (DP 1-6) as described by Baks et al. (2007), while the DE values from the model were calculated as described by Besselink et al. (2008).

2.4. Results and Discussion

2.4.1. Starch hydrolysis

The chromatograms of all carbohydrates as shown in Figure 2.1 depict different stages of enzymatic hydrolysis of wheat starch by Bacillus licheniformis α-amylase (BLA) at 50°C. The five chromatograms present the progress of starch hydrolysis after 0, 15, 60, 90, 180 and 360 minutes. A clear order in which the starch molecules were hydrolysed can be observed. In the beginning of the hydrolysis, just after the enzyme was added (time 0), hardly any carbohydrates present in the sample were smaller than DP 150 (below M_w of 25,000). The measurements correspond with the sizes known for wheat amylose (number-average DP of around 1200, spread over a wider range of DP 200-3,000 (Hanashiro and Takeda 1998)) and amylopectin (molecular weight in the range of 10^7 (Manners 1989)). It is not possible to distinguish between amylose and

amylopectin molecules in this type of analysis. The first peak at time 0 (retention time 16-19 minutes) represents the largest and the entangled soluble molecules that elute the fastest. Based on the observations of Park and Rollings (1994) these molecules correspond to high molecular weight amylopectin. The tall peak at time 0 (retention time 21 minutes) contained carbohydrates with apparent M_w of 40,000 to more than 100,000, built from roughly 200 to 650 glucose units based on the calibration.
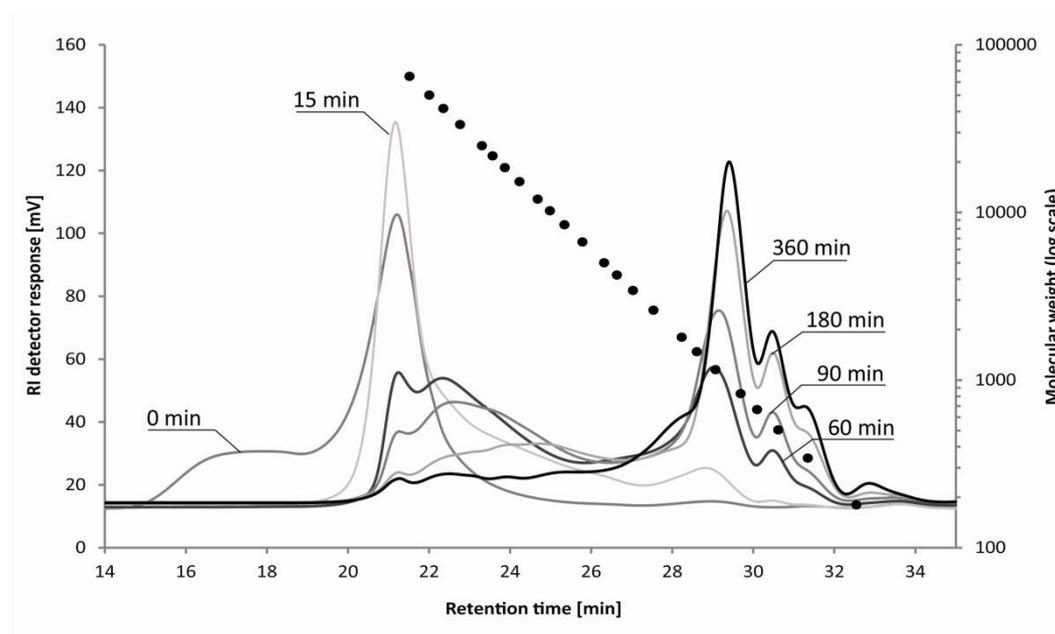

Figure 2.1. The chromatograms of wheat starch samples hydrolysed by BLA (50°C) at six different times t [minutes] during the reaction. The dotted plot represents the logarithm of molecular weight at the indicated retention times.

Hydrolysis of starch by α-amylases is not entirely random and the product distribution can be, at least partially, explained based on the highly ordered structure of amylopectin and the structure of the enzyme. The high molecular weight amylopectin (peak below the RT of 20 minutes) was degraded during the reaction's first 15 minutes and simultaneously the concentrations of all molecules larger than maltotetraose (DP4) started to increase. The majority of the products after the first 15 minutes of reaction were within the range of the large molecules (DP 250-900, average apparent M_w of 70,000), as a result of the attack of the enzyme on the accessible linear fragments of high molecular weight amylopectin. The α-amylase has a higher affinity for these linear fragments because they are more exposed (Park and Rollings 1994). The products of hydrolysis of the high molecular weight amylopectin will then be groups of clusters (domains) or single clusters. Simultaneously with amylopectin also amylose molecules

were digested. At this early stage of the reaction (15 min), the hydrolysis of amylose can be the source of the intermediate size carbohydrates (DP 40-150), since the molecular weight of amylose rapidly shifts to lower values due to the more random endo-action of α-amylase.

Despite the higher affinity of the enzyme for amylopectin, the rate of hydrolysis of amylose is overall higher (Park et al. 1988). This difference in the hydrolysis rate is a result of two factors, the first being the simpler structure of amylose, with less branching. Branch points and clusters of branches slow down the initially high reaction rate of amylopectin hydrolysis. The second factor is the different spatial conformation of the two polymers in solution. Park and Rollings (1994) reported that, at least in water-DMSO solutions, amylose conformation is that of a helix, whereas amylopectin forms a random coil. The enzyme can bind the exposed fragments of the more flexible amylopectin molecules easier, than the rigid amylose molecules. This leads to the higher preference of the enzyme towards hydrolysis of amylopectin. Additionally, the affinity of the enzyme for linear regions of amylopectin is based not so much on the structure of the polymer, but rather on the structure of the enzyme itself (Bertoft 2013). The reaction proceeded faster when all subsites of an enzyme are filled with glucose residues, which happens easier in the longer, linear regions of amylopectin.

Our observations for wheat starch and BLA are in line with those of Park and Rollings (1994), who observed chromatograms similar to ours after analysing a 50-50% mixture of isolated amylose and amylopectin degraded by BLA. Their figures presented a wide, tailing peak at the beginning of the reaction followed by a tall heading peak few minutes later. In the hydrolysis of waxy maize amylopectin by Bacillus subtilis α-amylase similar stages of hydrolysis were observed (Bertoft 1986).

In the following stages of the reaction, up until the 180th minute hydrolysis time, the enzymatic attack was mostly directed at the carbohydrates comprising the largest peak (at 21 minutes retention time). While the amount of these large carbohydrates continued to decrease, a group of intermediate size (shoulder at 22-23 minute) started to appear. An attack on the linear segments of amylopectin would quickly reduce the size of the molecules and produce smaller units - a cluster domain or a cluster.

Two main theories about the amylopectin structure are being debated in literature, the traditional cluster model (Hizukuri 1986) and the building block and the

backbone concept (Bertoft 2013). These theories, along with the research on chain length distributions of amylopectin enable estimating the sizes of clusters.
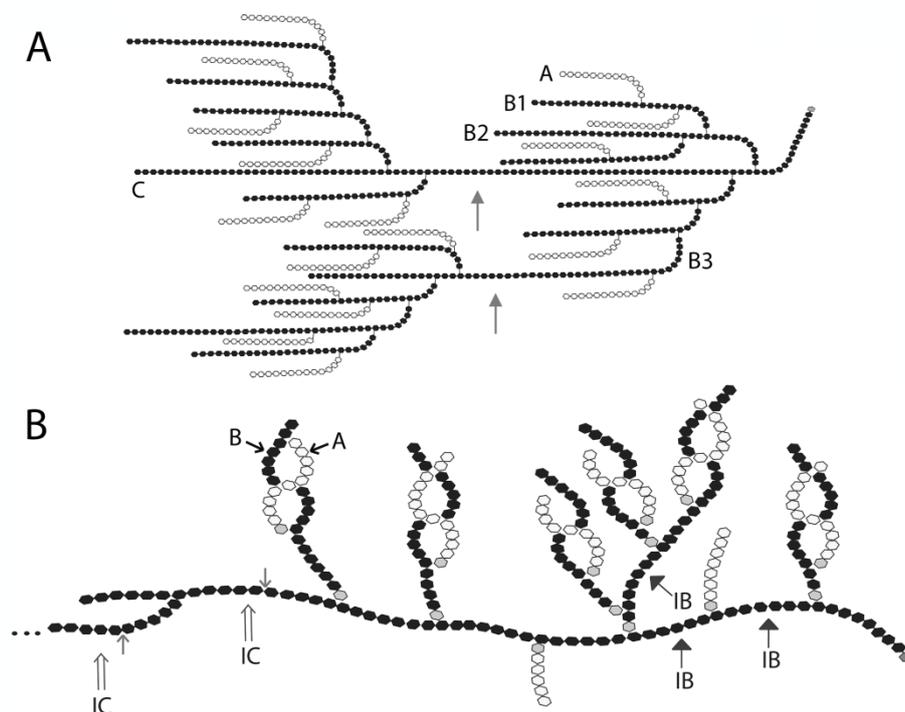

Figure 2.2. Structure of amylopectin fragments according to two models described in literature. (a) Traditional model with clusters. Three clusters are shown along with two possible points of the initial enzymatic attack (arrows). The sizes of the clusters are between DP of 220 and 270. A chains are presented in white, an exemplary B1, B2 and B3 chains and the C chain are indicated. The reducing end is shown in dark grey on the right-hand side of the figure. (b) Model described by Bertoft (2013). Only one cluster is shown (DP of 243 or 260, depending on the point of hydrolysis). A chains are shown in white, the first glucose unit of each branch is depicted in grey and the reducing end is shown in dark grey on the right-hand side of the figure. The IB are the inter-block segments, the double-arrows point out the inter-cluster (IC) segments. The small arrows show the possible point of enzymatic attack.

Cluster model. Based on the traditional model (Figure 2.2 A), Takeda and Hanashiro (2003) found that wheat amylopectin clusters contain on average 13 chains. The short chains (A and B1) build up the clusters, whereas the long chains (B2 and B3) connect them. When combining the information on the length of the isolated side chains of wheat amylopectin, the amount of A chains per B1 chain, and the mol% of the A, B and C chains (Hanashiro et al. (2002); Hizukuri and Maehara (1990)) one can calculate the size of a cluster. Based on this literature data, an average cluster in terms of degree of polymerization would be equal to approximately 240 glucose units (M_w of ~40,000). When taking into account the data collected by Hanashiro et al. (1996), a cluster of 13 chains would contain around 320 glucose units (M_w around 53,000). Of course, values chosen here are average and only used as an approximation. The produced clusters vary

in size, depending on the number of chains present per isolated cluster and the length of the chains.

Backbone model. In the second model, the amylopectin unit corresponding to a cluster in the traditional model has a different structure (Figure 2.2 B). The model described by Bertoft (2013) involves a backbone with small, branched units named building blocks. The building blocks are the smallest units of amylopectin (contrary to the clusters in the traditional model) that can be isolated from 'clusters' by extensive hydrolysis. For cereals, Bertoft and co-workers describe large clusters (on average 12 chains per cluster), with fewer long chains and a higher proportion of small chains, containing six to eight building blocks that are within a short distance from each other. The building blocks are occasionally separated by longer segments between the blocks – inter-cluster segments of more than nine glucose residues. These segments are thought to be the parts of the amylopectin molecule attacked first by the α-amylase to release the 'clusters'. We can make an estimation of the size of an exemplary cluster, based on the assumptions of Bertoft (2013) and the side-chain lengths following Hanashiro et al. (1996). Depending on the point of cleavage (e.g. indicated by one of the two arrows in Figure 2.2 B) and the number and lengths of side-chains, an average cluster could have the DP in the range from 240 (M_w ~39,000) to 260 (M_w ~42,000).

Regardless of the theory used and the average literature values of chain lengths, we arrive at the same explanation – the peak at roughly 23 minutes constitutes the clusters that are slowly degraded in the next stages of hydrolysis. This means that both of the models can be used to explain the hydrolysis profiles we obtained, therefore we do not exclude either of them.

The preference of the enzyme for the longer fragments of amylopectin not carrying branches, explains why the molecules constituting the tall peak (21 min) are being degraded rather slowly and through molecules that form the wider peak at 22-23 minutes (molecular weight ~25,000 to 50,000; DP ~150-300). The glycosidic bonds located outside of clusters, are easier accessible to the enzyme and are cleaved first. Therefore, in the first, most rapid stage of hydrolysis the domains containing various numbers of clusters are produced. That stage is followed by still quite fast hydrolysis of the domains into isolated clusters. After the clusters are separated, the bonds within the cluster become the main target of the enzyme. That reaction progresses slowly because

of the more spatial structure of the cluster, formed by the large number of branches. The linkages between the glucose molecules inside the cluster are less accessible than in the regions linking clusters. The structure of the substrate and the preference of the enzyme explain why the hydrolysis of amylopectin into clusters is a faster process than cleaving the shorter chains within the clusters (Bertoft 2013; Nielsen et al. 2008). Therefore, the hydrolysis of bonds in amylopectin cannot be described as entirely random.

The distribution of the smaller hydrolysis products throughout the hydrolysis is not entirely random either. Our data show clearly that the concentrations of carbohydrates of intermediate sizes (approximately DP 40 – 150 eluting between 24 and 27 minutes) remain at a steady level – they do not increase or decrease from 30th to 180-240th minute. We expect that this is the result of the non-random hydrolysis based on the structure of amylopectin clusters. The oligosaccharides are initially the product of the hydrolysis of amylose and the easier accessible outer chains of the clusters (Bertoft 1986). Eventually, also the inner chains of amylopectin clusters are degraded into oligosaccharides. The large enzyme molecules cannot easily penetrate the highly branched regions of clusters, this explains the lower rate of hydrolysis of the chains inside the highly branched clusters (Park and Rollings (1994); Park et al. (1988)). Based on the average chain lengths of the A and B chains of amylopectin, one can then expect more carbohydrates of smaller sizes, shorter than the A and B chains. Since the rate of hydrolysis of the α-1,6 bonds by BLA is low, the A and B chains will only be hydrolysed a few bonds away from the branch point. The clear pattern in the proportions of the oligosaccharides is a result of the characteristics of the enzyme (i.e. subsite map).

When the reaction reached 180 minutes, nearly all carbohydrates larger than DP 300 were digested. From that moment on only the concentrations of carbohydrates with around 16 or less glucose units continued to increase. With the largest molecules gone, the intermediate size carbohydrates of DP 30 to 300 were slowly hydrolysed. The slow hydrolysis of these molecules might suggest they are the branched remains that are left after cutting of the side-chains (A and B chains) of the clusters.

2.4.2. *Model development*

The model designed by Besselink et al. (2008) was used to predict the concentrations of the small carbohydrates (DP 1-7) and the dextrose equivalent over time. During the development of that model, the authors decided to focus on only small

carbohydrates for two reasons. First, running the model for all carbohydrates was computationally intensive and second, because they did not analyse the larger carbohydrates in their experiments. Due to increase in computational power available, the first was no longer a limitation, and thus we focused on the latter.

The model of Besselink et al. (2008) could be used to perform the calculations for carbohydrates of all sizes, predict their concentrations and even distinguish between the branched and the linear carbohydrates. Unfortunately, the currently available methods of analysis of starch do not provide such detailed separation. The methods that have a resolution of 1 glucose unit fail to analyse molecules larger than DP of 40-50 (e.g., capillary electrophoresis). Other methods, which can analyse the whole range of molecular weights, such as the HPSEC chosen here, struggle with lower resolution.

To partially compensate for the poor resolution of HPSEC the measured carbohydrates were divided into groups (Table 2.1). These groups are based on the shapes and trends in the chromatograms. The smallest carbohydrates (DP 1-4) were considered individually, while the larger ones were grouped together based on the retention times from the calibration. The groups were given strict DP borders, but one should bear in mind that this is only an approximation, especially for the larger carbohydrates.

The hydrolysis experiments were repeated four times. The differences between the four data sets and the model were minor (an average square root of the residual sum of squares (SRSS) of 639.5 with a 0.25% difference between the highest and the lowest SRSS value), therefore we randomly chose to depict only one of the experimental data sets in the graphs of this chapter.

After applying all of the initial modifications to the model, the first new version of the model was used to make predictions for one of the experimental data sets (Figure 2.3) using the same initial values of parameters as Besselink et al. (2008). As presented in Figure 2.3 A the fit of the modelled DE to the DE calculated from the experimental data is good; better than that of Besselink et al. (2008). This better fit is the result of lower DE values in our results (less extensive hydrolysis, resulting in lower concentrations of the oligosaccharides) and a more accurate fitting procedure.
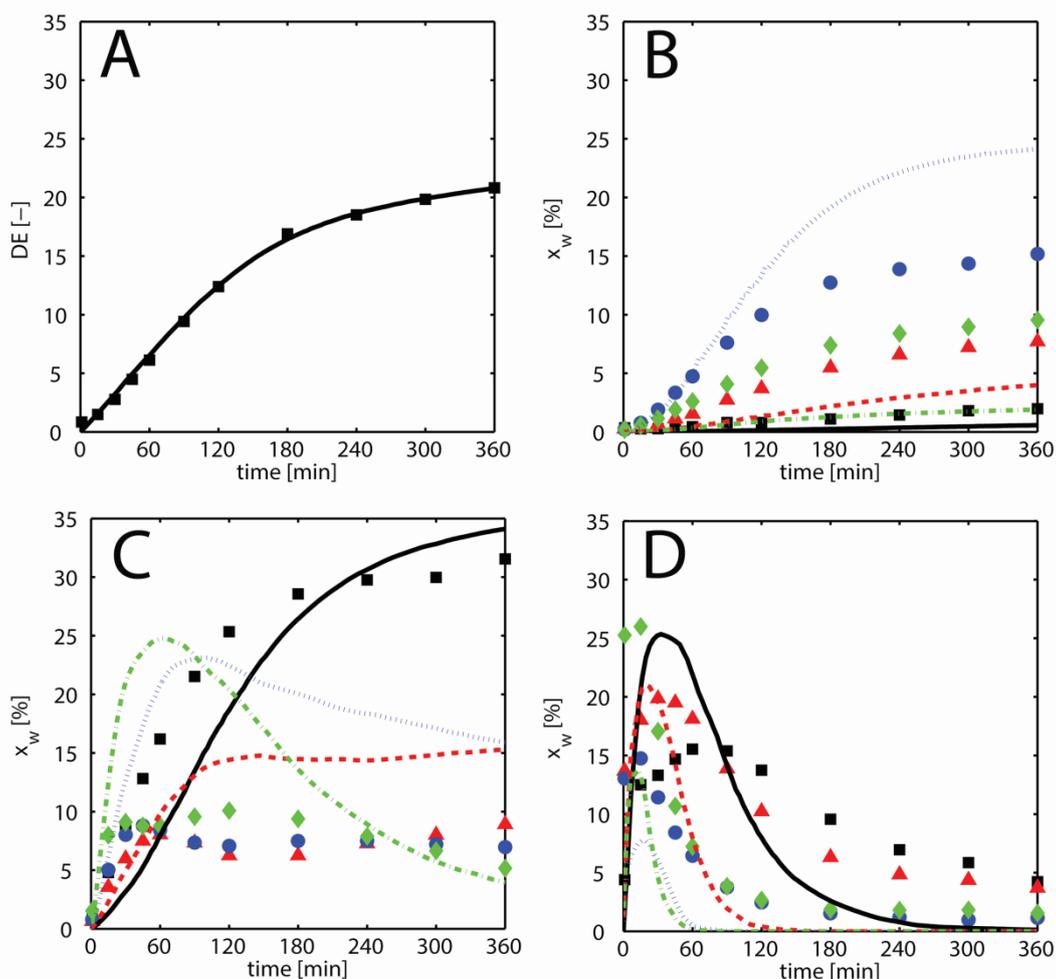
Figure 2.3 The model representing starch hydrolysis in time (lines) fitted to the experimental values (points). (A) Dextrose equivalent; (B) Glucose (black squares/solid line), maltose (red triangles/dashed line), maltotriose (blue dots/dotted line) and maltotetraose (green diamonds/dash-dot line); (C) Carbohydrate groups (by DP): 5-8 (black squares/solid line), 9-15 (red triangles/dashed line), 16-35 (blue dots/dotted line) and 36-80 (green diamonds/dash-dot line); (D) Carbohydrate groups (by DP): 81-180 (black squares/solid line), 181-350 (red triangles/dashed line), 351-450 (blue dots/dotted line) and 451-700 (green diamonds/dash-dot line).

Despite the good fit of the DE, the predictions of the concentrations of carbohydrate groups were less accurate. The shapes of the curves representing the model showed qualitative similarities to the trends in the experimental data, but the actual values of concentrations were either over- or underestimated. A clear example was maltotriose, for which after 360 minutes of hydrolysis the model predicted a concentration of nearly 25 w/w%, which was higher by 10 w/w% than in the experimental data.

The model also overestimated the production of molecules belonging to the groups DP 9-15, DP 16-35, DP 36-80 (Figure 2.3 C) and DP 81-180 (Figure 2.3 D). In the

beginning of the hydrolysis, the concentrations of these groups increased too rapidly, which would suggest that the hydrolysis in the model was not as restricted as in the actual reaction. Simultaneously the largest molecules in the model (Figure 2.3 D) were hydrolysed too fast. The concentrations of groups DP 351-450 and DP 451-700 were not only largely underestimated in the model, but their values also declined too rapidly. Similarly, groups DP 81-180 and DP 181-350 of the model were also hydrolysed too fast. In the experiments these groups were steadily hydrolysed over the whole duration of the reaction, whereas in the model they were no longer present after 120 (DP 181-350) or 240 minutes (DP 81-180). The fast hydrolysis of the largest molecules (Figure 2.3 D) lead to overestimating the concentrations of molecules of intermediate size (Figure 2.3 C). Thus, the model might not have had enough constraints set for the enzyme, and lack of those boundaries allowed for all the bonds to quickly be cleaved.

2.4.3. Subsite map changes

One of the causes of the inaccuracies in the predictions may be the assumed values of the subsite map. The differences in the binding energy of the subsites within a subsite map of an enzyme are responsible for the slightly different product distributions after enzymatic starch hydrolysis (Macgregor et al. 1994). Besselink et al. (2008) already pointed out that the subsite map developed by Kandra et al. (2006) had its shortcomings.

First, the enzyme can interact differently with the substrate used to create the subsite map than it would with the starch molecules. The presence of an aromatic ring (2-chloro-4-nitrophenyl) at the end of the substrate-analogue molecules could have influenced the interaction between the enzyme and the substrate. The pNPG aromatic ring can act as a glucose unit and thus interacts with the subsites of the enzyme. It was indeed shown to interact more favourably with some subsites, and less favourably with others (Macgregor et al. 1994).

Second, the analogue molecules Kandra et al. (2006) used were much smaller than a true starch molecule. Using linear maltodecaose, an analogue which does not reflect the complex structure of starch amylopectin, could have been overly simplistic and might have caused the calculated energy values to differ from the actual energy values for starch. Differences in the values of binding energies can cause significant discrepancies when they are used for predicting the yields of starch hydrolysis products.

Besselink et al. (2008) already observed that the model seemed to under- or overestimate the production of certain small carbohydrates. A change in the energy value of the subsite map (subsite +3 was changed from -5.8 to 0 kJ·mol^{-1}) resulted in an improved fit of the model. To verify the findings of Besselink et al. (2008) we also changed the value of subsite +3 to 0 kJ·mol^{-1} (Figure 2.4).
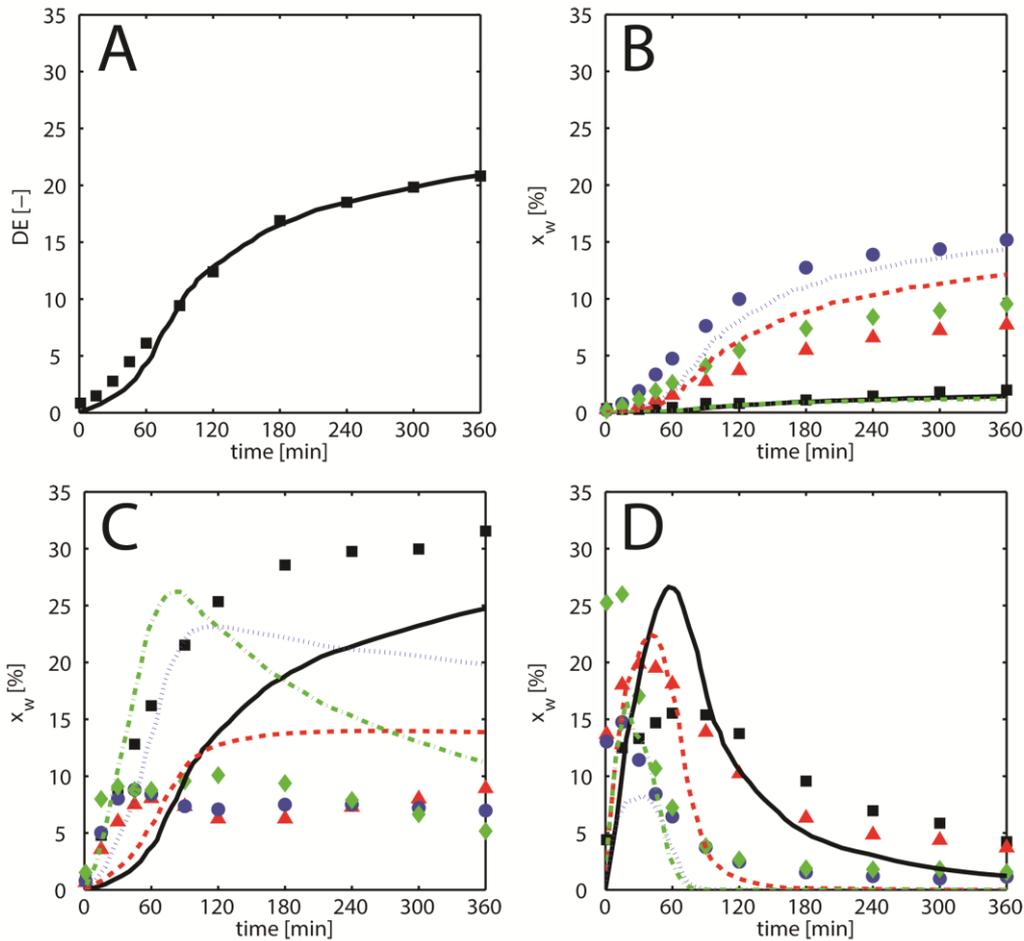

Figure 2.4 The model representing starch hydrolysis in time (lines) fitted to the experimental values (points). The value of subsite +3 was changed to 0 kJ·mol-1 in the model. (A) Dextrose equivalent; (B) Glucose (black squares/solid line), maltose (red triangles/dashed line), maltotriose (blue dots/dotted line) and maltotetraose (green diamonds/dash-dot line); (C) Carbohydrate groups (by DP): 5-8 (black squares/solid line), 9-15 (red triangles/dashed line), 16-35 (blue dots/dotted line) and 36-80 (green diamonds/dash-dot line); (D) Carbohydrate groups (by DP): 81-180 (black squares/solid line), 181-350 (red triangles/dashed line), 351-450 (blue dots/dotted line) and 451-700 (green diamonds/dash-dot line).

As a result of removing the influence of subsite +3, the fit of maltotriose and glucose generally improved (Figure 2.4 B), however the predictions for other carbohydrate groups remained the same or deviated even more from experiments as compared with the original model. The predicted concentration of e.g. maltose (Figure

2.4 B) increased nearly three-fold, and was almost two times higher than in our experiments. The change in the value of subsite +3 did not influence maltotetraose, but caused some deviation in the DP 5-8 group. The remaining groups of carbohydrates seemed to have been barely influenced by the change made to the subsite map.

When Besselink et al. (2008) changed the energy of subsite +3 to 0 they showed an improvement in the fit for glucose and DP 2, 3 and 5, as well as for the DE. The same action in our case improved the fit for only glucose and maltotriose. Our experimental analyses showed slightly different concentrations of carbohydrates compared to those found by Besselink et al. (2008). This discrepancy was most probably caused by the differences in the analysis method (e.g., different HPLC column).

Table 2.2 The residuals (SRSS) calculated for the model with unchanged subsite map and the model with the value of subsite +3 changed in the subsite map.

Group of carbohydrates (DP)	SRSS of the unchanged model	SRSS of the changed subsite +3
1	7.7	3.7
2	21.6	21.3
3	39.0	16.4
4	41.5	44.6
5-8	48.6	91.7
9-15	47.3	48.9
16-35	101.2	99.2
36-80	80.6	89.6
81-180	70.7	54.6
181-350	70.9	58.8
351-450	46.0	36.5
451-700	77.1	57.7
Total	**652.1**	**623.1**

Aside from the visual comparison of model and data, we used the residuals to put a value on how the changes in the model affect the predictions. These residuals were calculated as the square root of the sum of squares of residuals (SRSS) per group of carbohydrates and as a total for each model version (Table 2.2). Small differences in the residuals between the two versions of the model can be the effect of the Monte Carlo method that introduces randomness into both the synthesis and the hydrolysis in the model. That is especially true for the high molecular weight molecules, where a variation

within around 15% is normal. It thus seems that the concentrations of the large molecules are less dependent on the interactions within the catalytic site. More pronounced differences, such as for maltotriose and DP 5-8, are the effect of the change made to the subsite +3. Comparing the values of the residuals per group makes it easier to draw conclusions about the improvement of the fit, since some groups show clear improvement of the fit, while the fit of other groups becomes less accurate. Overall, the values of the SRSS confirm what can be observed in the plots of data versus the model when the value of subsite +3 is changed to 0: the improvement of fit for glucose and maltotriose, and the decline of fit for maltotetraose and DP 5-8.

Analysing the values of the individual subsites might bring more insight on how the energy values of subsite maps influence the hydrolysis products. Small changes made to the values of the subsite map in our model might show how detailed the calculations of subsite energy should be. Changing the values of subsite map to 0, might on the other hand help in confirming the theoretical deliberations about which groups of carbohydrates will be affected by a particular subsite. With a model like ours, where the subsite map values are used for making predictions about carbohydrate concentrations, a subsite map can be tested for its sensitivity and accuracy.

2.5. Conclusions

The products of wheat starch hydrolysis by *Bacillus licheniformis* α-amylase at 50°C show a pattern. This proves the enzyme does not attack the starch molecules randomly. This pattern can be explained based on the theories describing the structure of amylopectin and the subsite theory. The conclusions from our experiments can be used to further improve the structure of the substrate in the model.

An extended version of the model as proposed by Besselink et al. (2008), describes the hydrolysis of starch (both amylopectin and amylose) as a function of time. In contrast to these authors, we include not only the small oligosaccharides, but nearly all carbohydrates present in the reaction.

While the comparison of experimental values with the original model shows an acceptable fit, there is still a discrepancy: larger carbohydrates hydrolyse too fast, while smaller ones accumulate too fast. This may be due to inaccurate values of the subsite map. We showed that a change of the subsites will specifically modify the formation of

smaller carbohydrates, but did not yet explain the faster hydrolysis of the larger chains. Therefore, the model still needs to be adapted in different ways to make the prediction of the hydrolysis of the larger carbohydrates more accurate.

2.6. Acknowledgements

2.7. References

Allen JD, Thoma JA. 1976. Subsite mapping of enzymes - application of depolymerase computer model to two alpha-amylases. Biochemical Journal 159(1):121-131.

Baks T, Ngene IS, van Soest JJG, Janssen AEM, Boom RM. 2007. Comparison of methods to determine the degree of gelatinisation for both high and low starch concentrations. Carbohydrate Polymers 67(4):481-490.

Bertoft E. 1986. Hydrolysis of amylopectin by the alpha-amylase of B. subtilis. Carbohydr Res 149(2):379-387.

Bertoft E. 2013. On the building block and backbone concepts of amylopectin structure. Cereal Chemistry Journal 90(4):294-311.

Besselink T, Baks T, Janssen AE, Boom RM. 2008. A stochastic model for predicting dextrose equivalent and saccharide composition during hydrolysis of starch by alpha-amylase. Biotechnology and Bioengineering 100(4):684-97.

Hanashiro I, Abe J-i, Hizukuri S. 1996. A periodic distribution of the chain length of amylopectin as revealed by high-performance anion-exchange chromatography. Carbohydrate Research 283(0):151-159.

Hanashiro I, Tagawa M, Shibahara S, Iwata K, Takeda Y. 2002. Examination of molar-based distribution of A, B and C chains of amylopectin by fluorescent labeling with 2-aminopyridine. Carbohydrate Research 337(13):1211-1215.

Hanashiro I, Takeda Y. 1998. Examination of number-average degree of polymerization and molar-based distribution of amylose by fluorescent labeling with 2-aminopyridine. Carbohydrate Research 306(3):421-426.

Hizukuri S. 1986. Polymodal distribution of the chain lengths of amylopectins, and its significance. Carbohydrate Research 147(2):342-347.

Hizukuri S, Maehara Y. 1990. Fine structure of wheat amylopectin: the mode of A to B chain binding. Carbohydrate Research 206(1):145-159.

Kandra L, Remenyik J, Gyemant G, Liptak A. 2006. Effect of temperature on subsite map of Bacillus licheniformis alpha-amylase. Acta Biol Hung 57(3):367-75.

Macgregor EA, Macgregor AW, Macri LJ, Morgan JE. 1994. Models for the Action of Barley Alpha-Amylase Isozymes on Linear Substrates. Carbohydr Res 257(2):249-268.

Manners DJ. 1989. Recent developments in our understanding of amylopectin structure. Carbohydrate Polymers 11(2):87-112.

Marchal LM, Ulijn RV, De Gooijer CD, Franke GT, Tramper J. 2003. Monte Carlo simulation of the alpha-amylolysis of amylopectin potato starch. 2. alpha-amylolysis of amylopectin. Bioprocess and Biosystems Engineering 26(2):123-32.

Nielsen MM, Seo ES, Dilokpimol A, Andersen J, Abou Hachem M, Naested H, Willemoës M, Bozonnet S, Kandra L, Gyémánt G and others. 2008. Roles of multiple surface sites, long substrate binding clefts, and carbohydrate binding modules in the action of amylolytic enzymes on polysaccharide substrates. Biocatalysis and Biotransformation 26(1-2):59-67.

Park JT, Rollings JE. 1994. Effects of substrate branching characteristics on kinetics of enzymatic depolymerizaion of mixed linear and branched polysaccharides: I. Amylose/amylopectin α-amylolysis. Biotechnology and Bioengineering 44(7):792-800.

Park JT, Yu L-P, Rollings JE. 1988. Substrate structural effects on enzymatic depolymerization of amylose, amylopectin, and glycogen. Annals of the New York Academy of Sciences 542(1):53-60.

Takeda Y, Hanashiro I. 2003. Examination of the structure of amylose and amylopectin by fluorescent labeling of the reducing terminal. Journal of Applied Glycoscience 50(2):163-166.

An improved stochastic model for prediction of enzymatic starch hydrolysis

Abstract

All of the starch hydrolysis products can be predicted using a stochastic model provided the subsite map of the enzyme used for hydrolysis is known. A subsite map reported in literature for *Bacillus licheniformis* α-amylase at 50°C was tested in the model, resulting in a poor fit between the model and the experimental data. Therefore, the apparent binding energy values of each subsite in the subsite map were varied. The effect of these changes on the composition of the hydrolysis products was reported, along with a proposal of a new subsite map. The changes in the subsite map merely influenced the formation of the small oligosaccharides (DP 1-8). The predictions of the larger products also needed adjustments, therefore we investigated another element of the model – the inhibition. Our analyses show that rather high inhibition is necessary at 50°C to slow down the reaction in the model enough to reflect the experimental data. With a simple model we are able to near quantitatively predict the composition of wheat starch hydrolysis products.

3.1. Introduction

During enzymatic starch hydrolysis a complex mixture of products is being formed, which has a profound influence on the ultimate product. Modelling of the hydrolysis process can be helpful in anticipating the product composition and can even replace a costly empirical analysis. With polymers like starch, describing the hydrolysis reaction using merely a simple Michaelis-Menten equation is not sufficient. Detailed modelling becomes possible when the mechanism of the catalysis and the action pattern of the enzyme are known. If a subsite map, which is a characteristic trait of an enzyme, is known, hydrolysis products of a particular enzyme can be predicted (Suganuma et al. (1978), Torgerson et al. (1979), Kondo et al. (1980)).

Most of the available models describing starch hydrolysis focus only on a fraction of the hydrolysis products (Presečki et al. (2013), Murthy et al. (2011), Besselink et al. (2008), Marchal et al. (2003), Wojciechowski et al. (2001), Åkerberg et al. (2000), Paolucci-Jeanjean et al. (2000), Torgerson et al. (1979)). To our knowledge the only model that simultaneously studies all the separate products and substrates of the starch hydrolysis reaction and predicts the complete product composition and dextrose equivalent (DE) of starch liquefied by *Bacillus licheniformis* α-amylase (BLA) over time, is that described in chapter 2.

The predictions of that model did not yet provide satisfactory results for neither the small, nor the large carbohydrates. The concentrations of the individual groups of carbohydrates were predicted qualitatively, but not quantitatively (chapter 2). The largest products were predicted to be hydrolysed much faster than was seen in the experiments.

The differences in the concentrations of glucose and the oligosaccharides were most likely influenced by the energy values in the subsite map that was used. This subsite map (Kandra et al. 2006) was developed using only small substrate analogues (DP 10) and thus its use for starch is in fact an extrapolation. We expect that the subsite map containing nine subsites will especially influence the concentrations of molecules that contain up to five glucose units.

On the other hand, the fast disappearance of larger molecules led us to believe that additional inhibition or constraints are needed in the model to slow down this rapid hydrolysis. What is more, we could not exclude the possibility that the discrepancies in

the concentrations of the largest molecules might be attributed to the structure and chain length distribution (CLD) of the substrate that was used in the model.

In this chapter we investigate the different elements constituting the model for a better description of the experimental data by the extended model of starch hydrolysis. Our aim was to develop a new subsite map, based on the hydrolysis of starch, instead of the hydrolysis of substrate analogues that have relatively low molecular weight. We also investigated the influence of the inhibition caused by the proximity of a branch in amylopectin molecules. The inhibition caused by a branch point occurs when the model chooses a glucose unit joined by an α-1,6-glycosidic bond or a glucose unit in close proximity of a branch point. The search for the best fitting subsite map was coupled with the description of the effect of inhibition in the model on the products of the reaction and the rate of their hydrolysis.

3.2. Materials

Unmodified wheat starch (S5127) with 13.68 ± 0.24% of moisture was purchased from Sigma-Aldrich (Steinheim, Germany). Bacterial thermostable α-amylase (EC 3.2.1.1) Type XII-A from *Bacillus licheniformis* was purchased from Sigma (Thermamyl 120®, a product of Novozyme Corp.). Sodium hydroxide and calcium chloride di-hydrate (Merck, Germany) were at least analytical grade. MilliQ water was used for preparation of all the solutions. Carbohydrate standards for HPLC calibration (glucose, maltose, maltotriose, maltotetraose, maltopentaose, maltohexaose and maltoheptaose, all minimum 90% purity) and dextran analytical standards for gel permeation chromatography (5, 12, 25 and 50 kDa) were purchased from Sigma-Aldrich, Germany.

3.3. Methods

3.3.1. Starch hydrolysis

The method used for gelatinization and enzymatic starch hydrolysis was as described in chapter 2. A glass batch reactor (200 ml volume) was filled with a suspension of wheat starch in demineralized water (10 w/w%) with 5 mM $CaCl_2 \cdot 2H_2O$. The suspension was heated to 90°C and mixed (300 rpm) to gelatinize the starch. After one hour, the temperature of the gelatinized mixture was lowered to 50°C (± 1°C) and 0.01 w/w% of α-amylase from *Bacillus licheniformis* was added. During the reaction, 1.5

ml samples of hydrolysed starch mixture were pipetted into Eppendorf tubes and directly frozen in liquid nitrogen to stop the hydrolysis.

3.3.2. *Determination of the carbohydrate composition*

To 0.3 g of the frozen starch samples, 90 µl of 2 M NaOH and 1,110 µl of demineralized water were added. The samples were then dissolved and centrifuged for 10 minutes at 2,400 g and 4°C and the remaining supernatant was filtered using syringes and Minisart single use filters.

The carbohydrate composition was analysed by using size exclusion chromatography (HPLC-SEC). A Dionex Ultimate 3000 HPLC system was equipped with a Shodex Sugar KS-803 column with KS-G guard and an RI-detector. The column was operated at 80°C and used MilliQ water as eluent at a flow-rate of 0.3 ml·min^{-1}.

Using the equation from the calibration curve prepared with standards of known molecular weight and the average molecular weight of each group of carbohydrates, we estimated the elution times. The mass fractions of the carbohydrates were calculated as described in chapter 2.

The hydrolysis experiments were performed four times and the resulting data were compared with the model. We randomly chose one set of data for all of the analyses in this chapter, as the differences between the data sets were small and insignificant (data not shown).

3.3.3. *Model*

The model was used with the earlier described adaptations (chapter 2). For each model simulation a new substrate matrix was built. The experimental DE values were calculated from the mass fractions of the small carbohydrates (DP 1-6) as described by Baks et al. (2007), while the DE values from the model were calculated as described by Besselink et al. (2008). All presented model outputs were the average values based on five repetitions with the same set of parameters. Matlab 2012a (7.14) was used to perform all simulations.

Two inherent parts of the model were analysed in detail: the subsite map and the inhibition.

3.3.3.1. Improvement of the subsite map

The subsite theory used in the model describes the catalytic site of amylases as composed of several subsites, each interacting with one glucose residue of the substrate. The energy of the interaction between the enzyme and the substrate at each site can be quantified (Hiromi (1970), Thoma et al. (1970), Kandra et al. (2006)) and used to follow the hydrolysis of the substrate into various products over time (Allen and Thoma (1976), Suganuma et al. (1978)).

In our model the subsite map is used for calculating the association constants ($K_{r,n}$) for each binding mode of the enzyme-substrate complex (Figure 3.1), both productive (resulting in bond cleavage) and non-productive (not resulting in bond cleavage). The association constant for a particular complex is based on the sum of the binding energies of the subsites occupied by the substrate. The chance of hydrolysis (p) is equal to 1 when the sum of the binding energies has the lowest possible value, resulting in the highest value of the association constant (K_{max}). For more details about the model we refer to Besselink et al. (2008).

During the simulations the energy values of each subsite were first changed to 0, thereby removing the effect of the subsite completely. This change demonstrated which subsites affected the predicted concentrations of which carbohydrates. The energies assigned to the subsites were also varied within a specific range, to simplify determining of the apparent binding energy of each subsite. The relative quality of the fit between the prediction and the experiment was quantified using the square root of the sum of squares (SRSS) of the differences between the model and data.
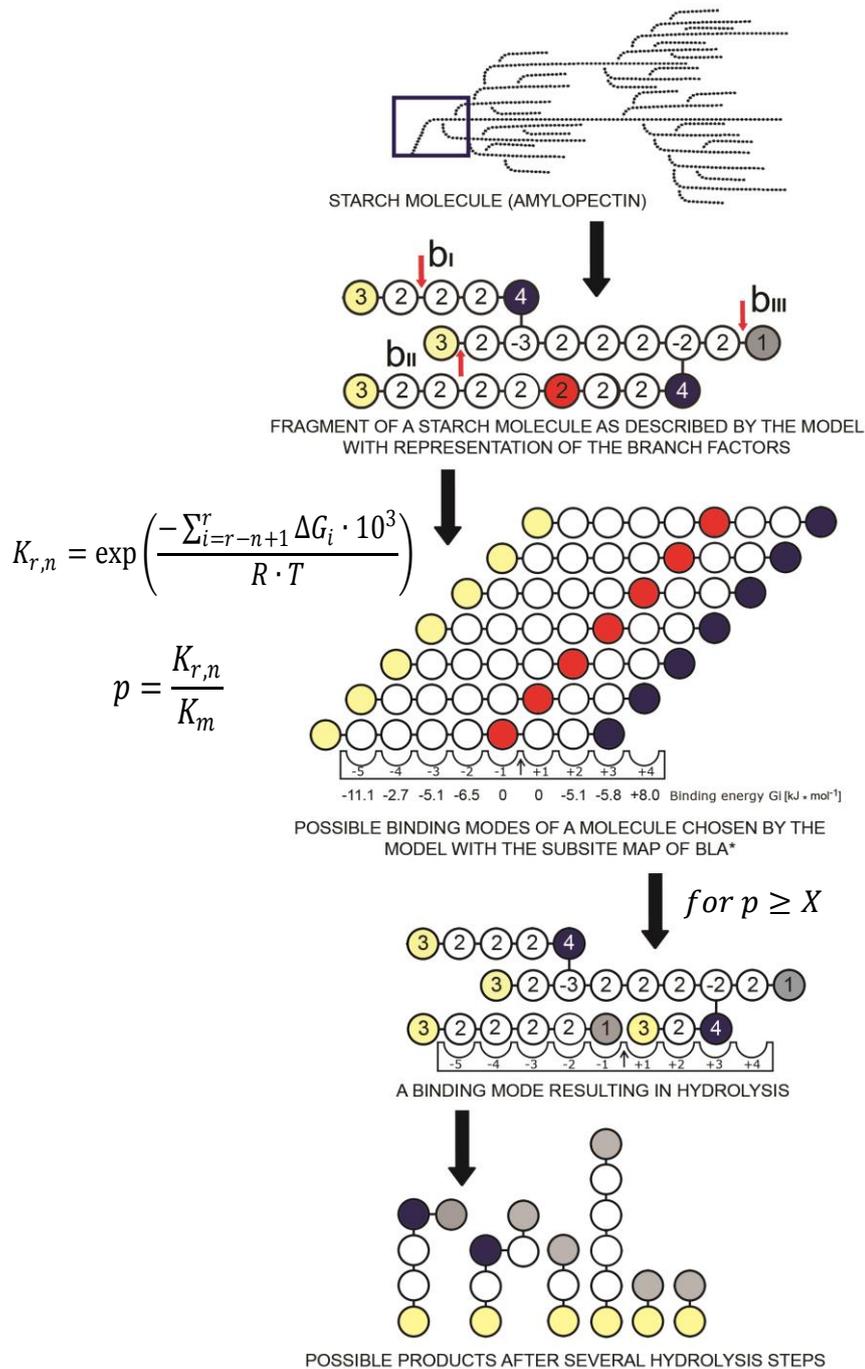
STARCH MOLECULE (AMYLOPECTIN)

FRAGMENT OF A STARCH MOLECULE AS DESCRIBED BY THE MODEL
WITH REPRESENTATION OF THE BRANCH FACTORS

$$K_{r,n} = \exp\left(\frac{-\sum_{i=r-n+1}^{r} \Delta G_i \cdot 10^3}{R \cdot T}\right)$$

$$p = \frac{K_{r,n}}{K_m}$$

-11.1 -2.7 -5.1 -6.5 0   0 -5.1 -5.8 +8.0 Binding energy GI [kJ • mol⁻¹]

POSSIBLE BINDING MODES OF A MOLECULE CHOSEN BY THE
MODEL WITH THE SUBSITE MAP OF BLA*

$for\ p \geq X$

A BINDING MODE RESULTING IN HYDROLYSIS

POSSIBLE PRODUCTS AFTER SEVERAL HYDROLYSIS STEPS

Figure 3.1 Schematic representation of the main elements of the hydrolysis model. $K_{r,n}$ is the association constant, ΔG_i is the binding energy [kJ·mol⁻¹] of subsite i, r is the subsite occupied by the reducing end in the subsite map, n is the degree of polymerization of the substrate, p is the chance of hydrolysis, K_{max} is the association constant with the lowest binding energy, X is a random number. The fragment of the starch molecule is used to both show the examples of bonds affected by three branch factors used in the model inhibition and the numbering of glucose units in the model. *Subsite map of Bacillus licheniformis α-amylase (BLA) at 50°C following Kandra et al. (2006).

3.3.3.2. Improvement of the model inhibition

In the discussed model inhibition is based on the inability of the enzyme to hydrolyse α-1,6-glycosidic bonds (branch points), which makes the α-1,4-glycosidic

bonds around these branch points less accessible for the enzymes to hydrolyse. If the model chooses a glucose unit that is affected by the proximity of an α-1,6-glycosidic bond, the chance of hydrolysis (p) of that unit's α-1,4 bond is multiplied by $(1 - k_{br\_in})$. The inhibition factor (k_{br_in}) is calculated with the following equation:

$$k_{br\_in} = 2 - e^{b_T y}$$

with b_T a branch factor that indicates the type of inhibition (location of the bond to be cleaved in the starch molecule relative to the branch point, as shown in Figure 3.1) and y is the number of glucose units between the cleavage site and the branch point (Marchal et al. 2003). The equation describes the decrease of the inhibition as the distance from the branch point increases. An inhibition factor (k_{br_in}) of 0 or below 0 is equivalent to no inhibition, whereas a value of 1 would indicate complete inhibition and impossibility of hydrolysis (Marchal et al. 2003).

Marchal et al. (2003) introduced three branch factors b_I, b_{II} and b_{III}, each corresponding to an α-1,4-bonds located in a different position with respect to the branch point (Figure 3.1). When the substrate is bound by the enzyme, the α-1,6 linked glucose unit can be located either on the non-reducing end side (b_{III}) or on the reducing end side of the catalytic site ('before' or 'after' the cleavage site). If the latter is true, the enzymatic attack can take place either at the chain containing the α-1,4,6-linked glucose unit (b_{II}) or at the chain with the α-1,6-linked glucose unit (b_I) – the branch. Besselink et al. (2008) used the values $b_I = 0.1$, $b_{II} = 0.2$ and $b_{III} = 0.4$, which influenced the hydrolysis only when the branch point was less than 6, 3 and 1 glucose units away from the cleavage site, respectively. To demonstrate to what extent the predictions are affected by this type of inhibition, the inhibition was either removed from the model or stronger as compared to the original values chosen by Besselink et al. (2008) and Marchal et al. (2003).

3.4. Results and discussion

3.4.1. Inhibition

The inhibition around branches was implemented in the model of Besselink et al. (2008), although its accuracy was not tested or discussed by the authors. We first used the same set of branch factors (0.1, 0.2 and 0.4) as Besselink et al. (2008) and Marchal et al. (2003). Figure 3.2 A-D shows that the model qualitatively resembled the trends of the

experimental data, but there was considerable quantitative discrepancy, as is also indicated by the large square root of the residual sum of squares (SRSS = 652).

We verified the extent of the influence of the branch factors (b_T) on the predictions of our model by removing the inhibition completely from the model or by changing the values of all inhibition factors to 0.1.

Removing the inhibition from the model resulted in a worse prediction of the concentrations of all groups of carbohydrates above DP 8 (Figure 3.2 E-H). Thus the inhibition mostly influences the hydrolysis rate of the larger starch molecules and has less effect on the concentrations of the mono-, di- and oligosaccharides (DP 1-8). Removal of the inhibition leads to the largest polysaccharides (Figure 3.2 H) being entirely hydrolysed within 90 minutes of the modelled reaction. The products of the intermediate size (Figure 3.2 G, except for DP 5-8) are then overestimated by nearly a factor of three at the early stages, and follow a different trend than the experimental data do. DP 1-4 are almost unaffected by the removal of inhibition, while DP 5-8 are only moderately influenced, mostly because of the change in concentrations of DP 6-8. These observations confirm that without any inhibition, the model cannot predict the experimental results correctly.

Changing the values of all the branch factors to 0.1 improves the predictions (Figure 3.2 I-L), with a significant decrease of the total SRSS from 652 (for the original set of factors) to 542. The lower value of the branch factors implies inhibition over larger distances from a branch point. With a b_T value of 0.1, glucose units up to 6 bonds away from the branch point were affected by the inhibition. Again, the smallest carbohydrates (Figure 3.2 J) were mostly unaffected by the change of the inhibition factors, but improvements were visible for all the groups containing larger molecules (Figure 3.2 K-L).

Overall the model predictions were especially better for the concentrations of molecules with DP from 9 to 350. It is reasonable that larger molecules are more strongly affected by this type of inhibition, as they contain more branches. The hydrolysis of linear molecules is not influenced by the inhibition factors. The predictions of the model with inhibition now better resemble the experimental data, which implies that the inhibition constraints are essential in the model.
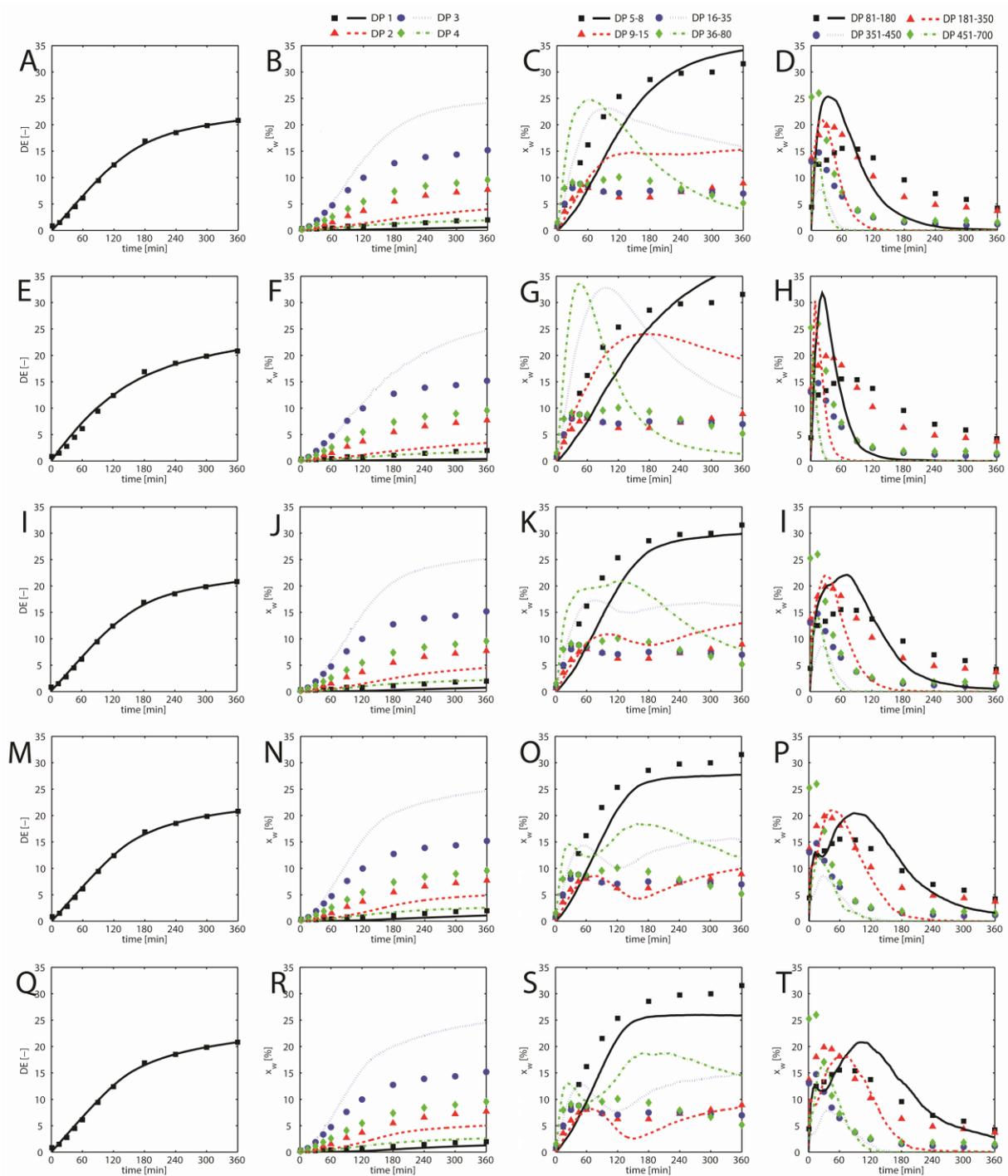
Figure 3.2. Predictions of the concentrations of hydrolysis products with varied amount of inhibition. (A-D) the original model (b_I = 0.1, b_{II} = 0.2, b_{III} = 0.4; SRSS 652); (E-H) no inhibition (SRSS 863); (I-L) all branch factors set to 0.1 (SRSS 542); (M-P) all branch factors set to 0.05 (SRSS 434); (Q-T) all branch factors set to 0.03 (SRSS 413). First graph in the row depicts the dextrose equivalent (DE), the remaining show the carbohydrates as described by the legend on top of each column.

Next, we optimized the branch factors, without distinguishing between them. The improvement of the predictions was verified by comparing the graphs (Figure 3.2) and the residuals (Table 3.1) of the original model with those using the new sets of branch

factors. The following set of branch factors: b_I = 0.03, b_{II} = 0.03 and b_{III} = 0.03 was chosen as optimal based on the values of the SRSS.

However, there are only minor differences between b_T = 0.03 and the one where all branch factors were set to 0.02, 0.04 or 0.05. When comparing the graphs visually, the differences between the predictions using these values of branch factors were not as pronounced, even if the SRSS values differed. As shown in Table 3.1, if all b_T were equal to 0.03 or equal to 0.05 the divergence in the SRSS was mostly caused by a worse fit of DP 16-35, 36-80 and 451-700 and a better fit for 181-350 of the b_T 0.05 and only these differences were noticeable in the graphs.

Table 3.1 The differences between the model and the data (represented as SRSS) caused by changes in the inhibition due to the variation of branch factors. The average of 5 computations is presented for each change. The inhibition factors are either removed ('Without inhibition'), taken from literature or all of them are set to one value.

DP group	Branch factors						
	No inhibition	0.1 0.2 0.4	0.1	0.05	0.04	0.03	0.02
DP_1	8.3	7.7	7.3	6.2	5.8	5.5	4.8
DP_2	23.8	21.6	19.5	17.1	16.8	15.8	15.3
DP_3	31.5	39.0	46.1	47.0	48.3	45.9	47.2
DP_4	42.6	41.5	40.4	38.6	38.6	38.5	38.1
DP_{5-8}	58.4	48.6	43.1	43.8	44.7	47.9	49.3
DP_{9-15}	98.2	47.3	25.2	7.9	8.0	11.0	15.0
DP_{16-35}	138.2	101.2	73.7	52.4	46.4	39.6	37.3
DP_{36-80}	114.5	80.6	77.5	61.1	58.3	55.6	48.8
DP_{81-180}	96.0	70.7	48.9	26.7	24.7	24.5	31.1
$DP_{181-350}$	106.7	70.9	52.1	40.0	42.5	46.4	50.2
$DP_{351-450}$	56.2	46.0	39.8	33.9	31.4	32.6	33.3
$DP_{451-700}$	88.2	77.1	68.3	59.5	51.0	49.8	50.7
Total SRSS	862.5	652.1	541.8	434.3	416.4	413.2	421.1

The results show that a rather high inhibition around the branches is necessary, for a reasonable prediction of the concentrations of the large molecules. The effect of inhibition would not be as visible if only the small hydrolysis products (DP < 8) were analysed and modelled, as it was the case in the study by Besselink et al. (2008).

Additionally, we altered the values of only one factor while keeping the remaining two constant. When the value of b_I was lowered while the others remained at 0.1, the improvement of the fit based on SRSS was attained by better predictions of all

carbohydrates above DP 5 (data not shown). When b_{II} was lowered, no difference was observed, but lowering b_{III} led to a lower total SRSS mostly due to a better fit of DP groups 9-15 and 16-35 (data not shown).

Branch factors b_I and b_{III} thus have the biggest impact on the outcomes of the model. Branch factor b_I influences the hydrolysis of the bonds in the outer branches of amylopectin. Decreasing the values of inhibition factors b_{II} and b_{III} increases the inhibition of cleaving the α-(1,4) bonds located in the same chain as the α-(1,4,6) linked glucose unit, with the branched glucose unit on the reducing end side (b_{II}) or the non-reducing end (b_{III}) of the cleavage site.

It is reasonable that the inhibition for the inner chains (B and C chains) of the amylopectin molecule should be larger, since the glycosidic bonds in those chains are less accessible to the enzyme than the bonds of the outer branches (the A chains). Both factors b_I and b_{III} are relevant in the binding modes affecting the outer parts of the amylopectin molecule. As the outer chains of the substrate are more accessible to the enzyme, the binding influenced by inhibition constants b_I and b_{III} will be more likely to occur, therefore these two factors will have a larger influence on the outputs.

The description of the inhibition used in our model solely acts on the ability of the enzyme to hydrolyse the α-1,6-glycosidic bonds and the α-1,4-glycosidic bonds adjacent to the branches. In the actual hydrolysis experiments more factors affect the performance of the enzyme. Our model does not take into account transglycosylation and condensation reactions, enzyme inactivation, product inhibition or multiple attack. Each of these factors may still help improving the fit of the model, albeit at the cost of significantly increasing the number of parameters of the model. Since these factors are not included in the current model, the increased inhibition caused by the presence of branches might partially compensate for these other effects.

3.4.2. Subsite map

The bond cleavage frequencies used for the calculations of the subsite map are determined using analogous substrates, at low substrate concentrations and only at the early stages of hydrolysis. This approach allows avoiding other reactions, e.g. transglycosylation, condensation - mechanisms that can affect product distribution (Suganuma et al. (1978); Kandra et al. (2002); Kandra et al. (2006)).

In our experiments, due to the different substrate and time scale, we cannot exclude that these mechanisms influenced the final results. The differences between our analyses and those of Kandra et al. (2006) may stem from us following the reaction over a longer period of time. Thus, our data include the effects of the subsequent enzymatic attacks. Our model is meant to predict the results of the actual enzymatic hydrolysis experiments, on a natural substrate - starch. As we demonstrated before (chapter 2), the subsite map in its current state (Kandra et al. (2002); Kandra et al. (2006), also given in Figure 3.1) does not accurately describe the carbohydrate profiles obtained in our starch hydrolysis experiments. Therefore, we followed a step-wise analysis of the subsite map energy values and of their influence on the composition of the hydrolysis products.

Analysing the binding energy values assigned to subsites gives insight into which hydrolysis products are favoured. Macgregor et al. (1994) pointed out that the hydrolysis of substrates long enough to cover all subsites is mainly determined by the two outer subsites on each side (-5, -4 and +3, +4). The large positive value of free energy at subsite +4 (8.0 kJ·mol^{-1}) suggests an unfavourable interaction between a glucose unit and that subsite. A repulsion at +4 causes the amylase to show preference for binding a reducing-end glucose residue at subsite +3 instead of at subsite +4. Hydrolysis of molecules bound in that way would cause more maltotriose to be formed from the reducing end side, and that is indeed what we observed in the concentrations of DP 3 predicted by the model using the original subsite map values – nearly 10 w/w% more maltotriose at 360 min compared to the experimental values.

Assigning a 0 kJ·mol^{-1} binding energy value to each subsite illustrates the effect of that subsite on the product composition. Combining the graphical outputs (Figure 3.3) with the SRSS values (Table 3.2) allows us to draw conclusions on the role of a particular subsite in over- or underestimating the predicted carbohydrate concentrations. The changes in subsites +1 and -1 are unnecessary, because for a bond cleavage to occur these subsites always have to be occupied. The yields of the carbohydrates are independent of binding energies of these subsites as long as their values remain constant.

Removing the effect of subsite +4 in the model by changing its value from +8 kJ·mol^{-1} to 0, influences the predicted concentrations of all of the oligosaccharides. The concentrations of maltotriose and maltotetraose become similar (Figure 3.3 C). For both

components the predicted concentrations correspond better with the experimental ones than those predicted with the unchanged subsite map.

Table 3.2 Differences expressed as SRSS between the data and the predictions (average of 5 computations) for each subsite individually changed to 0 and the original subsite map. The branch factors (b_T) were set to 0.05 for all the computations. One subsite was changed at a time. The extreme values of SRSS for each carbohydrate group are shown in bold. Underlined values are more than 15% higher or lower than in the original model, indicating the change is caused entirely by the change in the energy value of the subsite.

DP group	Subsite changed to 0							Original subsite map
	-5	-4	-3	-2	+2	+3	+4	
DP1	<u>**7.5**</u>	6.5	6.5	6.1	5.9	<u>**1.8**</u>	5.7	6.2
DP2	17.8	16.8	**16.0**	16.8	<u>22.5</u>	<u>**27.8**</u>	<u>22.2</u>	17.1
DP3	<u>**80.6**</u>	52.7	50.0	50.7	<u>36.1</u>	<u>**6.2**</u>	<u>33.7</u>	47.0
DP4	<u>**6.6**</u>	39.4	39.4	39.3	40.7	<u>**43.0**</u>	<u>12.3</u>	38.6
DP5-8	<u>**143.3**</u>	50.0	48.5	**41.9**	<u>52.7</u>	<u>83.5</u>	<u>75.9</u>	43.8
DP9-15	<u>17.0</u>	8.4	**8.0**	8.3	<u>12.3</u>	<u>18.1</u>	<u>**40.8**</u>	7.9
DP16-35	<u>**38.0**</u>	51.4	51.6	53.5	46.1	<u>39.4</u>	**59.5**	52.4
DP36-80	67.4	63.5	61.6	57.0	<u>72.9</u>	<u>**77.1**</u>	<u>**26.7**</u>	61.1
DP81-180	<u>41.5</u>	28.2	29.7	**27.9**	28.8	<u>44.5</u>	<u>**45.6**</u>	26.7
DP181-350	35.5	39.5	42.4	39.2	38.2	<u>**31.8**</u>	<u>**60.0**</u>	40.0
DP351-450	33.8	33.7	34.7	**31.5**	32.6	33.6	**38.2**	33.9
DP451-700	**51.0**	53.9	51.2	55.1	52.1	51.2	**62.3**	59.5
Total	540.0	443.9	439.5	**427.2**	440.9	458.0	482.9	434.3

When the energy of subsite +4 is removed, it becomes equally possible for a reducing end to bind at subsite +4 and at subsite +3, as the total energy of binding for both of the binding modes is the same. This causes a more equal distribution of DP 3 and 4. The concentrations of maltose and glucose are nearly unaffected by the change in subsite +4. On the other hand, the amounts of predicted DP 5-8 increase quite substantially - after 90 minutes the concentration of each of these carbohydrates is almost 5 w/w% higher, as compared with the output using the original value of subsite +4 (shown in the Figure A3 in the Appendix).
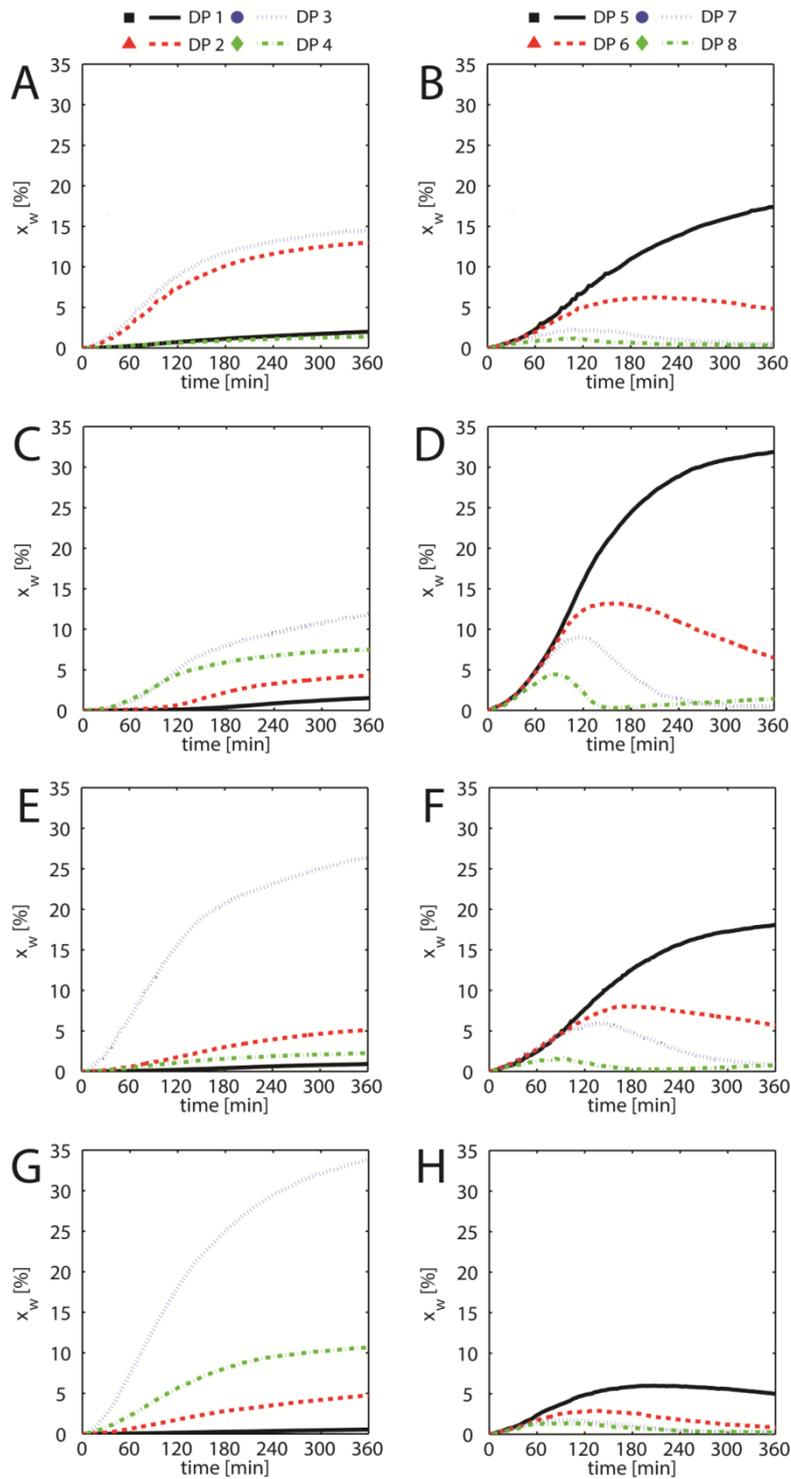
Figure 3.3. The influence of changing the binding energy of the subsites to 0 on the model predictions of DP 1-8. For each pair of graphs the binding energy of one subsite was changed to 0: (AB) subsite +3 to 0; (CD) subsite +4 to 0; (EF) subsite -4 to 0; (GH) subsite -5 to 0.

With the lack of the boundary at subsite +4, long molecules have more freedom in binding with the enzyme. These molecules will be free to overlap the whole subsite map in a different way. There will no longer be a preference of binding the reducing end at

subsite +3, and thus a lower amount of smaller molecules will be produced, since the hydrolysis events that produce molecules of larger sizes (>DP3) will be more common. The lack of the barrier at subsite +4 results therefore in a completely different pattern of products.

When the energy of subsite +3 is reduced from -5.8 to 0 kJ·mol^{-1}, the binding of a reducing end glucose at subsite +2 becomes more favorable. The mode of binding in which subsites from +2 to -5 are occupied will now be favored, which causes an increase in the concentration of maltose (Figure 3.3 A). However, the binding that includes subsite +3, now 0 kJ·mol^{-1}, will result in the same amount of total energy of binding as the binding mode that does not include subsite +3. This is why changing the energy of subsite +3 leads to similar concentrations of predicted maltose and maltotriose. The "lack" of subsite +3 causes only small changes in the concentrations of DP 5 and 6, but visibly decreases the concentration of DP 7.

Subsite -5 on the other end of the subsite map was given a high negative binding energy value (-11.1 kJ·mol^{-1}). Because of that, the non-reducing end glucose moiety will preferably bind with the subsite -5 and a large amount of DP 5 is expected. Also the quantities of DP 6 and 7 were expected to be quite high. In the model, changing the value of subsite -5 to 0 kJ·mol^{-1} significantly decreased the concentrations of DP 5, 6 and 7, and lead to a much higher production of DP 4 and DP 3 (Figure 3.3 G and H). The increase of the concentration of DP 4, was a result of the increased amount of binding over subsites -4 to +3 due to the low total energy of that interaction. Since subsite -4 had a rather low negative value assigned to it, the total energy of binding over subsites -3 to +3 would also become more common. This form of binding would further increase the production of maltotriose, at the expense of DP 5 and higher.

For subsites -2, -3 and -4 a change to 0 kJ·mol^{-1} had no effect on any of the modelled carbohydrates. The joined effects of these subsites (not shown) were tested in groups of two (-2 and -3, -3 and -4, -2 and -4) or all three simultaneously (-2, -3, -4). As long as the remaining subsites were assigned their original values, changing the values of energies of subsites -2, -3 and -4 to 0 gave minimal or no improvement of the fit between the model and the data as compared to the original subsite map used. Subsite -5 had a large negative binding energy that caused preference for binding modes including this subsite and diminished the effects subsites -2, -3 and -4 had on the

predictions. However, if the value of subsite -5 became less negative, the energy value of subsite -4 would become more relevant and would affect the product composition.

Changing the value of subsite +2 to 0 (Figure A1 in the Appendix) did not have a major effect on the concentrations of the carbohydrates, except for increasing the production of glucose in the model. When the energies of subsites -2, -3, -4 and +2 were changed to 0, only small differences in the concentrations of glucose and DP 5, 6 and 7 were observed (not shown). The outer subsites of this endo-acting amylase clearly impact the product profiles more than the inner subsites. This illustrated what also Macgregor et al. (1994) suggested - for the subsite map of BLA at 50°C only the outer subsites (-5, +3, +4) determine the distribution of the hydrolysis products.

Our model uses Monte Carlo simulation to increase the randomness of the events and by that simulates the random choice of the bond to be hydrolysed in the reaction. Because of that, variation of up to 15% in the SRSS's of the high molecular weight molecules is expected within the repetitions using the same set of parameters. Taking this into account, we assume that the difference in SRSS of more than 15% signifies an effect solely caused by the changes made to the model, e.g., different energy value of a subsite.

A number of subsites (-5, +4 and +3) specifically affect the predictions of the carbohydrates larger than DP 8. These subsites caused dramatic changes in the carbohydrate composition that in turn affect the value of the DE. Thus, by influencing the DE values predicted by the model, changes made to these subsites can influence the fitting procedure. The differences in the SRSS of the large carbohydrates will not be directly affected by the change in the subsite energy, but rather by the fitting of the model to the data, influenced by the change in the DE value.

Aside from changing the subsites individually we also tested the sensitivity of the subsite map. Many subsite maps are calculated within an accuracy of 0.1 kJ·mol^{-1}. To assess whether that is necessary we rounded the values of the subsite map to integers. The model with the integer subsite map resulted in no other distribution of carbohydrates, showing that small changes are not enough to affect the model predictions.

3.4.3. New parameters of the model

Finally, we combined the results of all optimizations in the subsite map (Table 3.3). Most of the small carbohydrates were affected by the changes made to three or more subsites. As an example, the concentration of DP 4 predicted by the model may be increased by either reducing the energy of subsite +4 (lowering the energy from +8 kJ·mol^{-1} towards 0) or increasing that of subsite -5. The concentration of DP4 predicted by the model, may be increased by increasing the value of subsite +3 (from -5.8 kJ·mol^{-1} towards 0). Maltooctaose (DP 8) was the only carbohydrate that was affected by one subsite (+4), but it was also influenced by the inhibition. A combination of properly chosen parameters was thus necessary for a correct prediction.

Table 3.3 The effect of subsites on the concentration of a particular DP. Subsites that cause a difference (an increase or a decrease) in the concentration of a particular saccharide are shown. Subsites in bold have a more pronounced effect.

Oligosaccharide	Subsite number	
	Subsites that increase concentration	Subsites that decrease concentration
DP1	**+2**	**-5**
DP2	**+3**	**+2**
DP3	**-5**	**+3 +4** +2
DP4	**+4 -5**	**+3**
DP5	**+4**	**-5** +3 -3 -4
DP6	**+4** +2 -3	**-5** +3
DP7	**+4** +2	**-5 +3**
DP8	**+4**	

Based on the rules we described, we started to look for a subsite map that could better describe the data we collected from our experiments. We varied the energy values of subsites -5, +3 and +4 in a wide range because these subsites had a major influence on the outputs. The graphs in Figure 3.4 show how changing the energy of a subsite can influence the model's predictions. The energy values of subsite +4 were varied from -1 to 10 kJ·mol^{-1}, while the energies of the other subsites remained unchanged (Figure 3.4 A). At high and low binding energy values the fit for maltotriose (DP3) and DP 5-8 became worse. Low binding energy of subsite +4 had a positive influence on the fit of DP 4, whereas at high energy values DP 9-15 fitted better. As the binding energy values
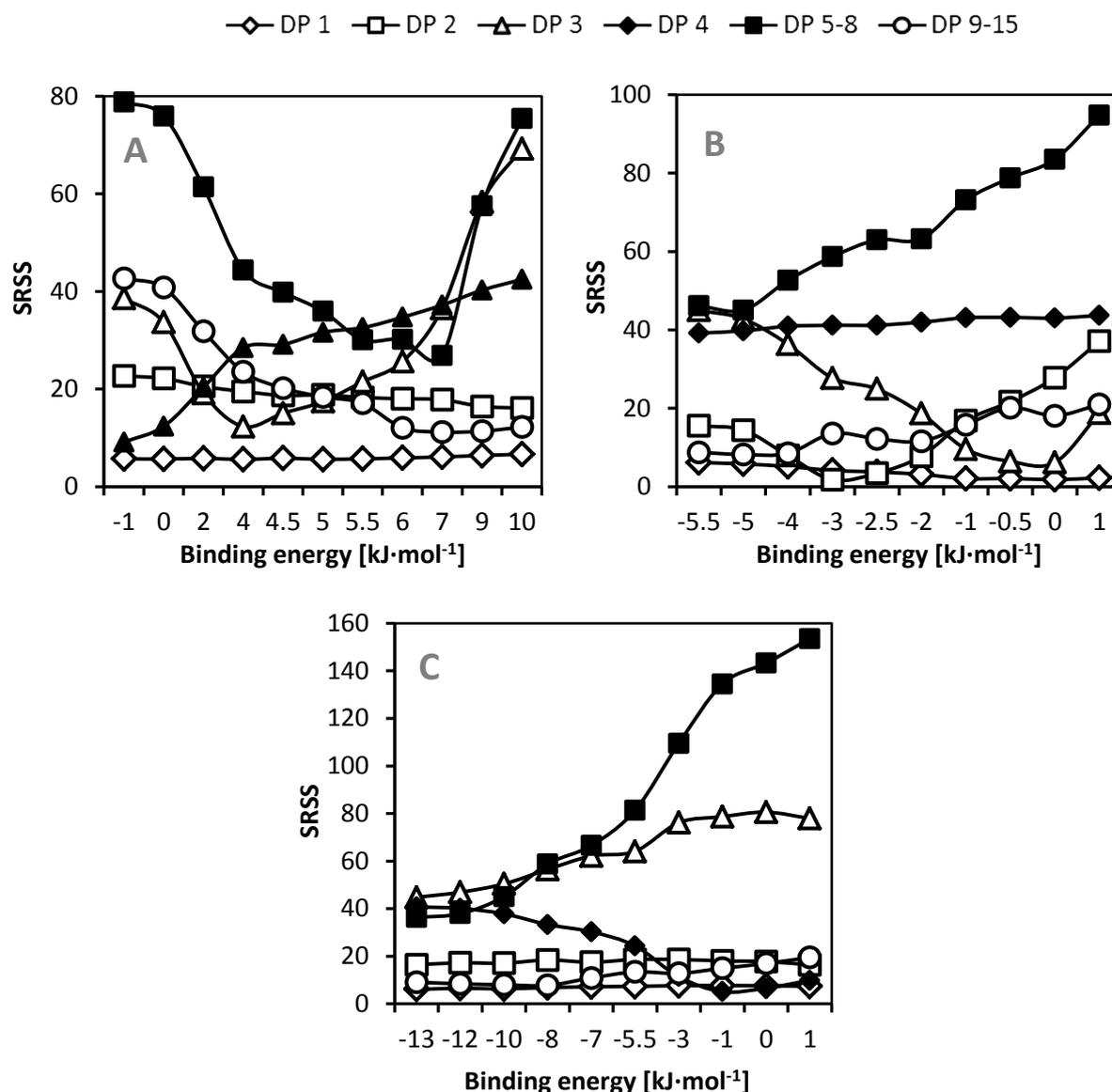
Figure 3.4. The changes in SRSS values of the different carbohydrates groups (DP 1-15) as the binding energies of subsites +4, +3 and -5 are varied in a range of values. (A) Energy of subsite +4 varied from -1 to 10 kJ·mol^{-1}; (B) Energy of subsite +3 varied from -5.5 to +1 kJ·mol^{-1}; (C) Energy of subsite -5 varied from -13 to +1 kJ·mol^{-1}. The remaining subsites were assigned the original energy values from the subsite map at 50°C (Kandra et al. 2006). The lower the SRSS values were, the more accurate predictions were obtained. The legend to all the plots is given above the figure, the lines in the plots were added to guide the eye. Note the differences in the scale between the plots (x and y axes).

approached 5 and 6 kJ·mol^{-1} the fit improved as shown by the overall lower SRSS values. Similar steps have been taken for subsites -5 (Figure 3.4 C) and +3 (Figure 3.4 B). Our observations led us to believe that subsite -5 should be assigned a relatively high negative value, as energies from -8 to -13 kJ·mol^{-1} lead to a better overall fit. DP 4 was the only carbohydrate that showed an opposite trend, fitting less accurately as the value

of subsite -5 became more negative. As for subsite +3, we concluded it should be assigned a small negative value of binding energy (-2 or -3 kJ·mol^{-1}).
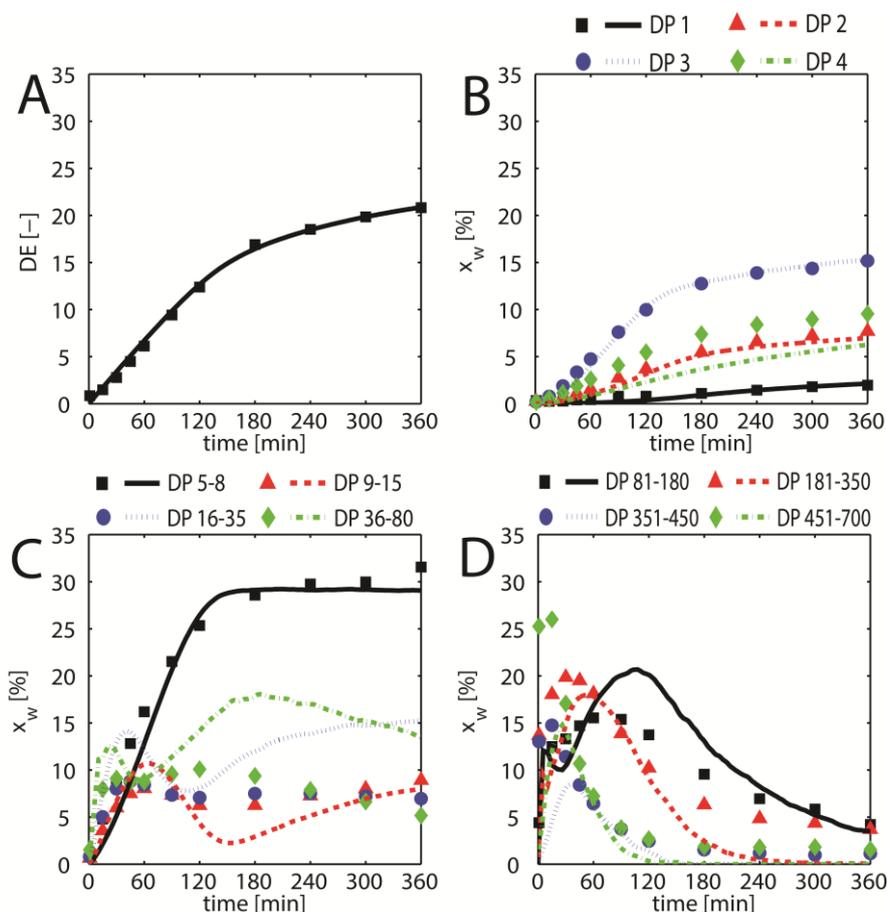

Figure 3.5. The predictions of starch hydrolysis products after adjusting the subsite map and the branch factors. Subsite map is given in Table 3.4, all branch factors were set to 0.03.

After comparing several combinations of subsite changes, a subsite map was defined that fitted our experimental data (Table 3.4). The new binding energy values improved the residuals from 413 (for only b_T = 0.03) to 326 (b_T = 0.03 and the subsite map as in Table 3.4) by improving the fit of DP 1-8. Changes made to energies of subsites +3, +4 and -5 were sufficient to obtain a good fit of DP 1, 2, 3 and 5-8, but in order to improve the fit of DP 4 also changes to subsite -4 needed to be made.

Table 3.4. The subsite map developed based on the data from wheat starch hydrolysis for BLA at 50°C.

Subsite number	-5	-4	-3	-2	-1	+1	+2	+3	+4
Binding energy [kJ/mol]	-8.0	-8.0	-5.1	-6.5	0.0	0.0	-5.1	-3.0	5.0

3.5.   Conclusions

The complete hydrolysis process of starch by α-amylase from *B. licheniformis* was modelled with a stochastic model. The model uses a combination of a subsite map with nine binding sites in the active centre of the enzyme and inhibition due to the proximity of an α-(1,6)-branching point in amylopectin. Our model is different from other models discussed in literature, as it follows all the products of starch hydrolysis over the whole course of the reaction and allows accurate predictions of their concentrations.

This model is also used to test and verify the accuracy of a subsite map, for predicting the composition of starch (both amylose and amylopectin) hydrolysis products. This is different from earlier approaches in which only hydrolysis of model oligosaccharides with DP up to 10 was considered for finding the binding energies of the subsite map.

Using a rational data-driven fitting procedure, we propose a subsite map that most accurately predicts the composition of our data set. For optimal results, especially if different analytical procedures are employed, further tuning of the subsite map might be required. However, we demonstrated that it is sufficient to vary values of subsites -5, +3 and +4 and, if necessary, the inhibition factor.

Changing the values of the apparent binding energies of the subsites mostly influences the composition of small carbohydrates (DP < 8). The concentrations of the large molecules (DP 9-181) over time were mostly affected by the inhibition factor. The structure of starch that is designed in the model most likely influences the initial concentrations of the largest carbohydrates (DP 350-700).

The inhibition factors do not significantly affect the concentrations of small oligosaccharides. Therefore, models that do not take into account carbohydrates larger than DP 5 cannot correctly estimate the actual values of the three inhibition factors. In those models the inhibition due to the presence of branches in amylopectin is not necessarily as important.

Transglycosylation and condensation can influence the composition of starch hydrolysis products. However, even without taking into account these mechanisms, our relatively uncomplicated model can quantitatively predict the product composition during starch hydrolysis by BLA, and may thus be used to predict the product quality as

function of the reaction conditions. Additionally, when hydrolysis data are available, this model can also be used to test the subsite maps of other enzymes.

3.6.   Acknowledgements

3.7.   References

Åkerberg C, Zacchi G, Torto N, Gorton L. 2000. A kinetic model for enzymatic wheat starch saccharification. Journal of Chemical Technology & Biotechnology 75(4):306-314.

Allen JD, Thoma JA. 1976. Subsite mapping of enzymes - depolymerase computer modeling. Biochemical Journal 159(1):105-120.

Baks T, Ngene IS, van Soest JJG, Janssen AEM, Boom RM. 2007. Comparison of methods to determine the degree of gelatinisation for both high and low starch concentrations. Carbohydrate Polymers 67(4):481-490.

Besselink T, Baks T, Janssen AE, Boom RM. 2008. A stochastic model for predicting dextrose equivalent and saccharide composition during hydrolysis of starch by alpha-amylase. Biotechnology and Bioengineering 100(4):684-97.

Hiromi K. 1970. Interpretation of dependency of rate parameters on the degree of polymerization of substrate in enzyme-catalyzed reactions. Evaluation of subsite affinities of exo-enzyme. Biochemical and Biophysical Research Communications 40(1):1-6.

Kandra L, Gyemant G, Remenyik J, Hovanszki G, Liptak A. 2002. Action pattern and subsite mapping of Bacillus licheniformis alpha-amylase (BLA) with modified maltooligosaccharide substrates. Febs Letters 518(1-3):79-82.

Kandra L, Remenyik J, Gyémánt G, Lipták A. 2006. Effect of temperature on subsite map of Bacillus licheniformis α-amylase. Acta Biologica Hungarica 57(3):367-375.

Kondo H, Nakatani H, Matsuno R, Hiromi K. 1980. Product distribution in amylase-catalyzed hydrolysis of amylose: comparison of experimental results with theoretical predictions. Journal of Biochemistry 87(4):1053-1070.

Macgregor EA, Macgregor AW, Macri LJ, Morgan JE. 1994. Models for the action of barley alpha-amylase isozymes on linear substrates. Carbohydrate Research 257(1):249.

Marchal LM, Ulijn RV, De Gooijer CD, Franke GT, Tramper J. 2003. Monte Carlo simulation of the alpha-amylolysis of amylopectin potato starch. 2. alpha-amylolysis of amylopectin. Bioprocess and Biosystems Engineering 26(2):123-32.

Murthy GS, Johnston DB, Rausch KD, Tumbleson ME, Singh V. 2011. Starch hydrolysis modeling: application to fuel ethanol production. Bioprocess and Biosystems Engineering 34(7):879-90.

Paolucci-Jeanjean D, Belleville MP, Zakhia N, Rios GM. 2000. Kinetics of cassava starch hydrolysis with Termamyl (R) enzyme. Biotechnology and Bioengineering 68(1):71-77.

Presečki A, Blažević Z, Vasić-Rački Đ. 2013. Mathematical modeling of maize starch liquefaction catalyzed by α-amylases from Bacillus licheniformis: effect of calcium, pH and temperature. Bioprocess and Biosystems Engineering 36(1):117-126.

Suganuma T, Matsuno R, Ohnishi M, Hiromi K. 1978. A study of the mechanism of action of Taka-amylase A1 on linear oligosaccharides by product analysis and computer simulation. Journal of Biochemistry 84(2):293-316.

Thoma JA, Brothers C, Spradlin J. 1970. Subsite mapping of enzymes - Studies on Bacillus subtilis amylase. Biochemistry 9(8):1768-&.

Torgerson EM, Brewer LC, Thoma JA. 1979. Subsite mapping of enzymes - Use of subsite map to simulate complete time course of hydrolysis of a polymeric substrate. Archives of Biochemistry and Biophysics 196(1):13-22.

Wojciechowski PM, Koziol A, Noworyta A. 2001. Iteration model of starch hydrolysis by amylolytic enzymes. Biotechnology and Bioengineering 75(5):530-539.
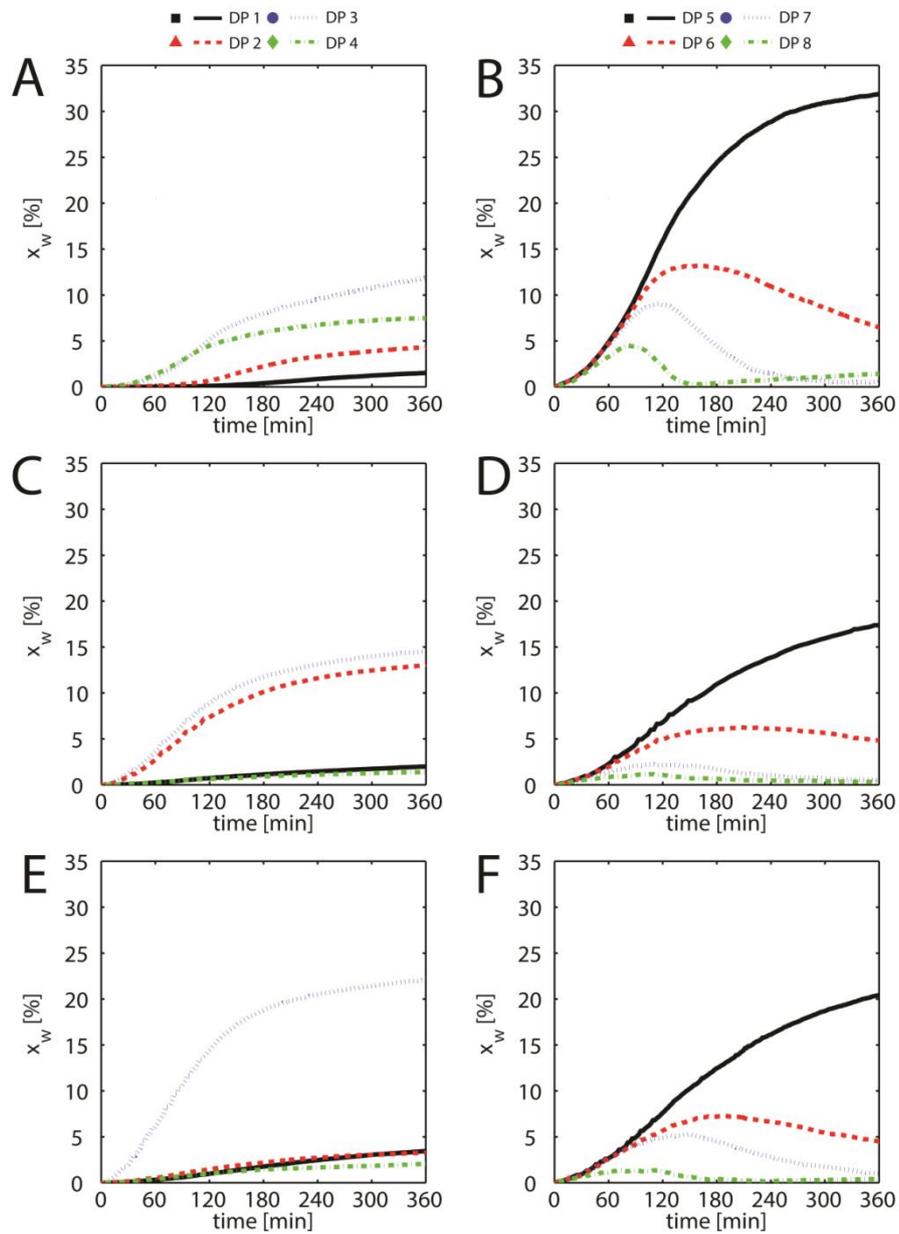
3.8. Appendix

Figure A 1. Predictions of the concentrations of DP 1-8 by the model as the binging energy of the indicated subsite was changed to 0: (AB) Subsite +4; (CD) Subsite +3; (EF) Subsite +2.
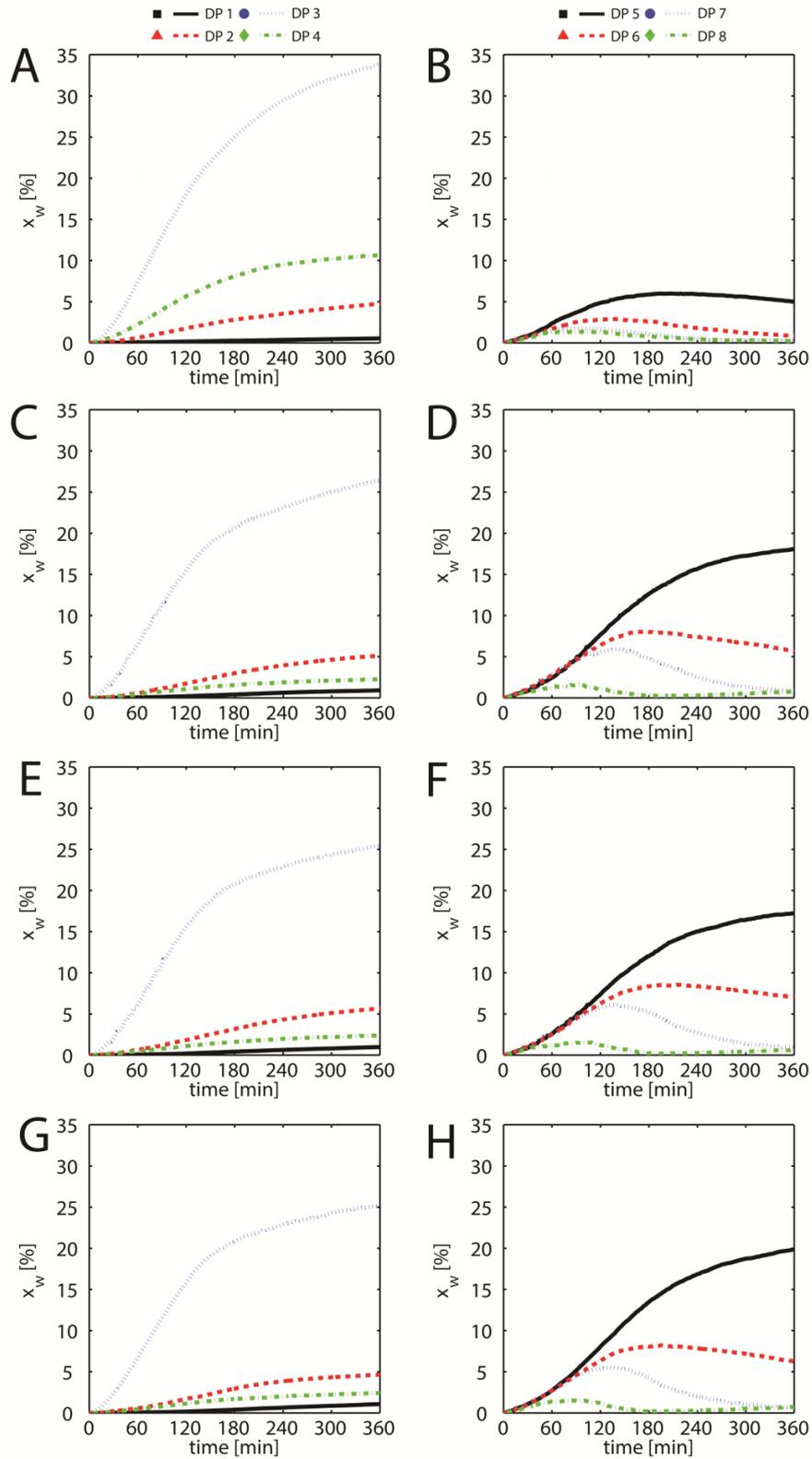
Figure A 2. Predictions of the concentrations of DP 1-8 by the model as the binging energy of the indicated subsite was changed to 0: (AB) Subsite -5; (CD) Subsite -4; (EF) Subsite -3; (GH) Subsite -2.
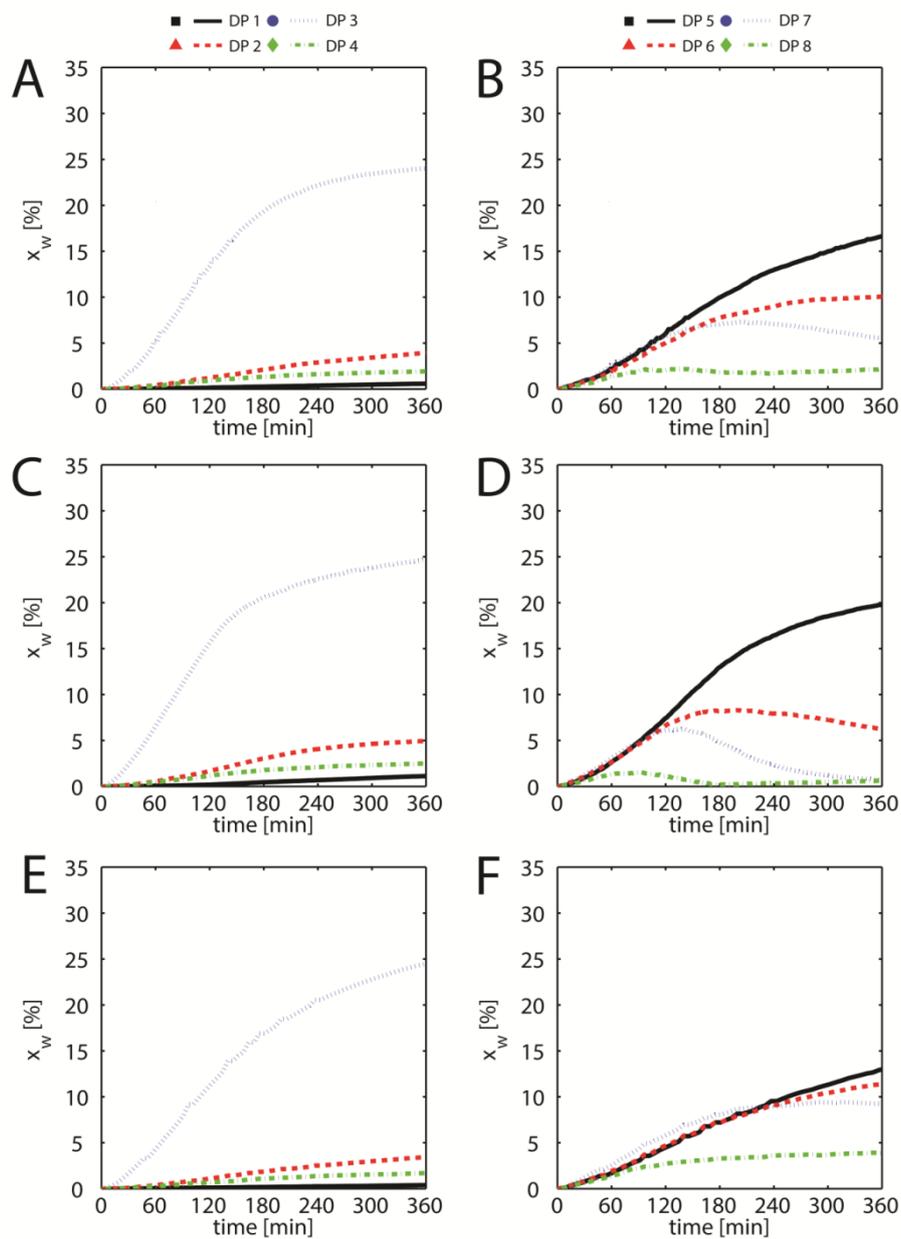
Figure A 3. Predictions of the concentrations of DP 1-8 by the model with different parameters of the model changed: (AB) The original subsite map and set of branch factors; (CD) Branch factors were set to 0.05; (EF) No inhibition (the code describing the inhibition was removed from the model).

The effect of temperature on the parameters of a stochastic model for enzymatic starch hydrolysis

Abstract

A stochastic model based on subsite theory, predicting the products of enzymatic wheat starch hydrolysis, was verified experimentally using data from starch hydrolysis at 80°C with *Bacillus licheniformis* α-amylase (BLA). The changes in the concentrations of all hydrolysis products were followed over time and analysed using size-exclusion chromatography (HPLC-SEC). The results from the hydrolysis at 80°C were first compared with previous results at 50°C to identify the effects of temperature on the product composition, and were subsequently used to verify the model predictions. Two parameters of the model, the inhibition and the subsite map, were analysed in more detail. The inhibition is important for correct estimation of the concentrations of the high molecular weight hydrolysis products. The subsite map of BLA at 80°C that was taken from literature was not accurate enough to obtain the correct prediction of the small hydrolysis products. A new subsite map was proposed using a step-wise procedure. Our procedure also led to identifying which subsites are important in determining the product composition. Contrary to what the literature data suggests, the binding energy values of the subsite maps at 50°C and at 80° were similar.

4.1. Introduction

With various factors influencing enzymatic hydrolysis it is not trivial to point out those that determine the distribution of hydrolysis products. The temperature and pH do not influence the mechanism of the reaction, but do influence the product composition. The structure of starch affects the final concentrations of the products. The branches in amylopectin, characterized by chain length distribution, along with the inability of the enzyme to cleave branches, force the enzyme to hydrolyse amylopectin in a different manner than it cleaves the linear substrates.

Changes in temperature influence enzymatic reactions by affecting the stability of the enzymes, competing reactions, affinity of the enzymes for activators, inhibitors and for the substrate, and the rate of substrate to product conversion (Whitaker 1993). *Bacillus licheniformis* α-amylase (BLA) is a thermostable hydrolase that reaches its highest activity around its optimum temperature of 90-95°C (Fitter et al. 2001). An enzyme is called thermostable when it has a high melting temperature (thermodynamic stability) or a long half-life of denaturation at a selected high temperature (kinetic stability) (Radestock and Gohlke 2008). The theories generally ascribe the differences in activity of thermostable enzymes at low and high temperatures to the flexibility of the protein (Vihinen (1987), Závodszky et al. (1998), Panasik Jr et al. (2000)). Enzymatic activity and conformational flexibility are indeed closely correlated (Závodszky et al. 1998). At high temperatures, with more collisions per unit of time, the increased mobility of the enzyme loosens up weak bonds and allows for more flexibility in some regions of the molecule, while still preserving its overall 3D structure. The active site may have different stability than the overall protein – a local reduction in the flexibility of the active site is the main feature in thermophilic amylases (D'Amico et al. 2003).

The active sites of hydrolases contain a number of binding sites that can hold the substrate. In case of amylases each subsite can bind one glucose unit of the starch chain. The final alignment of large molecules in the catalytic site is determined by the sum of the interactions over a number of glucose units and their binding sites. The specificity of the enzyme, the degree of polymerization of oligosaccharides produced during hydrolysis and the final carbohydrate profiles are determined by the number of subsites and the position and activity of the catalytic site (Synowiecki 2007).

In this chapter, we present the results of a study on the thermostable α-amylase from *Bacillus licheniformis* (BLA) and its action on wheat starch. One can use activity measurements to compare the action of an enzyme at different temperatures. These measurements are based on the overall composition of the reaction mixture, and show that reaction proceeds faster or slower, but do not give a detailed description of the products. We investigated the changes in product distribution during hydrolysis at 80°C and compared this distribution with previously described results (chapter 2) at a lower temperature (50°C). The data from starch hydrolysis at 80°C were then compared with simulations of a model that predicts the composition of starch hydrolysates over time (chapters 2 and 3). The model was based on the subsite theory, and its outcomes were mainly determined by the affinities of the subsites, and by the inhibition in the model. We describe the optimization of these parameters, such that the model reliably predicts the hydrolysis, and discuss their individual influence on the composition of wheat starch hydrolysates at 80°C.

4.2. Materials

Unmodified wheat starch (S5127) with 13.68 ± 0.24% of moisture (Sigma-Aldrich, Germany) was hydrolyzed with the thermostable α-amylase from *Bacillus licheniformis* (Sigma, Thermamyl 120®, a product of Novozyme Corp.). All chemicals (sodium hydroxide, calcium chloride di-hydrate (Merck, Germany)), carbohydrate standards for HPLC calibration (glucose, maltose, maltotriose, maltotetrose, maltopentose, maltohexose and maltoheptose, all minimum 90% purity, Sigma-Aldrich, Germany) and dextran analytical standards for gel permeation chromatography (5, 12, 25 and 50 kDa) were at least analytical grade. MilliQ water was used for the preparation of all solutions.

4.3. Methods

4.3.1. Starch hydrolysis

The method used for gelatinization and enzymatic starch hydrolysis followed the same protocol as described in chapter 2 with the exception of using a higher temperature (80°C) during hydrolysis. In summary, a suspension of wheat starch in demineralized water (10 w/w%) with 5 mM $CaCl_2 \cdot 2H_2O$ was heated to 90°C and mixed

(300 rpm) to gelatinize the starch for 1 hour. The temperature of the gelatinized mixture was then lowered to 80°C (± 1°C) and 0.01 w/w% of α-amylase from *Bacillus licheniformis* was added. During the reaction, samples of hydrolyzed starch mixture were pipetted into Eppendorf tubes and directly frozen in liquid nitrogen to stop the hydrolysis.

4.3.2. *Determination of carbohydrate composition*

A fraction of the frozen starch samples (0.3 g) was diluted with 90 μl of 2 M NaOH and 1,110 μl of demineralized water and centrifuged for 10 minutes at 2,400 × *g* and 4°C. The supernatant was filtered using syringes and Minisart single use filters.

The carbohydrate composition was measured by using size exclusion chromatography (HPLC-SEC). The Dionex Ultimate 3000 HPLC system was equipped with a Shodex Sugar KS-803 column with KS-G guard and an RI-detector. The column was operated at 80°C and used MilliQ water as eluent at a flow-rate of $0.3 \ \text{ml·min}^{-1}$.

Using the equation from the calibration curve, prepared with standards of known molecular weight, and the average molecular weight of each group of carbohydrates, we estimated the elution times. The mass fractions of carbohydrates within each group were calculated as described in chapter 2.

Hydrolysis experiments were performed in triplicate and the resulting data were compared with the model. We randomly chose one set of data for the comparison of the experiments with the model, as the differences between the individual data sets were small and insignificant. All data sets are shown in Figure A1 in the Appendix.

4.3.3. *Model*

The model described in chapter 2 was used with small adaptations. The model was adjusted for the increased temperature of the reaction (80°C) and consequently the subsite map that is available for *Bacillus licheniformis* α-amylase at 80°C was used (Kandra et al. (2006)).

All presented model outputs are the average values based on three repetitions with the same set of parameters. Since the model is stochastic in nature, the three runs do not give exactly equal results. However, the differences between the results were an indication of the reliability of the results, which was high. Matlab 2012a (7.14) was used

to perform all simulations. Two parameter sets of the model were analysed in detail: the subsite map and the branch factors. For more details on these parameters we refer to chapter 3 and Marchal et al. (2003).

The distribution of the hydrolysis products is affected by the values of apparent binding energies within the subsite map. The binding energy values of the original subsite map (Kandra et al. 2006) were treated as initial values. During the simulations one by one the energy of each subsite was set to 0, thereby removing the effect of that subsite completely. This procedure demonstrated which groups of saccharides were especially affected by a particular subsite. After changing the binding energies of a subsite within a range of values and comparing how the outcomes of the model differ, it becomes possible to choose the approximate binding energy of each subsite. The optimization of the subsite map was confirmed using the square root of the sum of squares of the differences between the model and the data (SRSS).

As demonstrated previously (chapter 3), inhibition had an effect on the concentrations of the high molecular weight products. Three different branch factors (Marchal et al. 2003) were necessary for a correct description of the inhibition. During the simulations of the subsite map these branch factors were set to: b_I = 0.1, b_{II} = 0.1 and b_{III} = 0.1. These values were subsequently varied until the best-fitting set of branch factors was obtained.

4.4. Results and discussion

4.4.1. *Experimental results of starch hydrolysis at different temperatures*

The product profiles of starch hydrolysis by BLA at 80°C (Figure 4.1) are qualitatively similar to profiles during hydrolysis at lower temperature (50°C) (chapter 2). Within the first minutes of the reaction the high molecular weight molecules (eluting at retention time 17-20 min) are hydrolysed. In the following 60 minutes the enzyme attacks mainly the largest molecules present in the solution, represented by the tall peak at 21 minutes (DP > 300). As these molecules are degraded to cluster-sized smaller structures (DP 150-300), also the quantities of the oligosaccharides begin to increase. The intermediate size molecules (DP 20 -150) remain on a steady level for as long as cluster-size molecules are available for the enzyme. Starting from the 120th minute, these intermediate size molecules become the enzyme's main target. As only the

quantities of maltose, glucose and DP 9-15 increase during that period, these intermediate size molecules seem to be hydrolysed by cleaving off small units (glucose, maltose). The branched structures that are left (DP < 20) can only be hydrolysed with difficulty. In the final stages of the reaction all molecules of DP > 20 are hydrolysed, and finally mostly mono-, di- and oligosaccharides remain.
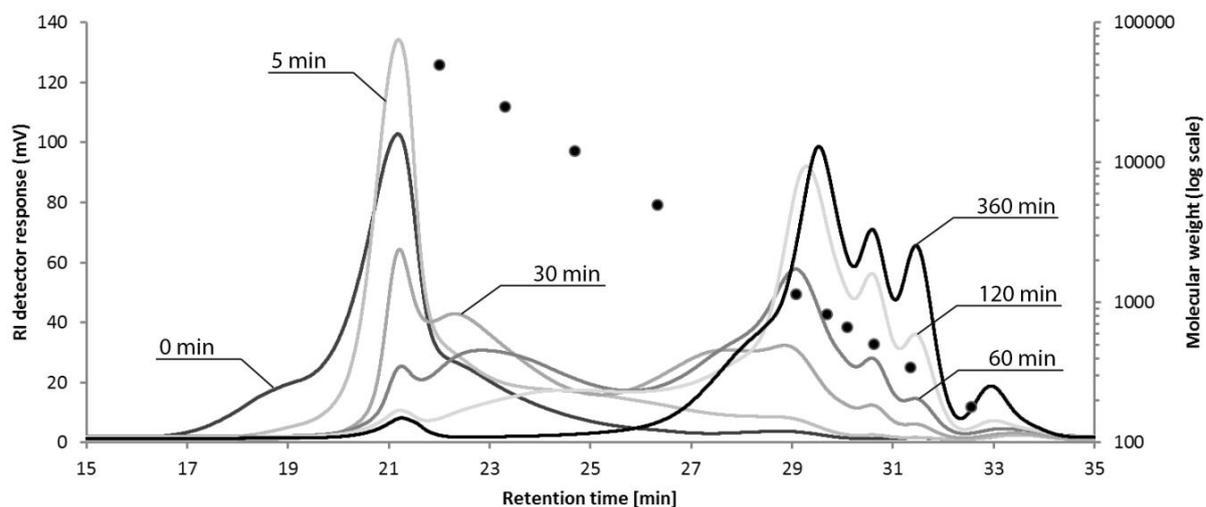

Figure 4.1 The chromatograms representing the products of wheat starch hydrolysis by BLA at 80°C during six stages of the reaction. The dotted plot represents the logarithm of average apparent molecular weight of the standards at the indicated retention times.

A comparison of the hydrolysis results shows that the reaction proceeds faster at 80°C than it does at 50°C (Figure 4.2 A-C). Up until the 120th minute of the reaction the concentrations of DP1-8 increase linearly for both temperatures. At 80°C, the slope of DE vs. time is larger, reflecting a 1.5 times higher product formation rate than at 50°C. In the advanced stages of hydrolysis the increased probability of non-productive bindings contributes to the decrease in the rate of reaction (Synowiecki 2007). The high molecular weight molecules are digested much more rapidly in the beginning of the reaction and at the higher temperature. At 80°C, these large molecules have been hydrolysed after 180 minutes.
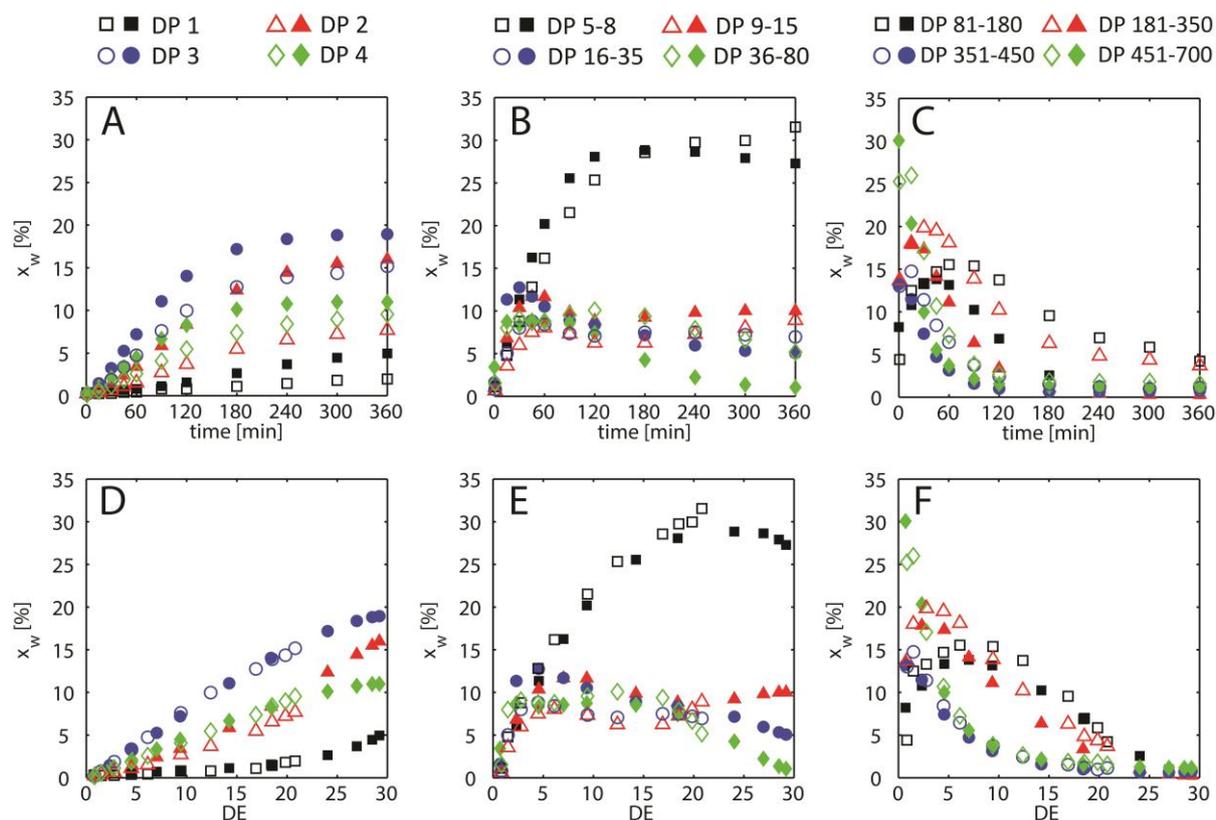
Figure 2. Products of starch hydrolysis at 50°C (empty symbols) and 80°C (filled symbols) plotted vs time or vs the dextrose equivalent (DE).

The increased reaction rate is not the only difference between the two sets of data. The relative concentrations of the oligosaccharides are also different at different temperatures. At 80°C the quantities of DP 5-8 are somewhat lower, and more maltotriose, maltose and glucose are produced, compared to the reaction at 50°C. However, when the data are plotted as a function of the dextrose equivalent (the degree of conversion), the differences become smaller. The concentrations of maltose (at DE 10-20), DP 9-15 and 16-35 (at DE 5-20) are slightly higher at 80°C, whereas the concentrations of DP 5-8 (at DE 15-20) keep increasing at 50°C, when at 80°C they begin to decrease (Figure 4.2 D-F). This shows that the hydrolysis temperature does not substantially affect the product composition after all.

A more flexible binding between the enzyme and the substrate in the active site at high temperatures may render an increased number of productive complexes between the enzyme and substrate, that leads to an increased reaction rate. However, this larger flexibility in binding of the substrate could also lead to decreased specificity

of the enzyme at high temperatures, which indeed has been reported (Marchal et al. (1999), Nakakuki et al. (1984), Saito (1973)).

Marchal et al. (1999) observed differences in product composition (DP 1-10) at different temperatures (50, 70 and 90°C) during the hydrolysis of amylopectin from potato with *B. licheniformis* α-amylase. The changes in temperature especially affected the final concentrations of maltose and maltopentose. The authors speculated that the differences could stem either from amylopectin molecules being more rigid at lower temperatures or the different capability of the enzyme to hydrolyse DP 5 and 6 at different temperatures. In addition, differences in product composition may result from increased transglycosylation. Some of these factors may affect the composition of our samples as well, but we see much smaller differences between 50°C and 80°C. The larger differences in the oligosaccharides concentrations in the samples of Marchal et al. (1999) as compared to our samples may also stem from the use of a different source of substrate (potato vs. wheat) and/or the lack of amylose.

The differences in the energy values of the subsite maps at 50 and 80°C (Kandra et al. (2006), summarized in Figure 4.6) imply that the composition of products differs at a higher temperature. In the subsite map for 80°C the barrier site (subsite +4) has a lower energy value (+4.1 kJ·mol^{-1}) than the same subsite at 50°C (+8 kJ·mol^{-1}). This smaller barrier allows for a more uniform distribution of the products, e.g. more DP 4 is produced at higher temperatures. In fact, nearly all subsites (with the exception of -5 and +2) are assigned much lower energy values at 80°C.

4.4.2. *Predicting the product composition*

BLA hydrolysis of starch at 80°C yields a higher final dextrose equivalent (DE 30) in the same amount of time as at 50°C. The higher DE is the result of a rapid accumulation of the small hydrolysis products. In our simulations, the DE, along with the number of formed enzyme-substrate complexes, are used to fit the model data to the experimental results. The model DE needs to be at least as high as in the experiments for a good description of the experimental results. With the same number of hydrolysis events occurring during the simulation (timespan), as chosen for the 50°C model (3·10^6 events), it was not possible to reach the DE of 30 when using the original subsite map at 80°C. The timespan was therefore extended to 2·10^7 events, or in some cases even to

$4 \cdot 10^7$ hydrolysis events. The intervals, at which the hydrolysis events were saved, were kept constant to make the comparisons possible, even if the total number of events was larger.

4.4.2.1. The subsite map

The overall fit of the model and data was not satisfactory with the original values of the subsite map at 80°C found in literature (Kandra et al. 2006). The subsite map influences the composition of oligosaccharides, but does not have a major effect on the molecules with a degree of polymerization (DP) larger than 8 (chapter 3). Figure 4.3 shows that the predicted concentrations of e.g. maltose and maltotetrose should be higher and the concentrations of maltotriose and DP 5-8 need to be lower.
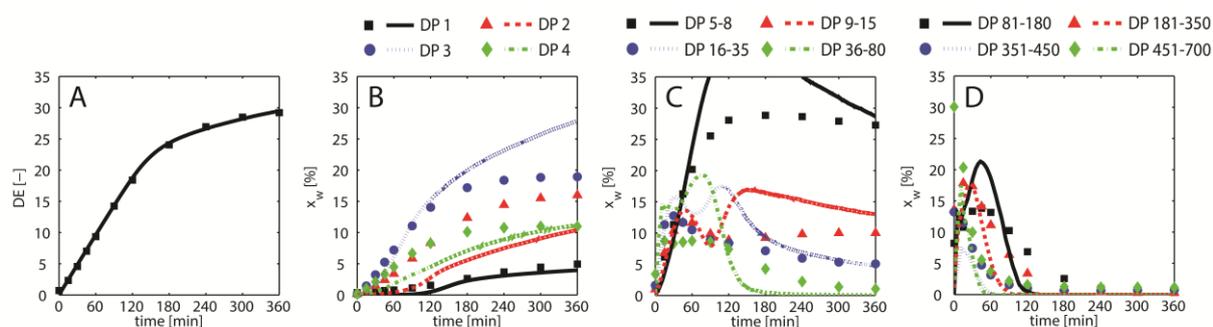

Figure 4.3. Predictions of the model (lines) with the original subsite map for 80°C (see Figure 4.6) compared with the experimental data (symbols) from wheat starch hydrolysis by BLA at the same temperature. Inhibition factors were set to 0.1.

We thus wish to improve the subsite map following similar steps as described previously (chapter 3). First, the values of energy assigned to every subsite were changed to 0, to assess the influence of each subsite on the composition of the products (Figure 4.4 and Table A1 in the Appendix). Three of the changes necessary for better predictions in Figure 4.3 B and C (DP 2, 3 and 5-8) already follow with the value of subsite +3 set to 0. Subsite -2 did not influence the composition when its energy was changed to 0: the simulation yielded the same outcome as with the original subsite map. Changing subsite -3 had only marginal effect on the product composition. However, varying the value of subsite +3 with simultaneous adjustment of the values of subsites +4 and -5 was most relevant for more accurate prediction of the hydrolysis products.
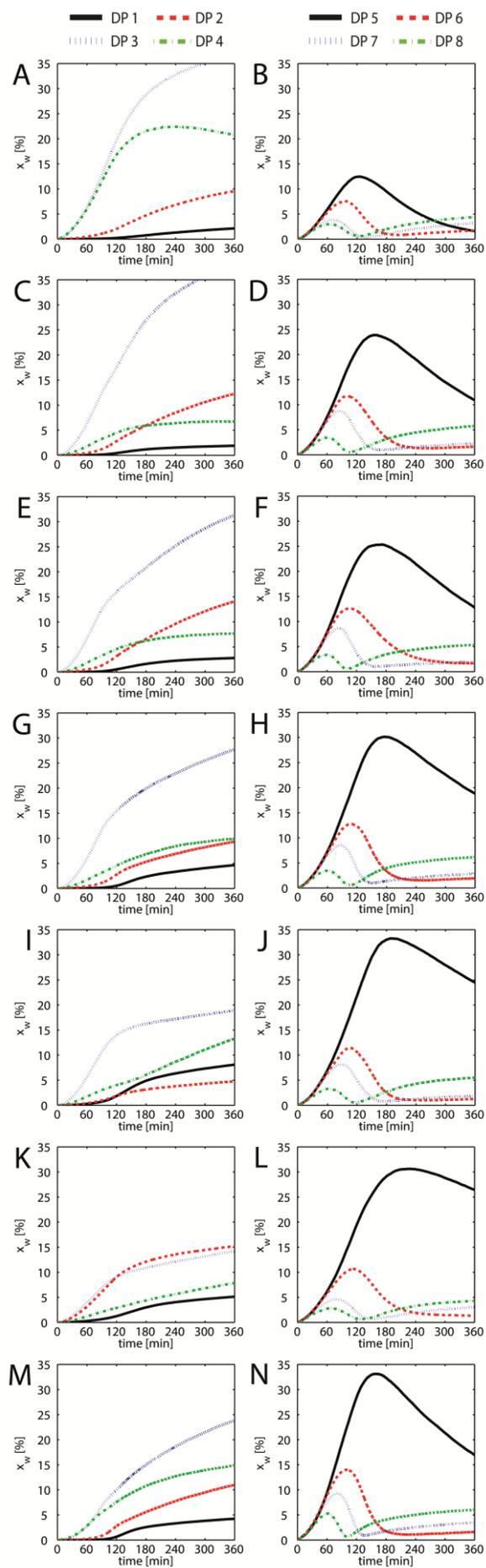
Figure 4.4 Predictions of the concentrations of DP 1-8 by the model after changing one of the subsites to 0: (AB) subsite -5; (CD) subsite -4; (EF) subsite -3; (GH) subsite -2; (IJ) subsite +2; (KL) subsite +3; (MN) subsite +4. The legend to the plots is placed above the column, plots on the left hand side represent DP 1-4, plots on the right hand side DP 5-8.

Table 4.1 summarizes the information about the influence of subsites on the concentration of a particular oligosaccharide, based on the results shown in Figure 4.4 and the SRSS values given in Table A1. The change of the energy of a subsite either caused an increase in the concentration of a carbohydrate, a decrease or had no influence at all. Subsites that appear most often in Table 4.1 influence the concentrations of numerous oligosaccharides in a substantial way.

Table 4.1 The effect of each subsite on the concentration of a particular oligosaccharide (DP 1-8). Out of all nine subsites, only those that influence (increase or decrease) the concentration of a saccharide are shown. Subsites indicated in bold have a more pronounced effect.

	Subsite number	
Oligosaccharide	**Subsites that increase concentration**	**Subsites that decrease concentration**
DP1	+2	-4 -5
DP2	**+3** -3	+2
DP3	-3 -**4** -**5**	+2 +3 +4
DP4	+4 -**5**	+3 -3 -4
DP5	+2 +3 +4	-3 -4 -**5**
DP6		-5
DP7		-**5** +3
DP8		-5 +3

The energy values of the outer subsites are more important when it comes to predicting the product composition. The composition of small hydrolysis products is especially affected by subsites -5, +2 and +3. Therefore, the binding energies of these subsites were varied in a range of values, to converge to more accurate energy values and thus improve the predictions. The results for two of the subsites are shown in Figure 4.5. By assigning different energy values to a subsite we were able to limit the apparent binding energy values to those resulting in an improved fit of the model and the data.
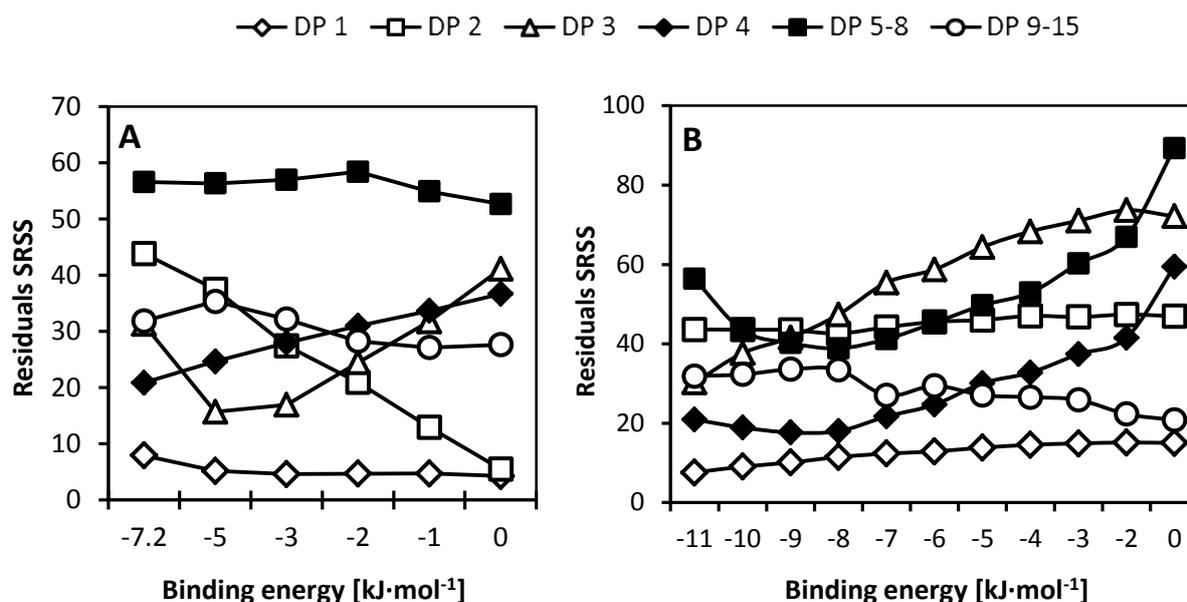
Figure 4.5. The changes in SRSS values as (A) the binding energy of subsite +3 is varied from 0 to -7.2 kJ·mol^{-1} and (B) the binding energy of subsite -5 is varied from 0 to 11.2 kJ·mol^{-1}. The remaining subsites were assigned the original energy values from the subsite map at 80°C (Kandra et al. 2006). The lower the SRSS values are, the more accurate predictions are obtained. The legend is given above the figure, the lines in the plots are added to guide the eye.

Subsite +2 was omitted from Figure 4.5, because the variation of this subsite between -3 and -7 kJ·mol^{-1} gave no changes, as long as the other subsites had their original values. We therefore chose to keep the original value (-5.6 kJ·mol^{-1}) of subsite +2, as we found no reason to change it. Only completely removing the subsite by assigning it a value of 0, caused a visible change in the concentrations of oligosaccharides.

The binding energy values of subsite +3 were varied between 0 and -7.2 kJ·mol^{-1} (Figure 4.5 A). For subsite +3 the analysis of the residuals is not entirely straightforward, since we observed better agreement with some oligosaccharides and poorer with others, as the energy was changed. The energy of subsite +3 must be high enough (0 or -1 kJ·mol^{-1}) for an optimal fit of maltose (DP 2) and yet low enough to facilitate a good fit of maltotriose (DP 3) and maltotetraose (DP 4).

Subsite -5 influences the concentrations of most of the oligosaccharides. The energy values of subsite -5 were varied from 0 to -11.2 kJ·mol^{-1} (Figure 4.5 B). When the energy value assigned to subsite -5 is close to 0 the fit of the model is poor for the oligosaccharides, but not for the DP 9-15 group. As demonstrated by the SRSS values in Figure 4.5 B, the predictions correspond better with the data when subsite -5 is assigned

a less negative energy value (-9, -8 kJ·mol⁻¹) than in the original subsite map (-11.2 kJ·mol⁻¹).

Overall, the model results at 80°C are more sensitive to changes in the subsite map than the outputs at 50°C. Changes to subsites -2, -3, -4 and +2 had little influence on the prediction of the hydrolysis products at 50°C (chapter 3). At 80°C, only subsites -2 and -3 seem to have no effect on the predictions (Figure 4.4 EF and GH). Energy changes in all of the remaining subsites lead to significant changes in the concentrations of the oligosaccharides. As at 80°C the hydrolysis was more extensive and yielded higher quantities of products within the same amount of time, the energy changes of more than just the three subsites could have actually influenced the model predictions.

The final predictions of the model, just as in the experimental hydrolysis, depend on the combination of the binding energies of multiple subsites. The optimal binding energy of each subsite is based not only on obtaining the best value of that subsite, but also the right energy values assigned to the subsites around it. The energy value of each subsite (e.g., subsite +3) can therefore be found only after it has been varied together with the energies assigned to other subsites (e.g., +2, +4 and -5).
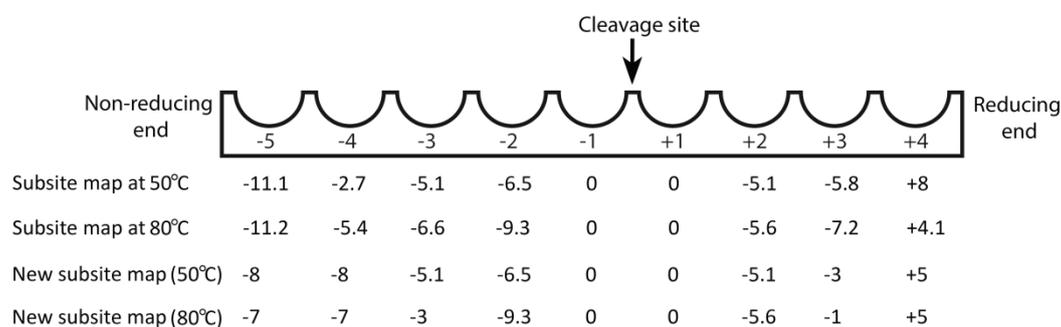

	-5	-4	-3	-2	-1	+1	+2	+3	+4
Subsite map at 50°C	-11.1	-2.7	-5.1	-6.5	0	0	-5.1	-5.8	+8
Subsite map at 80°C	-11.2	-5.4	-6.6	-9.3	0	0	-5.6	-7.2	+4.1
New subsite map (50°C)	-8	-8	-5.1	-6.5	0	0	-5.1	-3	+5
New subsite map (80°C)	-7	-7	-3	-9.3	0	0	-5.6	-1	+5

Figure 4.6. The binding energies [kJ·mol⁻¹] of subsites of *Bacillus licheniformis* α-amylase (BLA) at 50 and 80°C as given by Kandra et al. (2006) and the new subsite maps at 50°C (from Chapter 3) and at 80°C developed in this chapter based on the data from wheat starch hydrolysis.

Our step-wise analysis of the subsite map can be compiled into a new subsite map (Figure 4.6). This new subsite map allows accurate predictions of the concentrations of the oligosaccharides obtained during hydrolysis of wheat starch at 80°C by BLA. Compared to the original subsite map, the most important changes were made in the energies of subsites +3 and -5. Lowering the energy of subsite +3 to -1 kJ·mol⁻¹ and the energy of subsite -5 to -7 kJ·mol⁻¹ was already enough to substantially

improve the fit of the model. The total SRSS value reduced from 430 (out of which DP 1-8 was 161) for the original subsite map to 360 (74 for DP 1-8) for the new one, because of more accurate predictions of DP 2, 3, 4 and 5-8. Subsequent changes made to the energy values of subsites +4, -4, and -3 facilitated an improved fit of DP 2 and DP 5-8 and improved the shape of the model curves (total SRSS 365, out of which DP 1-8 was 68).
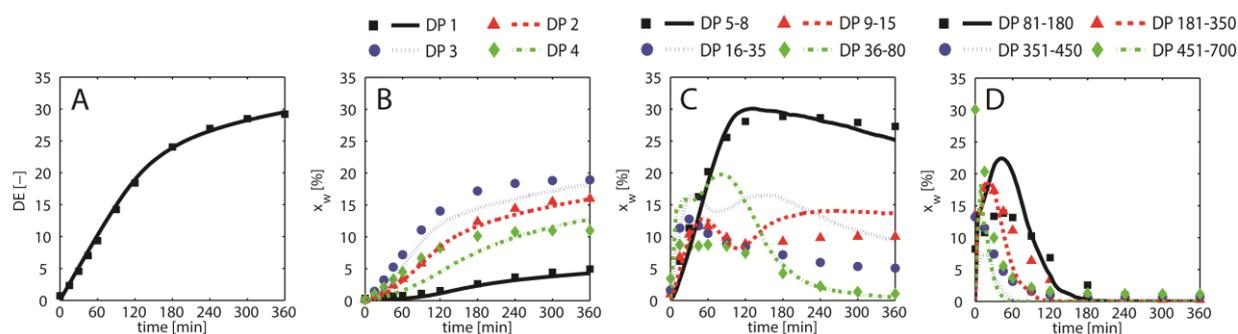

Figure 4.7 The outputs of the model (lines) simulating hydrolysis with the new subsite map (Figure 4.6) fitted with the experimental data (symbols) from wheat starch hydrolysis by BLA at 80°C.

The new subsite map gives more accurate predictions and the lowest SRSS values with DP 1-8. However, in the process of choosing the best subsite map several other combinations resulted in a similar outcome (Figure A2 in the supplementary material). Marchal et al. (2003) used a different approach for optimizing the subsite map of the α-amylase from *Bacillus amyloliquefaciens* for hydrolysis of potato amylopectin (50°C), but also noticed that more than one subsite map can lead to nearly the same model predictions with only small differences in SRSS values. In their most accurate four final subsite maps the energy values of subsites -6, -5, +2, +3 and +4 barely changed. The largest differences in the energy values were seen for the inner subsites -4, -3 and -2. We thus hypothesize that the values of the outer subsites are more important and therefore must remain stable (conserved subsites) for accurate predictions, whereas the energies assigned to the inner subsites had only minor effect on the final outcome of the model regardless of their values, and therefore could be changed without consequences.

When comparing the new subsite map with the one we obtained for BLA at 50°C, we notice that the energy values of the individual subsites are more similar than in the maps of Kandra et al. (2006). It is unclear whether the reason for the differences between our and the previously reported subsite maps lays in the different substrates (starch vs tagged maltooligosaccharides) or in the much longer time given for the

hydrolysis in our case. However, it is known that liquefying enzymes preferably degrade substrates with more than 15 glucose units (Synowiecki 2007), and therefore shorter substrates are expected to form different enzyme-substrate interaction. When Ermer et al. (1993) used two slightly different substrates in subsite mapping of glucoamylase from *A. niger*, *p*-nitrophenyl-α-maltooligosaccharides and maltooligosaccharides, they found that the subsite affinities differed substantially for these two short linear substrates. The difference in the structure between such substrates and starch will also result in different subsite maps. Neither our method, nor the one used by Kandra et al. (2006) are a direct measurement of the energy of binding within a subsite map, and thus it is difficult to draw absolute conclusions from either of these.

Studies on thermostable enzymes show that at room temperature most of these enzymes have low catalytic activity (Fitter et al. (2001), Radestock and Gohlke (2011)). The activity of thermostable enzymes increases with increasing temperature, reaching the highest point near the optimum temperature. The active site or the regions around it are more rigid at lower temperatures (so-called overly stable structure), and become more flexible only when the temperature increases towards the optimum value (Závodszky et al. (1998), Radestock and Gohlke (2011)). At high temperatures, the enzyme maintains its native structure by remaining rigid enough to avoid denaturation and flexible enough to remain active (Fontana et al. (1998)). However, the 'loosening' of the active site would suggest that the subsite binding would be weaker at higher temperatures. While the subsite map by Kandra et al. (2006) at 80°C shows overall stronger binding compared to the one at 50°C (all negative binding energies together shift from -36.3 to -45.3 kJ·mol^{-1}), in the new subsite maps the overall binding energies are reduced at higher temperatures (all negative binding energies shift from -35.7 to -32.9 kJ·mol^{-1}). This aligns better with the general idea of the 'loosening' of the active centre.

Thermostable enzymes can have multiple conformations at low temperatures that render them dysfunctional (Panasik Jr et al. 2000). Certain conformations of the enzyme prevent the substrate from fitting in the subsite or active site in the same way. The active site of BLA at temperatures below 60°C is rather rigid (Fitter et al. 2001), which can be one cause for lower activity at mesophilic temperatures. BLA has a relatively low activation energy measured in the range between 50 and 90°C and it

shows a strong increase in reactivity with temperature, as the relative activity of the enzyme nearly doubles within this temperature range (Fitter et al. 2001). The low activation energy also implies that binding energies at these temperatures will be high. The binding energies will then dictate the further course of the process, into cleavage and desorption of the reaction products.

4.4.2.2. Inhibition

While the subsite map is most important for correct prediction of the small oligosaccharides, the inhibition is a major determinant in the predictions of the larger carbohydrates. Endo-acting α-amylases yield larger oligosaccharides as products of starch degradation, mainly because hydrolysis of α-1,4-linkages close to the non-reducing end of a substrate molecule is slower for these enzymes (MacGregor et al. 2001). Lowering the values of the branch factors (b_T) increases the inhibition of hydrolysing α-1,4-bonds in the proximity of branches (glucose units with an α-1,4,6-bonds) (chapter 3). An increased inhibition in the model results in a higher number of non-productive complexes formed, by allowing more possibilities for non-productive binding. The higher the inhibition, the more hydrolysis events are required to reach the same DE value, which has impact on both the absolute rate of hydrolysis and on the relative concentrations of the reaction products.

In the model of starch hydrolysis at 50°C it was necessary to use low values of the branch factors (0.05 or 0.03), because at the lower temperature the hydrolysis rate was overall lower (Figure 4.2). At 80°C the active centre is more flexible, which is expected to lead to lower inhibition, and therefore the values of the branch factors are higher than those chosen in the model at 50°C.

The subsite map (80°C) was optimized with all three branch factors set to 0.1 (total SRSS of 365). Increasing the values of inhibition factors above 0.1 led to worse predictions and a higher SRSS value in all cases (examples are given in Table A2 in the supplementary material). With the proposed subsite map (Figure 4.6) and the branch factors set to 0.09 the fit of the model improved (total SRSS 336). The lower values of the branch factors resulted in lower SRSS values of all groups starting from DP 36-180 till DP 450-700.

Lowering the branch factors below 0.09 did not result in a lower SRSS value of the model in all but one case (Table A2). Decreasing the branch factors below 0.09 especially affected the predictions of DP 5-8, causing a decline in the concentration of this group as the inhibition in the model increased. With branch factors set to 0.09 the inhibition extended up to 8 bonds away from the branch point, weakening with increasing distance from the α-1,6-linkage. With branch factors set to values lower than 0.09 the extent and distance of the inhibition increased, which lowered the chance of detaching products that fall in the group of DP 5-8.

At any lower values of the branch factors the inhibition was nearly irrelevant to the fit of large carbohydrates and did not affect their concentrations. The only lower total SRSS (323) appeared when the branch factors were set to 0.06, but we deemed the resulting poor fit of the group DP 5-8 to be unacceptable. Other improvements became insufficient to make up for it. Since we could not realise further improvement using different sets of branch factors (Table A2), we assumed that the minimum value was reached. We concluded that as long as the branch factors were set to 0.1 or just below, the best possible fit between the model and the data was accomplished.

The flexibility of the active site along with the higher collision rates at higher temperatures contribute to faster hydrolysis of the largest molecules present in the solution. One may expect that a reduced specificity in the binding between enzyme and substrates also translates in less hindrance due to the proximity of branches. Therefore the inhibition should decay faster with the distance from an α-1,6-bond branch point, and thus the values of branch factors should become higher at higher temperatures. We speculate that in the extreme, this might even lead to the hydrolysis of α-1,6-bonds at higher temperatures. MacGregor et al. (2001) states that an α-1,4-linked glucose and an α-1,6-linked glucose cannot be bound the same way in the active site. The amylase capable of cleaving both α-1,4- and α-1,6- linkages needs to be quite flexible in the active site, especially at subsite +1. Since the flexibility of thermostable proteins increases at high temperatures (Panasik Jr et al. (2000), Závodszky et al. (1998), Vihinen (1987)), we cannot exclude that the chance increases that some of the α-1,6-bound glucose units might also be hydrolysed. It should be clear that this is not covered by our current model.

4.5. Conclusions

Hydrolysis of wheat starch by *Bacillus licheniformis* α-amylase (BLA) at different temperatures does not result in drastically different product composition.

The model discussed in this chapter can describe the distribution of all products during the course of wheat starch hydrolysis by *Bacillus licheniformis* α-amylase (BLA) at different temperatures. The predictions of the model at 80°C are accurate with a newly developed subsite map of BLA, based on the composition of products from starch hydrolysis. The binding energies in the new maps are less divergent than in the existing ones, and the differences in the binding energies between 50°C and 80°C are in accordance with the proposed mechanism of 'loosening' of the active centre.

Based on the data we collected, more than one subsite map can be accurate enough to predict the composition of oligosaccharides present in the product mixture. The energies of some subsites influence the reaction strongly and are therefore conserved (especially -5, +3 and +4), while the energy values of other sites do not have a major impact, and thus can differ considerably (e.g. subsite -2 or -3).

At 80°C the inhibition is lower than at 50°C. We expect that to be a result of higher flexibility of the enzyme at elevated temperatures, described for thermostable enzymes.

Both the intrinsic factors (properties of the enzyme) and the extrinsic ones (dependent on the substrate characteristics) influence the final composition of products during starch hydrolysis by α-amylase, and have to be taken into account when modelling enzymatic hydrolysis reactions.

4.6. Acknowledgements

4.7. References

D'Amico S, Marx J-C, Gerday C, Feller G. 2003. Activity-Stability Relationships in Extremophilic Enzymes. Journal of Biological Chemistry 278(10):7891-7896.

Ermer J, Rose K, Hubner G, Schellenberger A. 1993. Subsite Affinities of Aspergillus-Niger Glucoamylase-Ii Determined with P-Nitrophenylmaltooligosaccharides. Biological Chemistry Hoppe-Seyler 374(2):123-128.

Fitter J, Herrmann R, Dencher NA, Blume A, Hauss T. 2001. Activity and stability of a thermostable alpha-amylase compared to its mesophilic homologue: Mechanisms of thermal adaptation. Biochemistry 40(35):10723-10731.

Fontana A, Filippis VD, Laureto PPd, Scaramella E, Zambonin M, A. Ballesteros FJPJLI, Halling PJ. 1998. Rigidity of thermophilic enzymes. Progress in Biotechnology: Elsevier. p 277-294.

Kandra L, Remenyik J, Gyémánt G, Lipták A. 2006. Effect of temperature on subsite map of Bacillus licheniformis α-amylase. Acta Biologica Hungarica 57(3):367-375.

MacGregor EA, Janeček Š, Svensson B. 2001. Relationship of sequence and structure to specificity in the α-amylase family of enzymes. Biochimica et Biophysica Acta (BBA) - Protein Structure and Molecular Enzymology 1546(1):1-20.

Marchal LM, Ulijn RV, De Gooijer CD, Franke GT, Tramper J. 2003. Monte Carlo simulation of the alpha-amylolysis of amylopectin potato starch. 2. alpha-amylolysis of amylopectin. Bioprocess Biosyst Eng 26(2):123-32.

Marchal LM, van de Laar AMJ, Goetheer E, Schimmelpennink EB, Bergsma J, Beeftink HH, Tramper J. 1999. Effect of temperature on the saccharide composition obtained after α-amylolysis of starch. Biotechnology and Bioengineering 63(3):344-355.

Nakakuki T, Azuma K, Kainuma K. 1984. Action patterns of various exo-amylases and the anomeric configurations of their products. Carbohydrate Research 128(2):297-310.

Panasik Jr N, Brenchley JE, Farber GK. 2000. Distributions of structural features contributing to thermostability in mesophilic and thermophilic α/β barrel glycosyl hydrolases. Biochimica et Biophysica Acta (BBA) - Protein Structure and Molecular Enzymology 1543(1):189-201.

Radestock S, Gohlke H. 2008. Exploiting the Link between Protein Rigidity and Thermostability for Data-Driven Protein Engineering. Engineering in Life Sciences 8(5):507-522.

Radestock S, Gohlke H. 2011. Protein rigidity and thermophilic adaptation. Proteins: Structure, Function, and Bioinformatics 79(4):1089-1108.

Saito N. 1973. A thermophilic extracellular α-amylase from Bacillus licheniformis. Archives of Biochemistry and Biophysics 155(2):290-298.

Synowiecki J. 2007. The Use of Starch Processing Enzymes in the Food Industry. In: Polaina J, MacCabe A, editors. Industrial Enzymes: Springer Netherlands. p 19-34.

Vihinen M. 1987. Relationship of protein flexibility to thermostability. Protein Engineering 1(6):477-480.

Whitaker JR. 1993. Principles of Enzymology for the Food Sciences, Second Edition: Taylor & Francis.

Závodszky P, Kardos J, Svingor Á, Petsko GA. 1998. Adjustment of conformational flexibility is a key event in the thermal adaptation of proteins. Proceedings of the National Academy of Sciences 95(13):7406-7411.
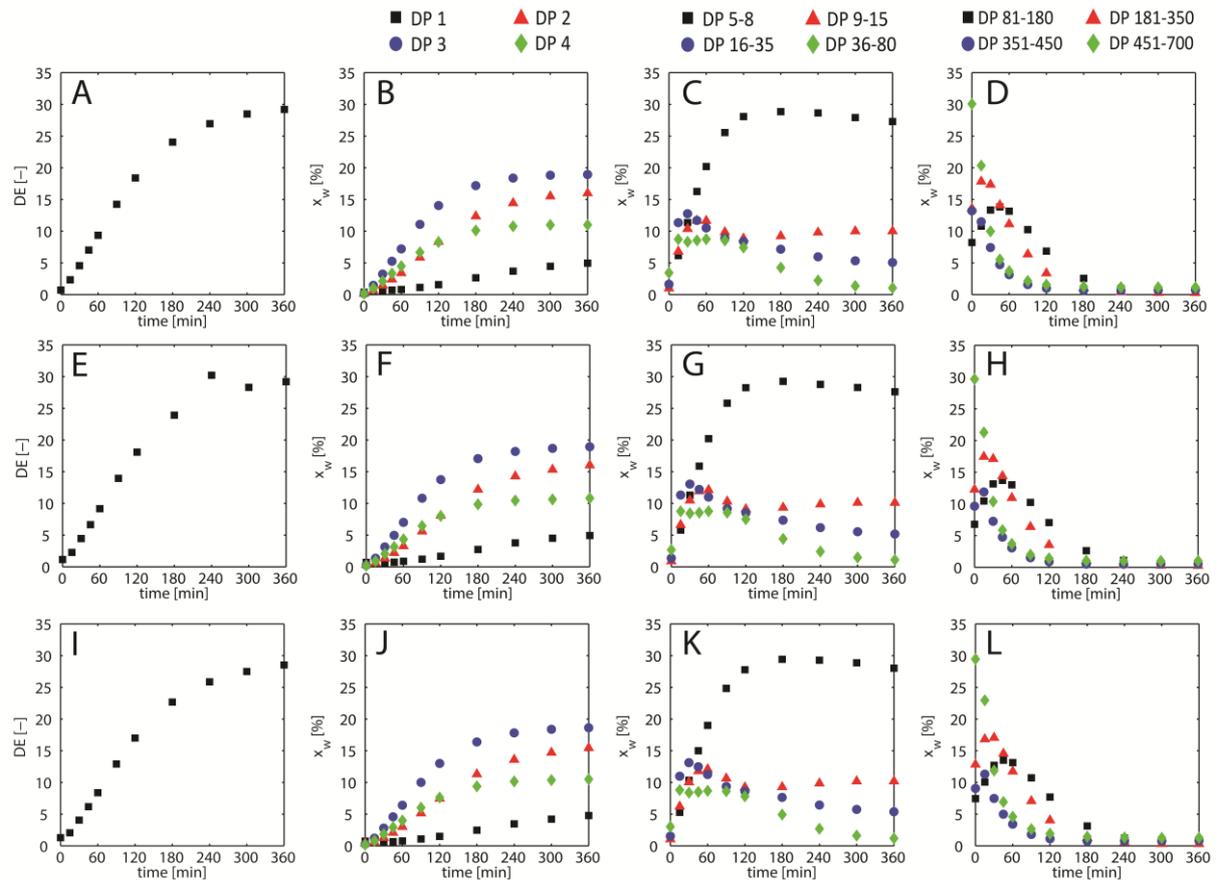
4.8.   Appendix

Figure A1. The data collected during hydrolysis of starch by BLA at 80°C – three repetitions of the experiment at the same reaction conditions.

Table A 1. The SRSS values for every carbohydrate group (DP group) as the binding energy of each subsite was changed to 0 kJ·mol^{-1}. The SRSS of the output computed with the original subsite map is given for comparison. The branch factors were set to 0.1 in all computations.

| DP group	Subsite changed to 0							Original subsite map
	-5	**-4**	**-3**	**-2**	**+2**	**+3**	**+4**	
DP$_1$	15.0	14.3	11.3	6.2	14.0	4.3	6.6	7.7
DP$_2$	47.0	38.3	32.5	46.1	61.0	5.5	42.7	43.3
DP$_3$	72.1	69.6	40.1	30.4	9.7	41.1	30.8	30.5
DP$_4$	59.5	31.8	28.8	23.3	24.0	36.7	17.0	20.5
DP$_{5-8}$	89.3	45.3	50.2	56.2	61.7	52.7	75.0	52.4
DP$_{9-15}$	20.8	34.5	37.5	33.9	33.9	27.7	37.6	33.4
DP$_{16-35}$	39.8	39.7	41.3	31.4	40.3	44.4	27.8	29.9
DP$_{36-80}$	48.5	46.1	47.6	47.8	49.5	56.6	46.4	49.3
DP$_{81-180}$	38.4	42.0	47.6	46.0	38.7	42.4	43.6	42.3
DP$_{181-350}$	30.9	35.5	38.1	34.8	36.6	33.1	39.2	36.6
DP$_{351-450}$	28.0	31.1	30.6	30.6	28.2	28.4	32.4	29.6
DP$_{451-700}$	58.0	60.2	59.4	56.6	55.1	55.9	62.4	54.5
Total	<u>547.3</u>	488.3	464.9	443.4	452.7	<u>428.8</u>	461.5	430.1
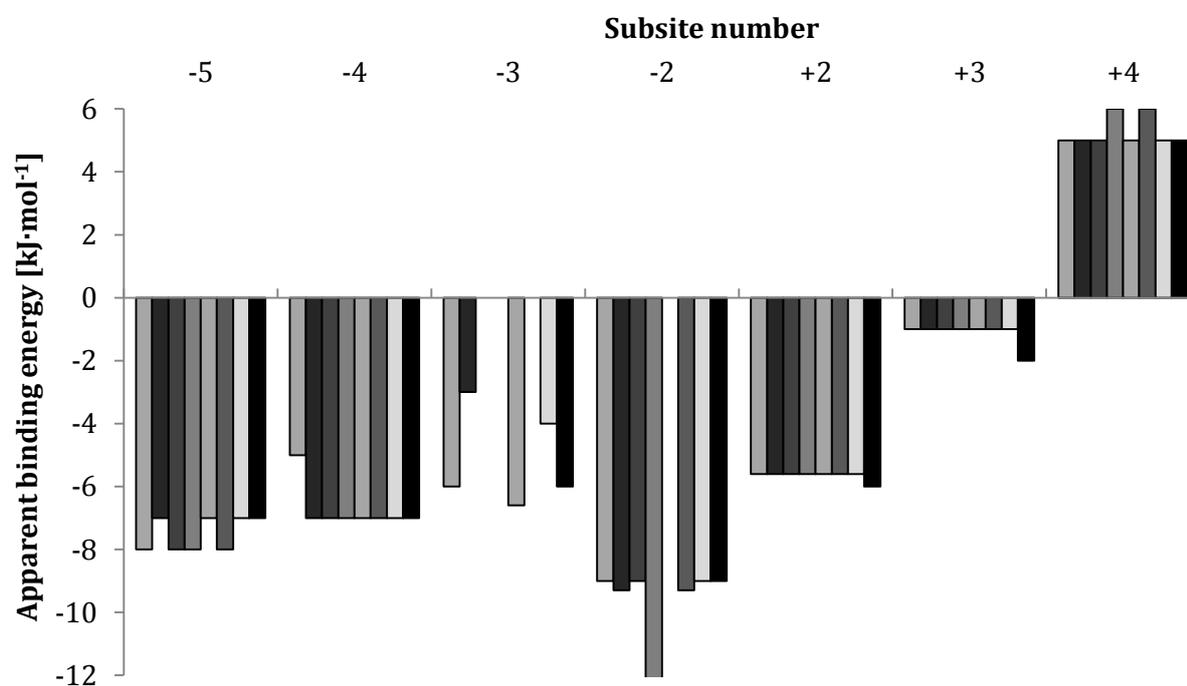
Figure A2. The binding energies of eight other subsite maps that resulted in similar predictions to the finally chosen subsite map (Figure 4.6). When the model outputs were simulated using these subsite maps the SRSS values varied from 355 to 380.

Table A 2. The values of SRSS, representing the differences between the model and the data, as the branch factors are varied in the model. The new subsite map (given in Figure 4.6) was used.

Branch factors			Total SRSS
b_I	b_{II}	b_{III}	
0.15	0.15	0.15	441
0.1	0.1	0.1	365
0.09	0.09	0.09	336
0.08	0.08	0.08	337
0.07	0.07	0.07	335
0.06	0.06	0.06	323
0.05	0.05	0.05	344
0.04	0.04	0.04	334
0.1	0.4	0.1	375
0.2	0.1	0.1	415
0.1	0.1	0.2	422
0.1	0.2	0.1	359
0.09	0.4	0.09	373
0.11	0.09	0.09	343
0.09	0.06	0.09	346
0.06	0.09	0.09	336
0.09	0.09	0.06	358
0.09	0.1	0.09	354
0.09	0.09	0.1	348
0.08	0.09	0.09	354
0.07	0.09	0.09	328
0.05	0.09	0.09	344
0.07	0.12	0.08	332
0.07	0.15	0.13	357

Modelling of the liquefaction and the saccharification of wheat starch at high substrate concentrations

Abstract

Complete hydrolysis of wheat starch was simulated using a liquefaction model and a saccharification model that are based on subsite theory and use the subsite maps of the enzymes to predict the hydrolysis products. The liquefaction model was also used to create the substrate matrices for the saccharification model. Available subsite maps of glucoamylase from *Aspergillus niger* (GA) were tested in the saccharification model to identify the best subsite map to describe the product composition of the hydrolysed samples. The experimental data were collected during a three-step process. Starch gelatinization in a shearing device preceded the two-step enzymatic hydrolysis reaction. The liquefaction of wheat starch by *Bacillus licheniformis* α-amylase (BLA) at 80°C and subsequent saccharification by glucoamylase from *Aspergillus niger* at 60°C were followed over time at high substrate concentrations (30-60 w/w%). All of the hydrolysis products were then analysed by using HPLC-SEC. The composition of the hydrolysis products in the experiments and the products predicted by the models were compared and the quality of the predictions was evaluated. We demonstrated that the models can predict the outcome of starch hydrolysis at different conditions (e.g., different substrate concentrations) and for different types of enzymes.

5.1. Introduction

Intensification of starch hydrolysis by reducing the water content can increase the volumetric productivity (van der Veen et al. 2006), and the stability of enzymes (de Cordt et al. (1994), Rosendal et al. (1979)), while lowering the energy costs during heating or evaporation (van der Veen et al. (2006), Grafelman and Meagher (1995)). However, these benefits come at a price, as increasing the concentration of starch leads to increased viscosity of the solution that impairs mixing, transport in convectional jet cookers (Grafelman and Meagher 1995) and also imposes mass transfer limitations (Baks et al. (2008), Sanromán et al. (1996)). On top of those effects, pH and temperature control, inhomogeneity (Curic et al. 1998), and longer reaction times add to the complexity of low moisture systems.

From a chemical standpoint, starch is merely a homopolymer of glucose. Unfortunately, when forming a solution in water starch exhibits complex physical properties. At high concentrations (>50%), starch slurries start to resemble pseudoplastic fluids and their apparent viscosity increases with increasing shear rate (Miranda et al. (1991), Sanromán et al. (1996)). This shear-thickening behaviour impairs mixing at increased dry matter concentrations.

Melting of starch granules is a requirement for complete hydrolysis of starch. However, gelatinization of starch becomes more difficult as the moisture content decreases. According to Baks et al. (2007), complete thermal gelatinization of 60 w/w% starch slurry is possible only when temperatures above 120°C are applied (45 minute-long treatment). For comparison, at 10 w/w% of starch, complete gelatinization can already be reached at 70°C (Baks et al. (2007)). However, both Baks et al. (2008) using an extruder and van der Veen et al. (2006) using a shearing device demonstrated that if during gelatinization of starch a mechanical and a thermal treatment are combined, complete gelatinization of starch can be achieved, even at low moisture contents and at lower temperatures. The mechanical stress (e.g., shear stress) facilitates the breakdown of starch granules and the loss of order of the granular structure which, just as gelatinization, increases the surface area of the substrate (Baks et al. 2008). Shearing also facilitates the homogeneity of the sample, contrary to samples obtained in a batch reactor (preliminary tests) or in an extruder (Curic et al. 1998).

When the starch concentration increases, the branched structure of amylopectin imposes steric hindrance that can decrease the rate of the reaction (Sanromán et al. 1996). While mass transfer limitations are decreasing when depolymerisation of starch occurs during hydrolysis, especially during liquefaction (Sanromán et al. 1996), most of the other effects remain. During saccharification, additionally the occurrence of the reversed reaction, product inhibition by glucose (Cepeda et al. 2001) and crystallization add to the complexity of concentrated systems.

With many opposite forces at play, a model describing enzymatic hydrolysis of starch at increased substrate concentrations can become complex. Available models of saccharification by glucoamylase focus on simulating the most optimal hydrolysis conditions (Cepeda et al. (2001), Curic et al. (1998)), predicting the concentration of glucose, sometimes also maltose and maltotriose (Nikolov et al. (1989), Lee et al. (1992), Nagy et al. (1992), Åkerberg et al. (2000), Cepeda et al. (2001), Rashid et al. (2006)) and describing the inhibition or enzyme activity (Zanin and De Moraes (1996), GonzalezTello et al. (1996), Polakovič and Bryjak (2002), van der Veen et al. (2005)). Most of these models are empirical or deterministic models, based on the Michaelis-Menten equation, usually simplifying the substrate by treating it as a single reacting component (Cepeda et al. (2001), Nagy et al. (1992),van der Veen et al. (2005)).

To our knowledge there are no publications describing the simulation of both liquefaction and saccharification, at substrate concentrations exceeding 30 w/w% or describing formation of all hydrolysis products. Van der van der Veen et al. (2005) obtained a satisfactory fit when following the hydrolysis of a 60% maltodextrins solution to glucose and maltose using a deterministic model with reversed reactions. Later on van der Veen et al. (2006) also described their predictions of the dextrose equivalent (DE) at 35-65%, yet they did not compare the simulations with any experimental results.

The stochastic model, described in our previous work (chapter 2, 3 and 4), follows the concentrations of all molecules present in the reaction mixture, in a range of reaction conditions, provided the DE values in time are known. The subsite map of the enzyme should also be known, but as subsite mapping of enzymes becomes easier using computer programs (Kandra et al. (2002), Gyémánt et al. (2002), Mótyán et al. (2011)), more subsite maps are becoming available.

This chapter reports on a model predicting all the products present during the enzymatic liquefaction and saccharification of wheat starch at high dry matter contents. The model predicting the products of wheat starch liquefaction by *Bacillus licheniformis* α-amylase (BLA) was already tested at relatively low (10 w/w%) substrate concentrations (chapters 2, 3 and 4). In this chapter we describe the use of the model to predict the outcomes of liquefaction at higher starch concentrations (30-60 w/w%), without implementing any additional limitations (i.e., mass transfer limitations, additional inhibition etc.). The outcome of this model – the matrix representing hydrolysed starch (maltodextrins), is used as the input for the saccharification model.

Subsequently, we describe the second step in starch hydrolysis, the saccharification of the liquefied starch by glucoamylase from *Aspergillus niger*. The model of starch liquefaction is converted into a model for an exoamylase and follows over time the hydrolysis of maltodextrins into glucose. Several subsite maps available for glucoamylase are tested in the model and compared with the results from hydrolysis at 30% dry matter. After the subsite map that best describes the formation of saccharification products is chosen, it is then used to predict the saccharification products at higher substrate concentrations.

5.2. Materials

Unmodified wheat starch (S5127) with the moisture content of 11.9 ± 0.2% was purchased from Sigma-Aldrich (Steinheim, Germany). *Bacillus licheniformis* bacterial thermostable α-amylase (EC 3.2.1.1) Type XII-A (Thermamyl 120®) and glucoamylase from *Aspergillus niger* were selected for hydrolysis, as both are commonly used in the commercial hydrolysis of starch. Both enzymes were ordered from Sigma-Aldrich (products of Novozyme Corp.). All the solutions were prepared using Milli-Q water. Analytical grade sodium hydroxide, sulphuric acid and calcium chloride di-hydrate originated from Merck, Germany. Carbohydrate standards for HPLC calibration (glucose, maltose, maltotriose, maltotetraose, maltopentaose, maltohexaose and maltoheptaose, minimum 90% purity) and dextran analytical standards for gel permeation chromatography (5, 12, 25 and 50 kDa) were purchased from Sigma-Aldrich, Germany.

5.3. Methods

5.3.1. Gelatinization

The overall method used for starch hydrolysis resembled the method proposed by van der Veen et al. (2006). Starch mixtures (30, 40, 50 or 60% dry matter) were prepared using MilliQ water with the addition of $CaCl_2$ (5mM). The gelatinization of starch took place in a conical shearing device with controlled shear rate, which used similar principles to a rheometer (for more details about the device we refer to van der Zalm et al. (2012). The space between the two cones of the shear cell was filled with the starch mixture. Starch was then mixed for 60 minutes with the speed of 50 rpm, at 90°C. The mechanical treatment (shear) along with the high temperature not only ensured complete gelatinization thanks to equally distributed temperature, but also allowed proper mixing and homogeneity of the samples with low water contents.

The dry matter contents of the samples were measured using a moisture analyser balance after gelatinization was finished. If necessary, water was added to maintain the right substrate concentration. The amount of water added to the gelatinized samples varied from 4 v/w% for the 30 w/w% samples up to 22 v/w% for the samples containing 60 w/w% dry matter.

5.3.2. Hydrolysis

The shear cell was not used for enzymatic hydrolysis for two reasons. Opening and closing the device caused evaporation of moisture, which was especially unwanted in the already concentrated samples. Second, the high mechanical stress applied during gelatinization could lead to enzyme inactivation, and that was why better results were obtained when the enzyme was added after the thermo-mechanical treatment (Baks et al. (2008), van der Veen et al. (2006)). To avoid enzyme inactivation and moisture loss, the hydrolysis of starch was carried out in a batch reactor equipped with an anchor stirrer. Although a batch system was more optimal for sampling and allowed for better preservation of enzymatic activity, it also had its drawbacks. The biggest challenges were the mixing in the top part of the reactor during liquefaction and avoiding the risk of crystallization during saccharification.

<u>*5.3.2.1. Liquefaction*</u>

After gelatinized starch was transferred from the shearing device into a batch reactor, the temperature of the mixture was adjusted. Subsequently, *Bacillus licheniformis* α-amylase (BLA) was added to the substrate mixture and liquefaction reaction took place for 5 hours (80°C, 150 rpm, E = 0.1 w/w%). Samples were taken regularly and frozen in liquid nitrogen to stop the reaction. Due to the limitation of the shear cell sample size, in the experiments that continued as saccharification only one sample, at 5h, was taken during liquefaction.

<u>*5.3.2.2. Saccharification*</u>

After the liquefaction process was finished, the temperature in the reactor was lowered to 60°C and the pH was adjusted to 4.5 using 1N H_2SO_4. At these conditions the α-amylase was no longer active (data not shown). During saccharification, just as in the liquefaction process a lower dosage of enzyme was used to facilitate a gradual hydrolysis. As the optimum conditions for glucoamylase from *Aspergillus niger* were reached, 0.023 w/w% of enzyme was added and the reaction was followed for 25 hours. According to Cepeda et al. (2001), 0.05 w/w% of enzyme would facilitate the 95% conversion of a suspension of 30 w/w% starch within 24 hours at 60°C. Thus we would have to double the dosage of enzyme and maintain a constant enzyme to substrate ratio if our goal would have been the complete conversion within 24 hours; however this was not the case.

Samples were taken regularly during the reaction, transferred into Eppendorf tubes and frozen in liquid nitrogen to terminate the reaction. The methods for sample treatment and the analysis of the carbohydrate composition were as described earlier (chapters 2, 3 and 4).

5.3.3. Modelling

<u>*5.3.3.1. Liquefaction model*</u>

The liquefaction model was fitted with the data from starch liquefaction at high dry matter contents and also used to create the substrate matrices (inputs) for the saccharification model. The starting DE values of the saccharification samples at each substrate concentration were used as an indicator for the end point of the liquefaction in

the model (Table 5.2). The DE, being the degree of conversion, was found by adjusting the timespan of the model. The model timespan is a measure of the number of reaction events occurring in the model, and therefore can be compared to the time of the reaction. However, since not all reaction events lead to hydrolysis the relation between the number of reaction evens and the degree of conversion is not linear. The model timespan was varied to create matrices of hydrolysed starch (maltodextrins) composed of molecules that are distributed similarly to the hydrolysates in the experiments.

5.3.3.2. Saccharification model

The saccharification model follows the same outlines as the liquefaction model. To create the model the basics of the subsite theory for depolymerizing enzymes were applied (Thoma et al. (1970); (1971), Hiromi (1970), Hiromi et al. (1973), (1976a; 1976b)). In short, subsites constitute the active site of enzymes and are locations where monomeric units of the substrate interact with the enzyme. Each subsite has a characteristic free energy of interaction with a glucose residue (MacGregor and MacGregor 1985). Binding of a substrate residue in a particular subsite can either cause the total free binding energy of enzyme-substrate (E-S) complex to increase (when the subsite binding energy is negative) or to decrease (if the binding energy of a subsite is positive) by the value assigned to the subsite (Besselink et al. 2008)). Subsites act independently, meaning that binding at one subsite will not affect the free energy of interaction in the neighbouring site (MacGregor and MacGregor 1985).

Based on the subsite theory, the association constant of the E-S complex (K_n) can be calculated using the binding energies of the subsites occupied by the substrate residues. The relationship between the association constant (K_n) and the subsite affinities (A_i) or free energy of binding (ΔG_i) of occupied subsites is summarized in the equation:

$$K_n = \exp\left(\frac{\sum_i^{cov} A_i}{RT}\right) = \exp\left(\frac{-\sum_i^{cov} \Delta G_i}{RT}\right)$$

where n is the DP of the molecule, $\sum_i^{cov}$ implies that the sum is calculated only for the occupied subsites starting from subsite i, R is the gas constant ($8.13 \frac{J}{mol \cdot K}$) and T is the reaction temperature [K].

With this equation one can calculate the association constants (K_n) of a particular substrate molecule (DP_n) with the enzyme. The ratio of the K_n of an E-S complex and the association constant (K_{max}) - calculated with the most negative value of binding energy possible, determines the chance of hydrolysis. If the formed E-S complex overlaps the catalytic site, hydrolysis will occur provided the chance of hydrolysis is larger than a random number between 0 and 1 (Besselink et al. 2008).

The first step in creating a stochastic model for saccharification is to describe the mode of action of a glucoamylase. The cleavage site described in the saccharification model is different than in the liquefaction model, because α-amylase (endoenzyme) hydrolyses bonds within the chain of starch, and glucoamylase (exoenzyme) hydrolyses the α-(1,4)-glycosidic bonds of the non-reducing end glucose unit. A saccharification model is therefore simpler than the model of liquefaction, since only one type of products (glucose) can be formed.

Second, not all bonds can be hydrolysed by glucoamylase. The part of the model defining the limitations of the enzyme is crucial, because glucoamylase can only detach the non-reducing end glucose units. The smallest DP of the substrate to be hydrolysed is thus set to 2, leaving glucose as the only product that cannot be hydrolysed. The α-(1,4)-bonds near a branch (an α-(1,6)-linkage) can be hydrolysed, but the hydrolysis of the branch itself is not permitted.

A number of subsite maps of *Aspergillus niger* glucoamylase have been reported (Table 5.1), using mostly isomaltooligosaccharides (IMOS) or maltooligosaccharides (MOS) as substrates for subsite mapping (Ermer et al. (1993), Meagher et al. (1989), Ono et al. (1988), Stoffer et al. (1997)). Based on our previous reports (chapters 3 and 4) and those of Ermer et al. (1993) or Meagher et al. (1989), the type of substrate used to create the subsite map influences the energy values of the subsites. Therefore some of the presented subsite maps might not allow for accurate predictions of the reaction products obtained with our specific substrate. The subsite maps we found were determined at different reaction conditions, but we kept the temperature in the model at 60°C (333K), as it was the optimal temperature used in our experiments.

We applied the principles of parsimony (van Boekel 2008) to limit the number of parameters in the model, and decided to keep the simulation as simple as possible by using the following assumptions:

- No reversed reactions take place,

- Mass transfer limitations are not included,

- Molecules of all sizes can be hydrolysed with the same rate (Ermer et al. 1993),

- No inhibition is implemented, including the inhibition caused by a branch point (an α-(1,6)-glycosidic bond).

Table 5.1 Subsite maps of different forms of *A. niger* glucoamylase and the conditions at which they were determined.

Enzyme form[a]	Subsite binding energies [kJ/mol] per subsite[b]							Conditions	Substrate[c]	Reference
	-1	+1	+2	+3	+4	+5	+6			
B-I	2.4	-19.0	-7.1	-2.6	-1.6	-0.7	-1.0	25°C, pH 5		
B-II	0.5	-21.0	-7.1	-3.5	-0.6	-1.2	-0.3	25°C, pH 4-4.5		
B-III	1.5	-21.7	-6.8	-1.8	-1.0	-0.9	-0.4	25°C, pH 4	MOS	Ono 1988
A-II	3.0	-22.0	-6.8	-1.8	-1.6	-0.5	-2.2	25°C, pH 3.5-4		
A-III	2.0	-21.7	-6.8	-1.0	-1.0	-0.4	-0.7	25°C, pH 4-4.5		
GA-I	1.4	-21.1	-6.4	-3.6	-1.1	-0.3	-0.4		MOS	
GA-I	5.6	-15.3	-4.9	-0.5	-1.0	-0.3		35°C, pH 4.5	iMOS	Meagher 1989
GA-II	5.1	-22.1	-6.1	-3.1	-1.4	-0.5	-0.5		MOS	
GA-II	3.8	-15.0	-4.9	-0.5	-1.2	-0.2			iMOS	
GA-II	-1.1 ± 0.8	-21.8	-7.3	-3.5	-0.3	-0.2	-0.3	25°C, pH 4.5	MOS	Ermer 1993
GA-II	-0.8 ± 0.4	-9.6	-9.8	-6.2	-1.0	-1.2	-0.4		p-NP-α-MOS	
wild type	-2	-20.8	-6.8	-2.8	-0.7	-0.4	0.5	45°C, pH 4.5	MOS	Stoffer 1997

[a] "Enzyme form" refers to the isoforms of glucoamylases reported for *A. niger*.
[b] Subsite numbering following Davies et al. (1997), with the non-reducing end located at subsite -1
[c] MOS – maltooligosaccharides, iMOS – isomaltooligosaccharides, p-NP-α-MOS - p-nitrophenyl-α-maltooligosaccahrides

Both the liquefaction and the saccharification model assume that only one molecule of enzyme is attacking any particular substrate molecule at the same time, which reflects the experimental conditions with the concentration of the substrate being much higher than that of the enzyme (Nakatani 1996).

The comparison between the model outputs and the data was possible by converting the model time into experimental time. This was done using non-linear fitting, based on the expectation that non-productive encounters between the enzyme and the substrate last nearly three times shorter than the productive ones (Wojciechowski et al. 2001). The procedure included only the experimental DE values and the time intervals at which they were measured and used all of the DE data points. Thus, the model incorporates the enzyme concentration and the influence of the

experimental conditions on the product composition in an indirect sense, through the changes in the DE.

5.4. Results and discussion

5.4.1. Gelatinization and hydrolysis at low moisture contents

After the thermo-mechanical treatment the gelatinized samples were compared to native starch samples under a microscope. The lack of granular structure in the gelatinized samples indirectly confirms that the gelatinization process was complete (not shown).

Increasing the substrate concentration (30-60%) and keeping the enzyme concentration constant during the liquefaction of wheat starch resulted in a lower reaction rate. In all cases, the reaction was still in an initial phase when it was terminated. As a result, the concentrations of DP 1-4 (Figure 5.1) were relatively low (<5 w/w%). The profiles of products were not that different at higher substrate concentrations, compared with data at low substrate concentrations. Baks et al. (2008) found that even as the concentration of starch is increased, the final product profiles (< DP 8) do not change, even though the final DE reached within the same time is somewhat lower. We observed similar effects in our liquefaction samples (Table 5.2, Figure 5.1). Baks et al. (2008) concluded that the concentration of substrate does not influence the selectivity of the enzyme. Additionally, the enzyme was more stable when starch was hydrolysed at higher substrate concentrations (50-70%), since BLA still remained active after several hours at low moisture contents.

The liquefaction samples containing 30 w/w% dry matter were hydrolysed most extensively, with the highest quantities of maltotriose (DP3), maltotetraose (DP4) and DP 5-8 (Figure 5.1, empty bars). A similar DE value of samples in both experiments at 50 and 60 w/w% dry matter reflects a similar distribution of the hydrolysis products, both oligo- and polysaccharides. At 40 w/w%, a lot of the large polysaccharide groups remained present after 5 hours despite the higher DE value, and similar amounts of oligosaccharides were produced to those at the higher starch concentrations.

5.4.2. Modelling of liquefaction by BLA

Modelling of wheat starch hydrolysis by BLA at low substrate concentrations (10%) showed that the model results were especially influenced by the binding energy values in the subsite map and by the inhibition in the model (chapters 3 and 4). In the liquefaction model for the higher starch concentrations, we used the subsite map at 80°C that we reported earlier (chapter 4). The three branch factors (determining the intensity of inhibition) were given values of 0.03 to express the low rates of hydrolysis at high substrate concentrations.
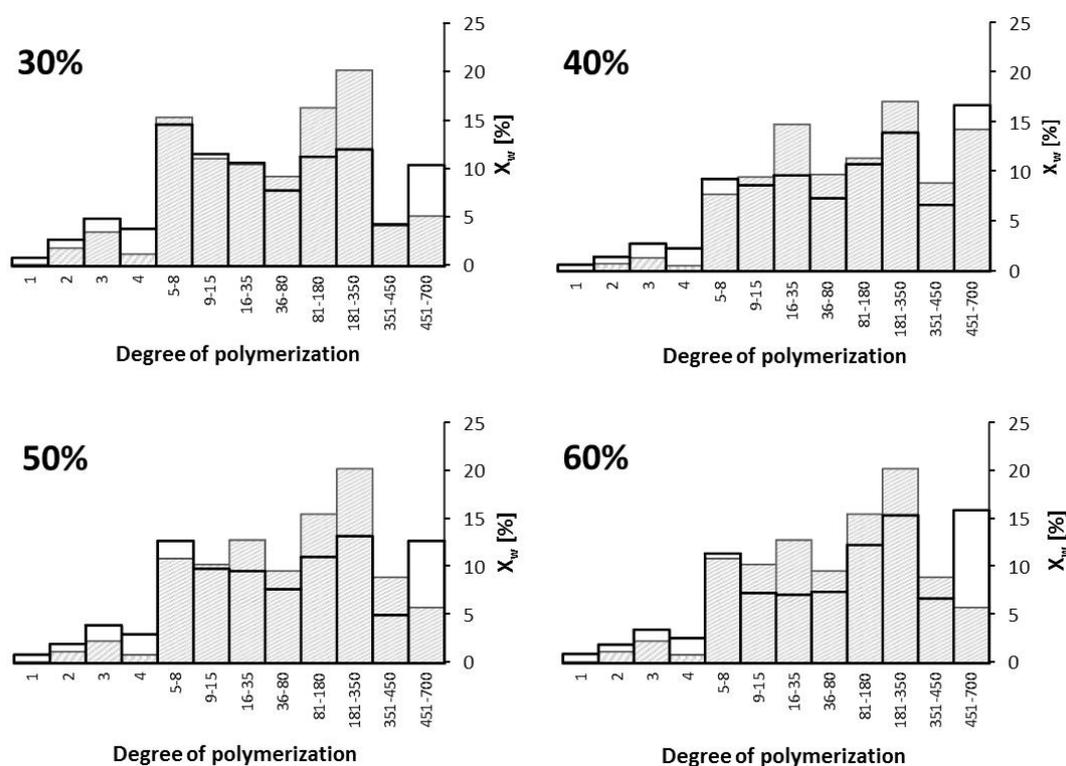
Figure 5.1. The final composition (w/w%) of starch liquefacts from the experimental hydrolysis at 5h (empty bars with black outline) compared with the final outputs of the liquefaction model (filled grey bars). The total values of the residuals (SRSS) at each substrate concentration were 223, 219, 219 and 265 for 30, 40, 50 and 60 w/w% respectively.

The predictions of the liquefaction model corresponded well with the experimental data at the end of the reaction (Figure 5.1). Despite the increase in the initial substrate concentration of the samples, the subsite map allowed for accurate predictions of the oligosaccharides. Most differences between the model and the data appeared in the concentrations of the larger polysaccharides. At 30, 50 and 60 w/w% the groups of DP 81-180, 181-350 and 451-700 were predicted least accurately. At 40 w/w% the predictions of DP 16-35 and 181-350 were least accurate. The former was

also an issue at 60 w/w%. Since the model and the experimental data at all concentrations were in fairly good agreement, we concluded that we could use the created model matrices of maltodextrins as the substrate for the glucoamylase model.

Table 5.2 Summary of data from saccharification obtained experimentally and by modelling.

Initial substrate concentration [w/w%]	Starting DE of the saccharification experiments	Timespan of the liqefaction model	Starting DE - saccharifiaction model	Final DE after saccarification	Initial saccharification rate [$mol_{glucose} \cdot h^{-1}$]	Final glucose concentration [$mol \cdot L^{-1}$]
30	7.9	$3.24 \cdot 10^4$	8.1	75	0.45	1.35
40	9.2	$3.94 \cdot 10^4$	9.4	76	0.42	1.84
50	5.8	$2.12 \cdot 10^4$	5.7	55	0.42	1.44
60	5.7	$2.12 \cdot 10^4$	5.7	50	0.04	1.58

5.4.3. Modelling of saccharification by glucoamylase

5.4.3.1. Subsite maps of A. niger glucoamylase

Some of the independently obtained subsite maps of *A niger* glucoamylase that are reported in literature are similar. Based on the similarities between the subsite maps given in Table 5.1 we categorised them into two major groups: the 'older' subsite maps (Meagher et al. (1989), Ono et al. (1988)) and the 'newer' subsite maps (Ermer et al. (1993), Stoffer et al. (1997)).

In the 'older' subsite maps, subsite -1 is assigned a positive binding energy value and the remaining subsites are given negative energy values. In the subsite maps created using maltooligosaccharides (Meagher et al. (1989), Ono et al. (1988)) subsite +1 has a high negative value (on average -20 kJ·mol⁻¹), and is followed by a relatively high negative value assigned to subsite +2 (on average -6.6 kJ·mol⁻¹). The subsite maps of Meagher et al. (1989) based on isomaltooligosaccharides, differ by having higher positive energy value of subsite -1 and less negative value of subsites +1 and +2. In all the 'older' subsite maps, the binding modes that result in the lowest energy of binding do not overlap the catalytic site. This suggests that a large number of non-productive complexes might be formed in the model and fewer E-S bindings will result in the production of glucose (i.e. hydrolysis).

In the 'newer' subsite maps, nearly all subsites are assigned negative binding energy values. Binding modes covering the catalytic site are more probable in these subsite maps, because of the negative energy values of subsite -1. One of the subsite

maps of Ermer et al. (1993) and the subsite map of Stoffer et al. (1997) (both based on maltooligosaccharides) show large similarities. These subsite maps are analogous with the 'older' subsite maps, but mainly differ in the negative energy value assigned to subsite -1. The remaining subsite map of Ermer et al. (1993) does not show similarities to any of the discussed categories, perhaps because it was obtained using a different type of substrate.

Table 5.3. The average subsite maps used in the model based on the subsite maps from Table 5.1. The references point to the source of the subsite maps used to calculate these average subsite maps.

Name	Subsite binding energies [kJ · mol^{-1}] per subsite							Reference
	-1	+1	+2	+3	+4	+5	+6	
SM 1	2.6	-21.3	-6.6	-2.7	-1.2	-0.6	-0.7	Ono et al. (1988), Meagher et al. (1989)
SM 2	4.7	-15.2	-4.9	-0.5	-1.1	-0.3	0.0	Meagher et al. (1989)
SM 3	-1.6	-21.3	-7.1	-3.2	-0.5	-0.3	0.1	Ermer et al. (1993), Stoffer et al. (1997)
SM 4	-0.8	-9.6	-9.8	-6.2	-1.0	-1.2	-0.4	Ermer et al. (1993)

The model substrate matrices that were prepared with the liquefaction model for a specific concentration, were used as input for all simulations of that one substrate concentration. With the large number of subsite maps available we decided to test the different subsite maps in groups to verify whether they result in reliable predictions (Table 5.3). The glucose concentrations predicted by the model using the average subsite maps (Table 5.3) are shown in Figure 5.2.

As expected, the positive energy values assigned to subsite -1 in the subsite maps SM 1 and SM 2 imply that more non-productive complexes are formed in the model. To reach a DE of 75, each simulation requires around 66,000 productive complexes, but at that stage the number of non-productive complexes increases from just above a $1.2 \cdot 10^6$ for SM 3, up to $3.6 \cdot 10^6$ for SM 1 and even $6 \cdot 10^6$ for SM 2. The binding modes that overlap the catalytic site have the highest values of total binding energy in subsite maps SM 1 and SM 2 and are not favoured by the model. When subsite maps SM 1 and SM 2 are used in the model, the concentration of glucose increases steadily, but within the given number of events it does not exceed 60 w/w% or 40 w/w%, respectively (Figure 5.2). These subsite maps are therefore most likely to result in inaccurate predictions. With all subsites assigned negative energy values, the subsite maps SM 3 and SM 4 allow for the highest concentration of glucose within the given number of hydrolysis events.
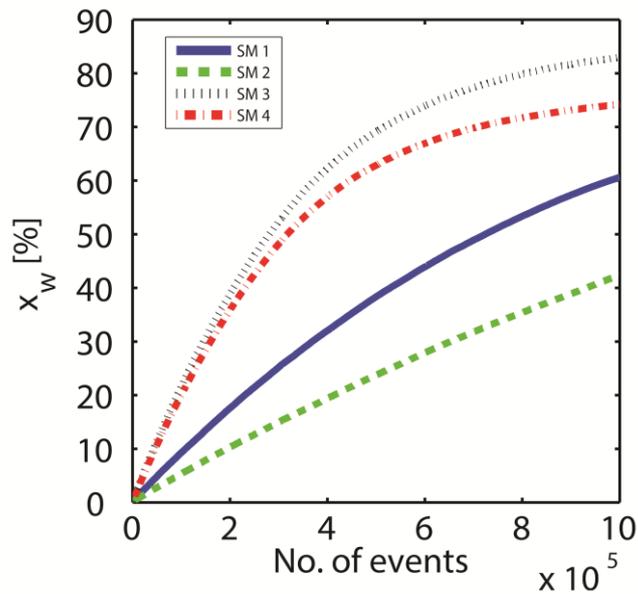
Figure 5.2 Outputs of the saccharification model computed using the maltodextrin matrix from the liquefaction model at 30% starch, with four average subsite maps from Table 5.3. Only the concentration of glucose is shown. The complete output is shown in the Appendix (Figure A1).

Thus the use of different substrates during subsite mapping of the enzyme influences the values of binding energies of the subsites, and therefore also the predictions of the product composition. If the aim of subsite mapping is to estimate the binding energies of the enzyme with its actual substrate (starch), then choosing the right kind of substrate for subsite mapping is crucial.

5.4.3.2. Saccharification model vs. the experimental data at 30 w/w% dry matter

The model outputs of saccharification with the four subsite maps from Table 5.3 were fitted to the experimental data collected from starch hydrolysis at 30 w/w%. The accuracy of the predictions was evaluated using both the graphical outputs (Figure 5.3) and the SRSS values (the Square Root of the Sum of Squares of the residuals). For the model outputs that used subsite maps SM 1 and SM 2 the number of events in the model needed to be increased threefold as compared to the subsite maps SM 3 and SM 4 to compensate for the large number of non-productive events.
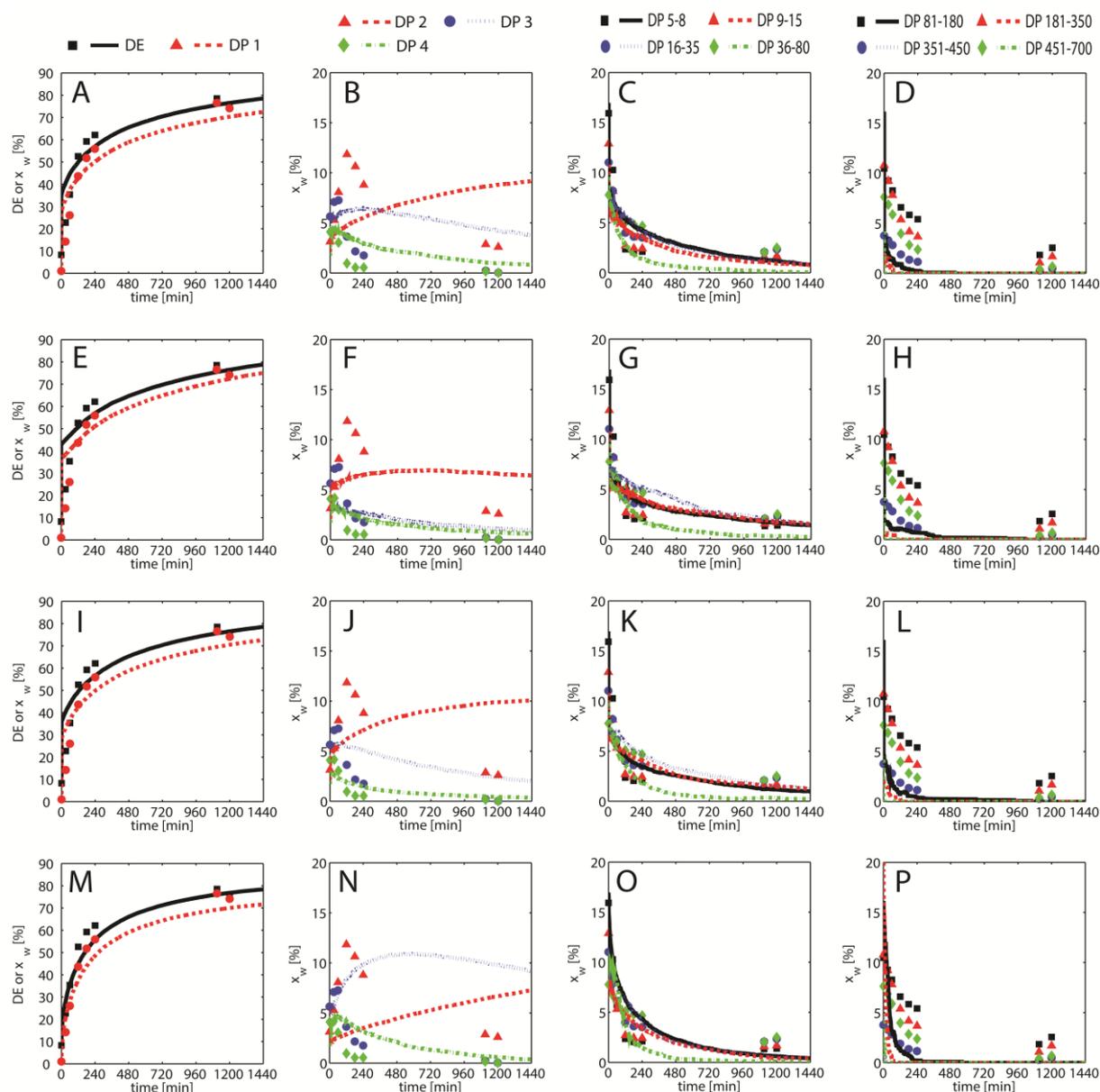
Figure 5.3 Model predictions using the four subsite maps from Table 5.3 fitted to the experimental data from saccharification of maltodextrins by *Aspergillus niger* glucoamylase at 30 w/w%. (A-D) SM 1; (E-H) SM 2; (I-L) SM 3; (M-P) SM 4. The legends for plots in each column are given on top. The points represent the experimental data and the lines represent the model simulations. Note: the plots in the first column have a different scale and represent both glucose concentration and the DE.

The saccharification model was very accurate. Repetitions of the simulation were nearly equal, because there was less randomness involved in only cleaving off the non-reducing end glucose units, compared to the action of α-amylase during the liquefaction. The saccharification reaction always follows the same pattern, in which one of the products is always glucose, and the size of the residual substrate reduces by 1 DP. The liquefaction model is different in this respect, because the large molecules can be

cleaved in a number of positions, leading to numerous possibilities and much larger product diversity between the simulations.

Based solely on the graphical outputs in Figure 5.3, reasonable fits between the data and the model are achieved with all of the subsite maps. The predictions of all molecules larger than DP 5 barely differ between the different subsite maps. It is somewhat surprising that the concentrations of maltose, maltotriose and maltotetraose (Figure 5.3 B, F, J and N) in the model vary more than the concentrations of glucose. For example, the more negative binding energies assigned to subsites +2 and +3 in SM 4 are responsible for the differences in concentrations of glucose, maltose and maltotriose between the simulations with subsite maps SM 3 and SM 4. The model outputs using SM 3 and SM 4 will resemble each other more, if the energy values assigned to subsites +2 and +3 in SM 4 are proportionally more similar to the values of the same subsites in SM 3, e.g., – 3.3 and -1 kJ·mol^{-1} respectively. The concentration of glucose will then increase and the amounts of maltose and maltotriose produced by the model will decrease.

Table 5.4 The differences, expressed as SRSS, between the simulated and the experimental data for saccharification samples at 30 w/w% of starch. Four subsite maps from Table 5.3 are compared.

DP	SRSS values per subsite map			
	SM 1	SM 2	SM 3	SM 4
DP_1	49.6	53.6	49.8	37.6
DP_2	27.9	19.9	22.0	34.7
DP_3	20.3	13.2	15.3	33.4
DP_4	12.7	9.5	8.2	13.9
DP_{5-8}	12.6	12.4	11.0	13.6
DP_{9-15}	9.3	11.1	10.1	9.4
DP_{16-35}	7.2	8.1	8.4	8.3
DP_{36-80}	14.8	12.8	11.6	16.1
DP_{81-180}	38.1	37.6	36.2	28.5
$DP_{181-350}$	40.2	40.5	39.5	38.1
$DP_{351-450}$	11.1	11.1	11.1	11.1
$DP_{451-700}$	25.5	25.5	25.5	25.5
Total	269.3	255.3	248.8	270.2

The differences between the modelled and the experimental data (SRSS) in simulations using the four subsite maps are given in Table 5.4. SM 3 with the lowest total SRSS of 249 is the subsite map allowing for the most accurate overall predictions. However, if the choice of the subsite map was based solely on the ability of the model to predict the concentration of glucose, subsite map SM 4 would be more accurate. The

more inaccurate predictions of maltose and maltotriose (Figure 5.3 N) increase the total residuals with subsite map SM 4.

The residuals of the model outputs with SM 1 and SM 3 are similar. The worse fit of DP 2, DP 3 and DP 4 with SM 1 is also evident in the higher total SRSS than that of the model using SM 3. The subsite map SM 2, although based on isomaltooligosaccharides, also allows for accurate predictions. The predicted concentrations of maltose and maltotriose are more alike the experimental data with this subsite map, however they are offset by the worse fit of glucose.

Based on the comparison of all the average subsite maps with the saccharification data form experiments at 30 w/w% dry matter content, subsite map SM 3 is most optimal, as it allows for most accurate predictions of the model and results in the lowest total value of SRSS.

5.4.3.3. Saccharification model and experiments at higher substrate concentrations

Subsite map 3 yields the most accurate predictions when the data from experiments at 40, 50 and 60 w/w% of starch are compared with the model outputs using the same four groups of subsite maps from Table 5.3.

Table 5.5 The differences between the model and the experimental data during saccharification at 40, 50 and 60 w/w% of substrate expressed as SRSS. Each model simulation was performed using SM 3.

DP	SRSS		
	40%	50%	60%
DP_1	43.7	29.1	14.4
DP_2	24.6	19.4	8.6
DP_3	11.0	7.2	12.1
DP_4	7.5	7.9	10.4
DP_{5-8}	15.4	12.6	19.6
DP_{9-15}	8.7	5.3	4.1
DP_{16-35}	4.7	22.4	26.6
DP_{36-80}	10.1	9.3	11.7
DP_{81-180}	28.1	17.7	12.9
$DP_{181-350}$	35.2	50.1	19.7
$DP_{351-450}$	11.8	21.3	11.1
$DP_{451-700}$	29.8	48.5	47.2
Total	230.6	251.0	198.4

At 40 w/w% substrate the saccharification reaction proceeds at a similar initial rate as at 30 w/w% or 50 w/w% of substrate (Table 5.2). The reaction rates of the three experiments deviate as the hydrolysis progresses, leading to differences in final

concentrations of glucose at each substrate concentration, with the highest glucose concentration reached at 40 w/w% starch (Table 5.2). Even though in the beginning of the reaction the concentration of maltose is quite high in our samples (even up to 15 w/w% at 40 w/w% dry matter), as the reaction progresses maltose is being hydrolysed and only around 5 w/w% of maltose remains when the reaction is stopped (Figure 5.4 B). Comparing the results with SM 3 to the experimental data at 40 w/w% (Figure 5.4 A-D) the total SRSS was equal to 231 (Table 5.5).

At 50 w/w% of starch, we observe a high initial reaction rate, but the reaction quickly slows down. The DE value reached at the end of the reaction is a third lower than the DE at 30 w/w% (Figure 5.3) or 40 w/w% (Figure 5.4). The lower rate of the reaction also reflects in the gradual hydrolysis of the largest molecules available for the enzyme in the mixture (Figure 5.4 G and H). The benefit of hydrolysis at 50 w/w% dry matter is that less maltose is formed in the reaction, which can result in a better yield of glucose if the reaction continues towards higher degree of conversion.

At 50 w/w% the system starts to behave differently. The reaction rate decreases, leading to long reaction times, while the concentrated, very viscous reaction mixture makes reproducible sampling complicated. However, the model using the SM 3 also at this substrate concentration predicts the products reasonably well. The total SRSS of 251 is reached when the model is fitted with the data of the 50 w/w% starch saccharification. The fit is less accurate compared with the experiments at other substrate concentrations (Table 5.5), because of discrepancies with the larger saccharides. However, the production of glucose, maltotriose (DP 3) and maltotetraose (DP 4) is predicted satisfactorily.

When an even higher amount of substrate is used during hydrolysis (60 w/w%) the reaction remains in the initial stage within the given time. The hydrolysis begins slowly at these conditions, with an initial reaction rate 10 times lower than in the other saccharification experiments. However, at 25 hours the dextrose equivalent reaches nearly the same value as at 50 w/w% starch, and the final glucose concentration is even higher than at 50 w/w% (Table 5.2). The rates of reaction reported by Curic et al. (1998) varied little when the substrate concentration was increased from 45 to 60 w/w%, but we do not observe this. The reaction rate clearly decreases as starch concentration increases from 50 to 60 w/w%.
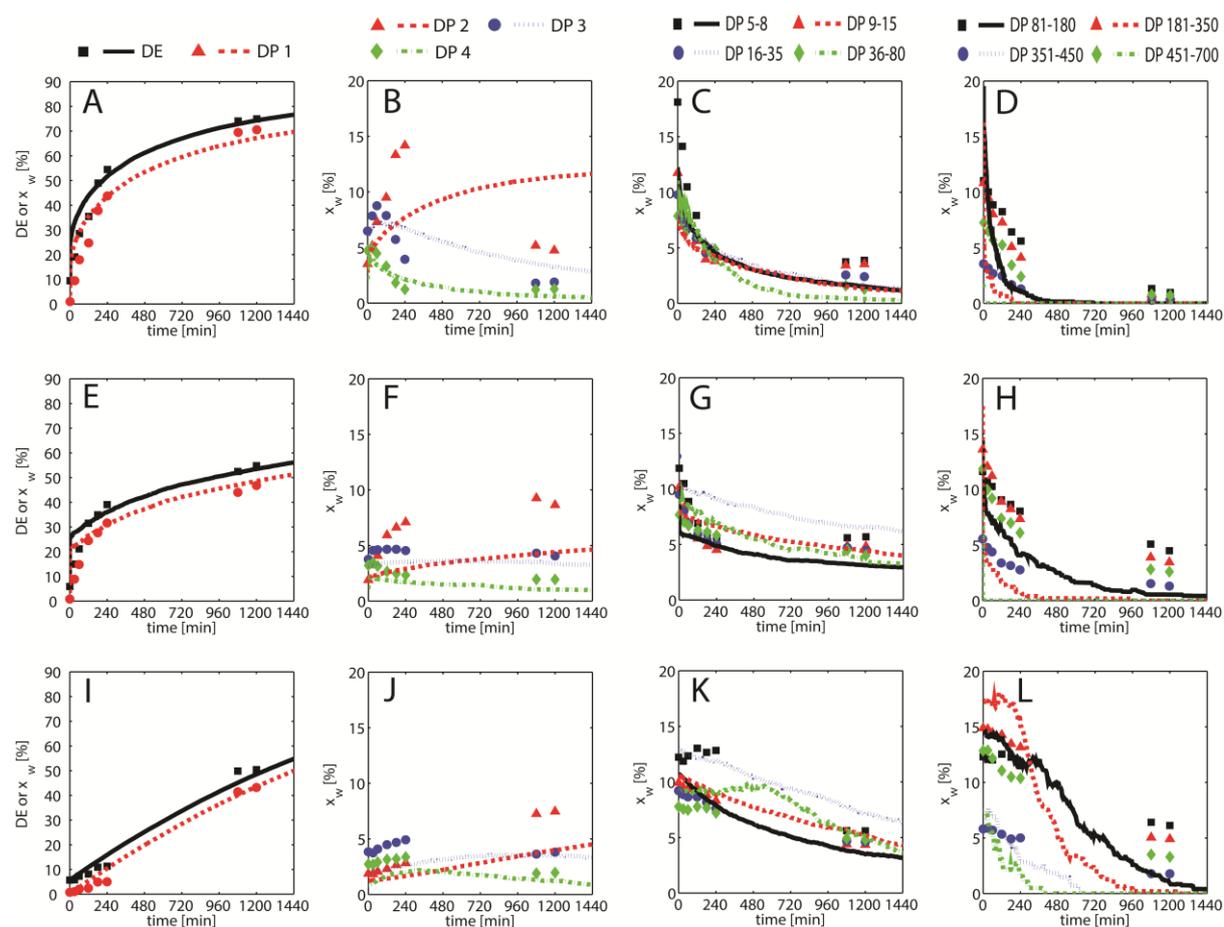
Figure 5.41 Model predictions using the average subsite map SM 3 fitted to the experimental data from saccharification of maltodextrins by *Aspergillus niger* glucoamylase at: (A-D) 40 w/w%, (E-H) 50 w/w% and (I-L) 60 w/w% dry matter content. The legends for each column are given above the plots, with points representing the experimental data and the lines representing the model simulations. Note: the plots in the first column have a different scale and represent both glucose concentration and the DE.

The results of van der Veen et al. (2005) show that lower glucose concentrations and a higher final degree of polymerization of products are common at increased dry matter contents. In our experiments, the glucose concentration decreases as the substrate concentration increases above 40 w/w%, but the other products (maltose, isomaltose, maltotriose) appear in lower quantities. Maltose and isomaltose are generally undesired products in the production of glucose syrup as they can decrease the final yield of glucose. The model predictions of maltose do not agree very well with the experimental data: most subsite maps allow for higher maltose production in the model than in the experiments (Figure 5.3 B, F, J, N).

The experimental DE values that are obtained at the different substrate concentrations drop from nearly 80 at 30 and 40 w/w% of substrate to 55 and 50 for 50 and 60 w/w% of substrate, respectively. The model of van der Veen et al. (2006)

proposed that the DE values are similar at all substrate concentrations (35-65%) after 24 hours of the reaction. Clearly, high DE values in the experiments at low moisture contents can only be reached at very long reaction times, or at high enzyme concentrations with a constant enzyme to substrate ratio. For our kinetic experiments it is better to work at a constant enzyme concentration, which is different from experiments where the goal is to reach the highest possible conversion.

Despite the differences in the rate of glucose production and in the final DE values, the model predicts the results of the experimental hydrolysis at 60 w/w% of starch quite well (Figure 5.4 I-L). Both the gradual degradation of the large substrates and the less rapid build-up of glucose are well described by the saccharification model. The model fits the data reasonably well, resulting in a low total SRSS value (Table 5.5).

Both the subsite map of Ermer et al. (1993) and that of Stoffer et al. (1997) were developed using maltooligosaccharides (DP < 9) as substrates. The maltodextrin derived from starch by enzymatic liquefaction are similar substrates to their maltooligosaccharides, with the addition of some larger, branched molecules. Based on our data, the maltooligosaccharides are relevant substrates for the estimation of binding energies of glucoamylases' subsites. In addition, using the same set of binding energy values we were able to predict the product distribution during hydrolysis with glucoamylase at different initial substrate concentrations.

The binding energies in the subsite map of BLA are assumed to be dependent on the temperature of the reaction (Kandra et al. 2006). It is not clear whether that is also the case for glucoamylase. The subsite mapping experiments of Ermer et al. (1993) and Stoffer et al. (1997) were conducted at 25 and 45°C respectively. It is likely that these authors used the same enzyme and thus came to nearly the same values of binding energies, irrespective of the 20°C temperature difference. We could successfully use these values for our predictions at 60°C, which shows that the energies in the subsite maps of glucoamylase from *Aspergillus niger* might not be strongly dependent on the temperature. Therefore, the conclusions from the hydrolysis by α-amylase form *Bacillus licheniformis* (chapter 4) that the binding energies do not differ substantially at different reaction temperatures, can be generalized to the glucoamylase originating from *Aspergillus niger*.

5.5.   Conclusions

The liquefaction and saccharification products in time during hydrolysis of wheat starch at high dry matter contents were predicted using a stochastic model based on the subsite theory. With more subsite maps emerging in literature and new methods available to estimate the binding energies (Gyémánt et al. (2002); Mótyán et al. (2011)), our models will become more and more useful in predicting the outcome of hydrolysis and thereby characterizing the mechanisms of the enzymes in more detail, even under extreme reaction conditions.

The saccharification model predicts the hydrolysis products without obtaining numerous parameters, but does not need inclusion of inhibition and condensation reactions, or restriction to specific conditions, e.g., enzyme concentrations. The hydrolysis products can be predicted by the model based solely on the subsite map. Also, by fitting just the values of DE in time, the composition of the products over time can be predicted.

Although both models were applied to completely gelatinized samples, we expect that these models can also predict the outcomes of reactions taking place at any enzyme-substrate ratio, and at any substrate concentration. The changes of the reaction conditions will reflect in the DE, which is used for fitting of the carbohydrate concentrations, and therefore are included in the model.

5.6.   Acknowledgements

5.7.   References

Åkerberg C, Zacchi G, Torto N, Gorton L. 2000. A kinetic model for enzymatic wheat starch saccharification. Journal of Chemical Technology & Biotechnology 75(4):306-314.

Allen JD, Thoma JA. 1976a. Subsite mapping of enzymes - application of depolymerase computer model to two alpha-amylases. Biochemical Journal 159(1):121-131.

Allen JD, Thoma JA. 1976b. Subsite mapping of enzymes - depolymerase computer modeling. Biochemical Journal 159(1):105-120.

Baks T, Bruins ME, Janssen AEM, Boom RM. 2007. Effect of pressure and temperature on the gelatinization of starch at various starch concentrations. Biomacromolecules 9(1):296-304.

Baks T, Kappen FHJ, Janssen AEM, Boom RM. 2008. Towards an optimal process for gelatinisation and hydrolysis of highly concentrated starch–water mixtures with alpha-amylase from B. licheniformis. Journal of Cereal Science 47(2):214-225.

Besselink T, Baks T, Janssen AE, Boom RM. 2008. A stochastic model for predicting dextrose equivalent and saccharide composition during hydrolysis of starch by alpha-amylase. Biotechnology and Bioengineering 100(4):684-97.

Cepeda E, Hermosa M, Ballesteros A. 2001. Optimization of maltodextrin hydrolysis by glucoamylase in a batch reactor. Biotechnology and Bioengineering 76(1):70-76.

Curic D, Karlovic D, Tripalo B, Jezek D. 1998. Enzymatic conversion of corn starch in twin-screw extruder. Chemical and Biochemical Engineering 12(2):63-71.

Davies GJ, Wilson KS, Henrissat B. 1997. Nomenclature for sugar-binding subsites in glycosyl hydrolases. Biochemical Journal 321(Pt 2):557-559.

de Cordt S, Hendrickx M, Maesmans G, Tobback P. 1994. The influence of polyalcohols and carbohydrates on the thermostability of α-amylase. Biotechnology and Bioengineering 43(2):107-114.

Ermer J, Rose K, Hubner G, Schellenberger A. 1993. Subsite affinities of Aspergillus-niger glucoamylase-II determined with P-nitrophenylmaltooligosaccharides. Biological Chemistry Hoppe-Seyler 374(2):123-128.

GonzalezTello P, Camacho F, Jurado E, Guadix EM. 1996. A simple method for obtaining kinetic equations to describe the enzymatic hydrolysis of biopolymers. Journal of Chemical Technology and Biotechnology 67(3):286-290.

Grafelman DD, Meagher MM. 1995. Liquefaction of starch by a single-screw extruder and post-extrusion static-mixer reactor. Journal of Food Engineering 24(4):529-542.

Gyémánt G, Hovánszki G, Kandra L. 2002. Subsite mapping of the binding region of α-amylases with a computer program. European Journal of Biochemistry 269(21):5157-5162.

Hiromi K. 1970. Interpretation of dependency of rate parameters on the degree of polymerization of substrate in enzyme-catalyzed reactions. Evaluation of subsite affinities of exo-enzyme. Biochemical and Biophysical Research Communications 40(1):1-6.

Hiromi K, Nitta Y, Numata C, Ono S. 1973. Subsite affinities of glucoamylase: Examination of the validity of the subsite theory. Biochimica et Biophysica Acta (BBA) - Enzymology 302(2):362-375.

Kandra L, Gyemant G, Remenyik J, Hovanszki G, Liptak A. 2002. Action pattern and subsite mapping of Bacillus licheniformis alpha-amylase (BLA) with modified maltooligosaccharide substrates. Febs Letters 518(1-3):79-82.

Kandra L, Remenyik J, Gyémánt G, Lipták A. 2006. Effect of temperature on subsite map of Bacillus licheniformis α-amylase. Acta Biologica Hungarica 57(3):367-375.

Lee CG, Kim CH, Rhee SK. 1992. A kinetic model and simulation of starch saccharification and simultaneous ethanol fermentation by amyloglucosidase and Zymomonas mobilis. Bioprocess Engineering 7(8):335-341.

MacGregor EA, MacGregor AW. 1985. A model for the action of cereal alpha amylases on amylose. Carbohydrate Research 142(2):223-236.

Meagher MM, Nikolov ZL, Reilly PJ. 1989. Subsite mapping of Aspergillus niger glucoamylases I and II with malto- and isomaltooligosaccharides. Biotechnology and Bioengineering 34(5):681-688.

Miranda M, Murado MA, Sanroman A, Lema JM. 1991. Mass transfer control of enzymatic hydrolysis of polysaccharides by glucoamylase. Enzyme and Microbial Technology 13(2):142-147.

Mótyán JA, Gyémánt G, Harangi J, Bagossi P. 2011. Computer-aided subsite mapping of α-amylases. Carbohydrate Research 346(3):410-415.

Nagy E, Belafibako K, Szabo L. 1992. A kinetic study of the hydrolysis of maltodextrin by soluble glucoamylase. Starch-Starke 44(4):145-149.

Nakatani H. 1996. Monte Carlo simulation of multiple attack mechanism of alpha-amylase. Biopolymers 39(5):665-669.

Nikolov ZL, Meagher MM, Reilly PJ. 1989. Kinetics, equilibria, and modeling of the formation of oligosaccharides from D-glucose with Aspergillus niger glucoamylases I and II. Biotechnology and Bioengineering 34(5):694-704.

Ono K, Shintani K, Shigeta S, Oka S. 1988. Comparative studies of various molecular species in Aspergillus niger glucoamylase. Agricultural and Biological Chemistry 52(7):1699-1706.

Polakovič M, Bryjak J. 2002. Modelling of the kinetics of thermal inactivation of glucoamylase from Aspergillus niger. Journal of Molecular Catalysis B: Enzymatic 19-20(0):443-450.

Rashid R, Jamaluddin H, Saidina Amin NA. 2006. Empirical and feed forward neural networks models of tapioca starch hydrolysis. Applied Artificial Intelligence 20(1):79-97.

Rosendal P, Nielsen BH, Lange NK. 1979. Stability of bacterial alpha-amylase in the starch liquefaction process. Starke 31(11):368-372.

Sanromán A, Murado M, Lema J. 1996. The influence of substrate structure on the kinetics of the hydrolysis of starch by glucoamylase. Applied Biochemistry and Biotechnology 59(3):329-336.

Stoffer BB, Dupont C, Frandsen TP, Lehmbeck J, Svensson B. 1997. Glucoamylase mutants in the conserved active-site segment Trp170-Tyr175 located at a distance from the site of catalysis. Protein Engineering 10(1):81-87.

Thoma JA, Brothers C, Spradlin J. 1970. Subsite mapping of enzymes - Studies on Bacillus subtilis amylase. Biochemistry 9(8):1768-&.

Thoma JA, Rao GVK, Brothers C, Spradlin J, Li LH. 1971. Subsite mapping of enzymes - correlation of product patterns with Michaelis parameters and substrate-induced strain. Journal of Biological Chemistry 246(18):5621-&.

van Boekel MAJS. 2008. Models and Modeling. Kinetic Modeling of Reactions In Foods: CRC Press.

van der Veen ME, van der Goot AJ, Boom RM. 2005. Production of glucose syrups in highly concentrated systems. Biotechnology Progress 21(2):598-602.

van der Veen ME, Veelaert S, Van der Goot AJ, Boom RM. 2006. Starch hydrolysis under low water conditions: A conceptual process design. Journal of Food Engineering 75(2):178-186.

van der Zalm EEJ, Berghout JAM, van der Goot AJ, Boom RM. 2012. Starch–gluten separation by shearing: Influence of the device geometry. Chemical Engineering Science 73(0):421-430.

Wojciechowski PM, Koziol A, Noworyta A. 2001. Iteration model of starch hydrolysis by amylolytic enzymes. Biotechnology and Bioengineering 75(5):530-539.

Zanin GM, De Moraes FF. 1996. Modeling cassava starch saccharification with amyloglucosidase. Applied Biochemistry and Biotechnology 57-58(1):617-625.

5.8. Appendix
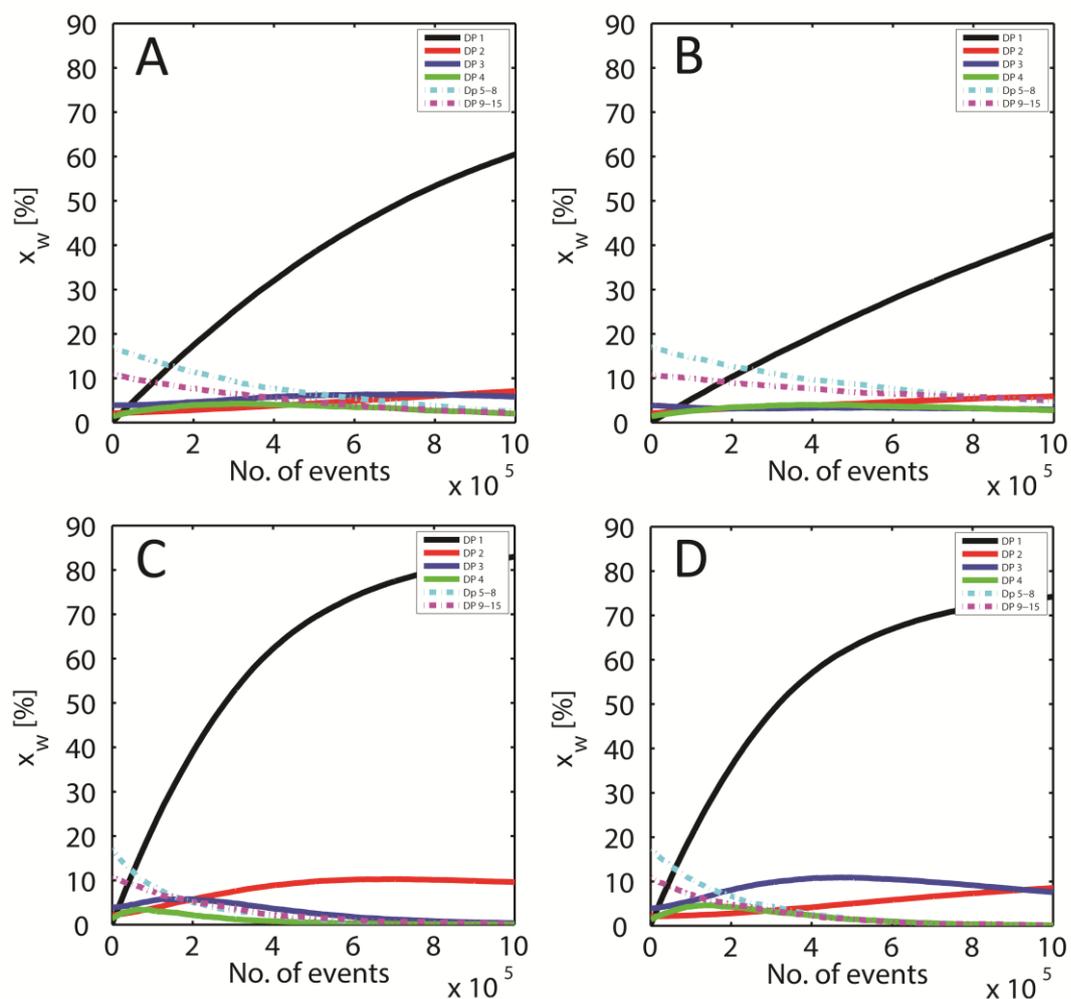
Figure A1. The predictions of the saccharification model (DP 1-15) computed using the maltodextrin matrix from the liquefaction model at 30% starch, with the four versions of the subsite maps given in Table 5.3: (A) SM 1; (B) SM 2; (C) SM 3; (D) SM 4.

General discussion

6.1. Outline of the thesis

The aim of the work described in this thesis was to create stochastic models that simulate the full extent of enzymatic hydrolysis of wheat starch. By comparing how wheat starch hydrolysis products develop over time at different experimental conditions, we also explored the mechanisms of the two industrially relevant enzymes used in starch processing: α-amylase form *Bacillus licheniformis* (BLA) and glucoamylase from *Aspergillus niger*.

In chapters 2, 3 and 4 the model that simulates the concentrations of all the hydrolysis products is described. In chapter 2 we discussed the gradual hydrolysis of wheat starch by the α-amylase (BLA) at 50°C. We confirmed that the reaction products do not appear in a random manner, but that they are the result of both the action pattern of the enzyme and the structure of the substrate. Subsequently, we described the stochastic model simulating the decomposition of starch into all of the hydrolysis products. The parameters available in literature were not adequate in accurately predicting the experimental product composition.

In chapter 3 we examined in detail whether by developing a new subsite map we were able to predict the products of starch hydrolysis at 50°C. We looked into the energy values assigned to each subsite to find out how the binding energy influences the simulated products. Subsites that are most relevant for accurate predictions were selected and their energy values were changed accordingly. As the subsite maps mostly influence the small hydrolysis products (< DP 8), we concluded that inhibition is the parameter that influences the hydrolysis of the large molecules in the solution. After testing several branch factors, we chose the set that resulted in the best fit between the modelled and experimentally collected data from chapter 2.

Wheat starch liquefaction at 80°C is described in chapter 4. First, the changes in product composition during the hydrolysis reaction were discussed. All the hydrolysis products were then simulated using the subsite map taken from Kandra et al. (2006), again showing a poor fit with the experimental data. Just as in chapter 3, a new subsite map and a set of branch factors were derived. Despite the differences in temperature, the interactions within the active site were not as different as the results of Kandra et al. (2006) might suggest.

After demonstrating that liquefaction of starch can be accurately described with the model at two different temperatures, the liquefaction at low moisture contents was studied. In chapter 5 we used the liquefaction model, with the parameters developed in chapter 4, to predict the composition of starch hydrolysates after 5 hours of hydrolysis at 30, 40, 50 and 60 w/w% of dry matter. The model results were compared with hydrolysis data, showing good agreement. However, the main focus of chapter 5 was to describe a stochastic model that can predict the products of the next step in starch hydrolysis – the saccharification. The maltodextrins produced during wheat starch liquefaction by BLA, were further hydrolysed with a glucoamylase from *Aspergillus niger*. Using the saccharification model, we tested several subsite maps found in literature and chose the one that best describes the data. The concentration of all hydrolysis products during the course of complete hydrolysis of starch could be well described using a single type of model.

In the next section, we will discuss the method for optimisation of the parameters of the liquefaction model that we used in chapter 5. After a motivation of the choice of the subsite map and inhibition values, we will verify our liquefaction model with independently collected data. Experimental data from literature, that represent wheat starch liquefaction by α-amylases from *Bacillus licheniformis* (BLA) at 5 and 60 w/w% of substrate and α-amylases from *Bacillus amyloliquefaciens* (BAA) at 5 w/w% of substrate, will be used to verify our choice of model parameters and the method used for their optimization. Further on, factors affecting the saccharification of maltodextrins at low moisture content will be discussed. We will conclude this chapter by describing the potential of mechanistic modelling and the applications of the models that we developed.

6.2. The liquefaction model at low moisture contents

In chapter 5 only the final results of the liquefaction model at different initial substrate concentrations were shown. The liquefaction model was used mainly to simulate the matrix representing the maltodextrins, which was used as the substrate for the saccharification model. Here we demonstrate how the parameters chosen for the liquefaction model at low substrate concentrations (subsite map, branch factors) were

evaluated for each of the experiments at increased dry matter content (30, 40, 50 and 60 w/w%).

The liquefaction model at 80°C (chapter 4) was adapted to predict starch hydrolysis products at higher dry matter concentrations (chapter 5). The timespan of the model was shortened (Table 5.2 in chapter 5). With a shorter timespan the reaction in the model reached the same final dextrose equivalent (DE) as the experimental data at each substrate concentration.

The results at low substrate concentrations (10 w/w%) were influenced by the binding energy values of the subsite map and by the inhibition (chapter 3, 4). The best subsite maps and branch factors from chapter 4 were tested against the experimental data from hydrolysis at high substrate contents (30-60 w/w%). With the low concentrations of DP 1-4 (<5 w/w%) and the large molecules being hydrolysed gradually, we expected that the branch factors would have more impact than the subsite map, since the inhibition is the major determinant of the hydrolysis of larger polysaccharides. Every stochastic simulation was executed five times and the average model output was compared with the experimental data using graphical outputs and the square root of the residual sums of squares (SRSS).

6.2.1. Subsite maps

The model with the original subsite map of BLA at 80°C (Kandra et al. 2006) could not correctly predict the composition of starch hydrolysates (chapter 4). However, the agreement between the predicted DP 1-8 and DP 1-8 from the experimental data (chapter 4) when using this subsite map was quite accurate in the initial stages of the reaction (< 1h). To ensure better description of the experiments during the whole observed reaction time, we changed the energy values of five subsites. Increasing the binding energy values of subsites -5 (from -11.2 to -7 kJ·mol^{-1}) and +3 (from -7.2 to -1 kJ·mol^{-1}) substantially improved the predictions of the model and led to more accurate results. Three additional smaller changes in the binding energies of subsite -4, -3 and +4 resulted in the final new subsite map (referred to as SM 1). Both subsite maps, the original one and SM 1, were used in the liquefaction model and compared with the data from starch hydrolysis at 30, 40, 50 and 60 w/w%.

Table 6.1 A comparison of the model predictions using two subsite maps: SM 1 and Org – the original subsite map of Kandra et al. (2006). The average values of SRSS are given. The simulations were repeated five times with branch factors set to 0.05.

DP	SRSS							
	30%		40%		50%		60%	
	Org	SM1	Org	SM1	Org	SM1	Org	SM1
DP_1	4.5	4.1	4.4	4.0	5.2	4.8	5.7	5.4
DP_2	8.3	3.4	7.0	2.8	8.6	4.2	8.1	4.7
DP_3	6.2	7.0	5.4	5.9	8.0	8.6	5.3	6.3
DP_4	10.3	10.6	8.5	8.6	10.3	10.5	8.0	7.9
DP_{5-8}	14.3	14.7	11.3	12.8	10.8	19.5	14.7	14.1
DP_{9-15}	8.3	4.3	10.6	6.3	16.0	12.6	17.9	14.5
DP_{16-35}	24.1	22.3	27.2	24.2	36.6	33.4	45.4	41.5
DP_{36-80}	36.6	36.8	38.0	39.6	32.3	29.2	43.7	39.5
DP_{81-180}	28.6	27.7	29.1	28.7	23.9	21.3	27.4	28.1
$DP_{181-350}$	20.1	23.8	20.0	24.3	28.5	26.7	16.1	18.8
$DP_{351-450}$	11.6	13.0	15.6	14.7	8.6	13.7	20.5	16.7
$DP_{451-700}$	75.3	73.3	85.2	83.8	53.7	57.1	90.4	89.0
Total	248.3	240.8	262.2	255.9	242.4	241.5	303.1	286.5

The simulations with the original subsite map predicted the outcomes of hydrolysis at high dry matter contents better than at 10 w/w% substrate concentrations (chapter 4). Despite a satisfactory description of the experimental data, the predictions using the original subsite map were not as good as the predictions using the new subsite map (SM 1). For all the starch concentrations, the lower residuals of DP 2 and DP 9-15 (Table 6.1) summed up to lower total SRSS values. Only at 50 w/w% the total SRSS's were similar, because the predictions of DP 5-8 were less accurate at this substrate concentration when SM 1 was used.

The predictions of DP 1-8 could not be further improved, however there was still potential for improvement in the predictions of the higher molecular weight polysaccharides. With the branch factors set to 0.05, the inadequate predictions of the large molecules, especially DP 16-35, 36-80 and 450-700, increased the total SRSS.

6.2.2. Inhibition (branch factors)

At the initial stages of hydrolysis, with the abundance of large molecules, the enzyme easily attacks the large, accessible targets. Only small amounts of oligosaccharides were produced within the reaction time we followed (Figure 6.1) and the energies assigned to the subsites will have less effect than the inhibition on the overall predictions.

The inhibition in the liquefaction model was based on the inability of the enzyme (BLA) to hydrolyse the α-(1,6)-glycosidic bonds and on the impairment of the hydrolysis when these bonds were in the proximity of the catalytic site. The inhibition influences the rate of hydrolysis indirectly, by increasing the frequency of non-productive hydrolysis events. A decrease of the values of the branch factors leads to stronger inhibition in the model. More non-productive complexes are formed and the rate of hydrolysis will be lower.

The influence of the branch factors on the model predictions during starch hydrolysis at 80°C was not as pronounced as in the outputs of the model at 50°C. As long as the branch factors were given values of 0.1 or less, the predictions of the hydrolysis model at 80°C were close to the trends of the experimental data. For our analyses at increased substrate concentrations, we designed the values of 0.1, 0.07, 0.05 and 0.03 to the branch factors with each of the subsite maps (Table 6.2). As the values of the branch factors decreased, the fit of the model improved (Table 6.2). Based on the SRSS values alone, the branch factors needed to be set to 0.03, as then the differences between the model and the data were the smallest for every starch concentration.

Table 6.2. Total SRSS values of computations using different branch factor values with two subsite maps (SM1 and the original subsite map) and for all of the substrate concentrations. The values in bold mark the lowest SRSS per experiment.

Branch factors	Total SRSS							
	30%		40%		50%		60%	
	Org	SM1	Org	SM1	Org	SM1	Org	SM1
0.1	300	301	312	308	313	312	372	377
0.07	278	269	284	279	281	281	330	325
0.05	248	241	262	256	242	242	303	287
0.03	234	**232**	245	**243**	**220**	228	272	**267**

The SRSS values given in Table 6.2 indicated that, in all but one case, the best agreements between the model and the experimental data was found with branch factors set to 0.03 and with SM 1. At 50 w/w% the two subsite maps yielded nearly identical results for each set of branch factors, and difference can only be discerned
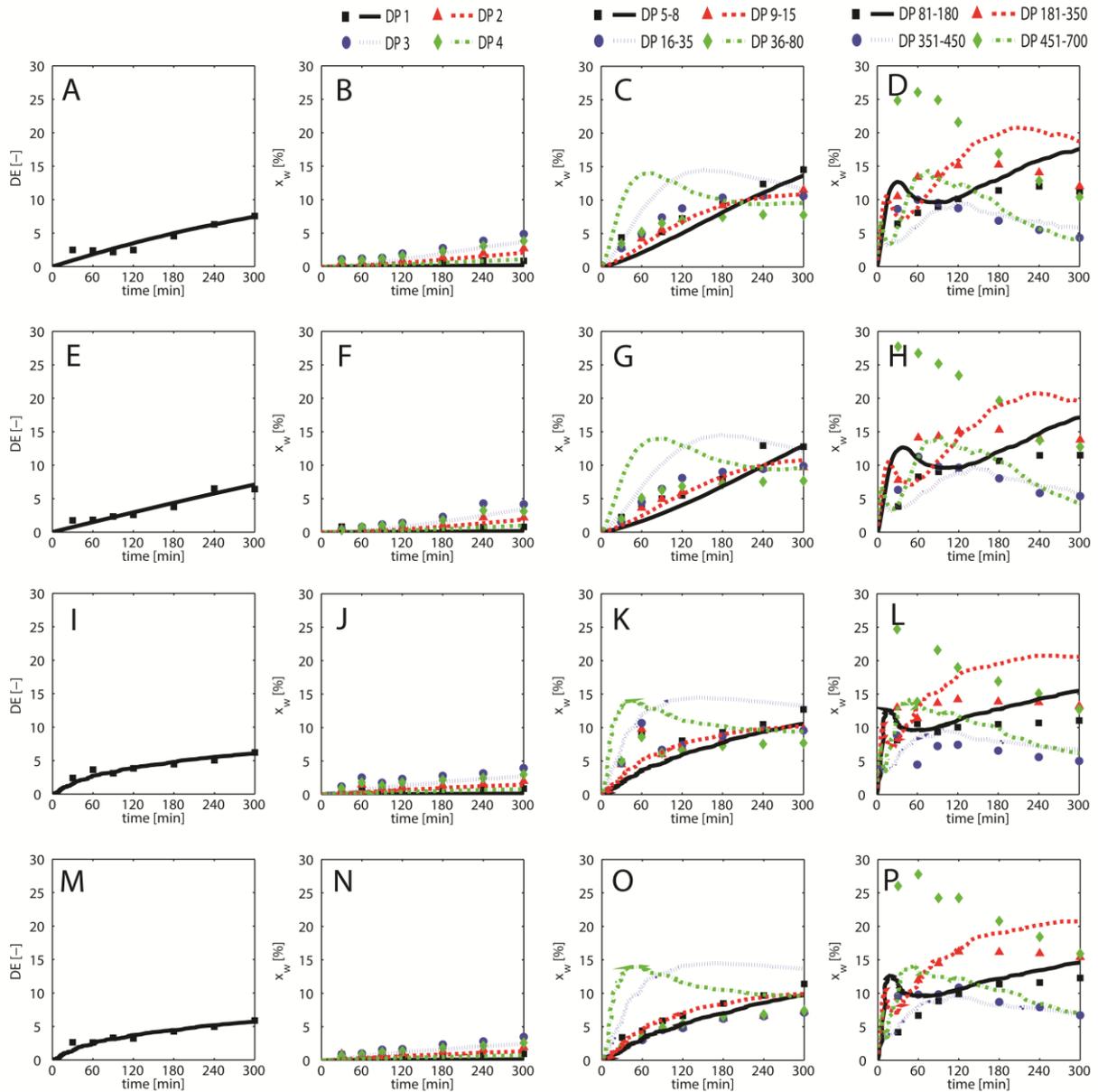
Figure 6.1 Outputs of the liquefaction model (lines) plotted against the experimental data (points) from starch hydrolysis by BLA at 80°C and at: (A-D) 30 w/w%; (E-H) 40 w/w%; (I-L) 50 w/w%; and (M-P) 60 w/w% of dry matter. Subsite map SM 1 and branch factors of 0.03 were used during simulations.

when the branch factors were set to 0.03. This was a result of a better fit of DP 351-450 in the computations using the original subsite map. The new subsite map (SM 1) described the oligosaccharides better at each substrate concentration, and therefore we graphically show the results of the model with SM 1 (Figure 6.1) and the branch factors set to 0.03. The differences between the experimental data sets at the four substrate concentrations were rather small, and the model fitted all data sets. With the branch factors set to 0.03, the somewhat lower concentrations of DP 16-35, 36-80, 81-180

predicted by the model correspond better with the experimental data (Figure 6.1) than at higher values of branch factors.

The final output of the liquefaction model was used as the substrate for the saccharification model. The plots in Figure 6.1 show that in the final stage of hydrolysis the model and the experimental data lead to similar outcomes. This implies that the maltodextrin matrices resulting from the liquefaction model are a good representation of the saccharification substrates.

6.3. Testing the liquefaction model and its parameters with independently collected data

The goal of our model was the simulation of all hydrolysis products, and therefore we chose a carbohydrate analysis method that identified and quantified molecules of all sizes. The method, however, had more limitations than the model we designed. The model can of course distinguish the molecules with a difference of 1 DP, and even differentiate between branched and linear products. An analytical method able to distinguish such subtle differences between the products, would be necessary to fully test our model. However, such detailed separation was not possible to achieve with the analysis method (HPLC-SEC) and the column we chose.

HPLC-SEC does allow the analysis of all the hydrolysis products, at the cost of a somewhat lower resolution. Additionally, the size-exclusion method separates on the apparent molecular weight, and cannot resolve between linear and branched molecules with the same degree of polymerization. Separation methods achieving higher resolution are available (e.g., HPAEC), but most of them fail to analyse larger polymeric molecules.

The composition of small oligosaccharides produced during wheat starch hydrolysis by BLA has been examined before (Baks et al. (2008), Besselink et al. (2008)) at a higher resolution. This allowed us to test the subsite maps we developed against these independently collected hydrolysis data.

6.3.1. Starch hydrolysis by BLA at 5 w/w% of substrate

Besselink et al. (2008) used the subsite map proposed by Kandra et al. (2006) in their predictions of hydrolysis products (DP 1-7) from 5 w/w% wheat starch hydrolysis by BLA at 50°C. The simulations were not accurate for any of the seven oligosaccharides.

Besselink et al. (2008) proposed a single change – the energy of subsite +3 was set to 0, which somewhat improved the predictions of DP 1, 2, 3 and 4.

The fit of the modelled oligosaccharides with the data by Besselink et al. (2008) (Figure 6.2 A-C) was not as good as with our own data in chapter 3. Our reaction took place at a higher substrate concentration (10 w/w% instead of 5 w/w%), but with the same enzyme dosage. At these conditions the conversion did not exceed a DE of 20, whereas Besselink et al. (2008) reached a much higher final DE (35). In addition, we did not separately consider the concentrations of DP 5, 6 and 7 when choosing the final subsite map. Yet, till the DE of 20 (90 minutes of the reaction), the predictions using our subsite map correspond well with the experimental data. The predictions of some of the oligosaccharides begin to deviate from the experimental data by Besselink et al. (2008), at conversions larger than we considered in our own experiments.
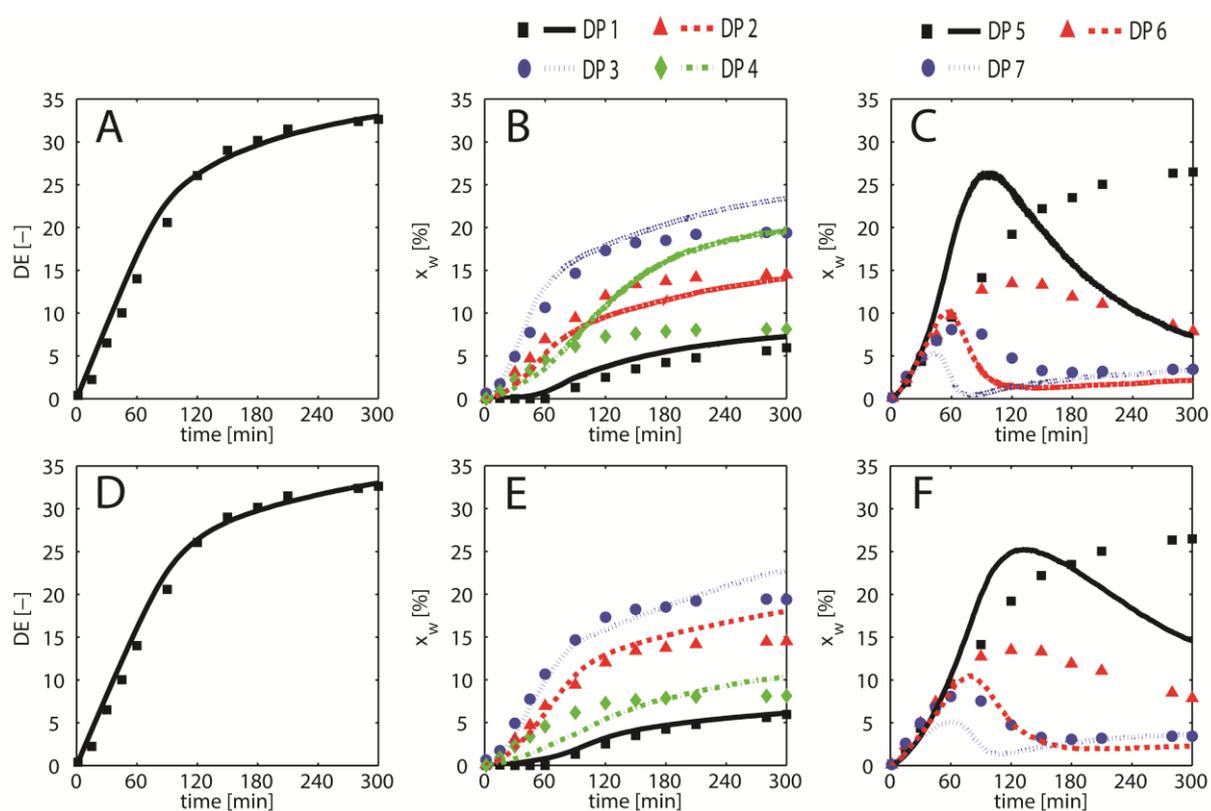
Figure 6.2 Data collected by Besselink et al. (2008) during hydrolysis of 5 w/w% wheat starch by BLA at 50°C compared with the model simulations using: (A-C) the new subsite map (chapter 3); (D-F) an updated subsite map for 50°C.

As we have tested a number of possible subsite maps before we applied the one given in chapter 3, we could now compare the model results using some of those subsite

maps with the data of Besselink et al. (2008). The key to a better description of the data by the model was in a decreased production of DP 4 and slower hydrolysis of DP 5, 6, 7.

Table 6.3. The binding energies [kJ·mol^{-1}] of subsites of *Bacillus licheniformis* α-amylase (BLA) at 50°C. The new subsite map developed in chapter 3 is shown, along with its updated version that better describes the data at higher degree of conversion. The changed binding energies are given in bold

Subsite map	Binding energy kJ·mol^{-1}								
	-5	**-4**	**-3**	**-2**	**-1**	**+1**	**+2**	**+3**	**+4**
New subsite map at 50°C (Chapter 3)	-8.0	-8.0	-5.1	-6.5	0.0	0.0	-5.1	-3.0	5.0
Updated subsite map at 50°C	-8.0	**-4.0**	-5.1	-6.5	0.0	0.0	-5.1	**-1.0**	5.0

The subsite map that predicted the data of Besselink et al. (2008) best (Figure 6.2 D-F) did not differ substantially from the one we originally proposed (Table 6.3 and Table 6.4). To reduce the concentration of DP 4 the energy of subsite -4 has to be less negative (-4 kJ·mol^{-1} instead of -8 kJ·mol^{-1}). Initially we decreased the energy value of that subsite, because in the beginning of the hydrolysis, the concentration of maltotetraose was more accurately predicted (also visible in Figure 6.2 B). Additionally we increased the energy of subsite +3 from -3 kJ·mol^{-1} to -1 kJ·mol^{-1}, to compensate for the increase in the concentration of DP 3 that occurs after the change in subsite -4.

Table 6.4 The residuals (SRSS) of the experimental hydrolysis data at 5 w/w% of wheat starch and the model computations using the new subsite map we proposed in Chapter 3 or the updated version of the subsite map, more suitable based on extensively hydrolysed wheat starch.

Subsite map	DP$_1$	DP$_2$	DP$_3$	DP$_4$	DP$_5$	DP$_6$	DP$_7$	Total
New subsite map	10.3	16.6	20.5	54.6	90.3	67.8	22.2	282.3
Updated subsite map	2.7	18.0	15.7	16.5	47.1	56.4	18.9	175.2

The concentrations of DP 1-4 (Figure 6.2 E) were predicted nearly quantitatively with the model using the updated subsite map. However, in the first 2 hours of the reaction (DE < 25), DP 4 fitted less accurately with the changed subsite map. This was the reason why we previously did not consider this subsite map as most optimal. The concentrations of DP 5-7 (Figure 6.2 F) were not predicted quantitatively, but the shapes of the curves and the proportions of DP 5 to DP 6 to DP 7 were as in the experimental

data. Changing the energy values of subsites did not improve the predictions of these groups anymore.

The lack of further improvement in the predictions of DP 5, 6 and 7 by subsite map changes may also convey that the increased production of these oligosaccharides was the result of transglycosylation. Another reason can be the structure of the subsite map itself. Recently, the presence of an additional subsite on the glycone side of the subsite map, subsite -6, was suggested by Tran et al. (2014). The predictions for DP 6 might be more accurate if one more subsite was considered.

We conclude, that an accurate subsite map can be obtained when the data are hydrolysed more extensively, reaching a higher DE value. Additionally, measurements of the extent of transglycosylation during the reaction could improve the choice of the subsite map.

6.3.2. *Starch hydrolysis at 60 w/w% dry matter by BLA*

Baks et al. (2008) hydrolysed 60 w/w% wheat starch at 90°C using BLA with an enzyme-to-substrate ratio of 0.1% (six-fold higher BLA concentration than in our analyses in Chapter 5). The composition of the oligosaccharides (DP 1-7) was analysed with a higher resolution. Our liquefaction model with the subsite maps at 80°C and 100°C (Kandra et al. 2006), and our new subsite map for 80°C from chapter 4 (Table 4.5) were compared to the data by Baks et al. (2008) in Figure 6.3.

Table 6.5 The binding energies [kJ·mol^{-1}] of subsites of *Bacillus licheniformis* α-amylase (BLA) developed by Kandra et al. (2006) for 80 and 100°C along with the new subsite map at 80°C that we proposed in chapter 4.

Subsite map	Binding energy kJ·mol^{-1}								
	-5	-4	-3	-2	-1	+1	+2	+3	+4
Kandra at 100°C	-5.9	-5.0	-3.1	-7.2	0.0	0.0	-17.5	-3.6	4.2
Kandra at 80°C	-11.2	-5.4	-6.6	-9.3	0.0	0.0	-5.6	-7.2	4.1
New subsite map at 80°C (Chapter 4)	-7.0	-7.0	-3.0	-9.3	0.0	0.0	-5.6	-1.0	5.0
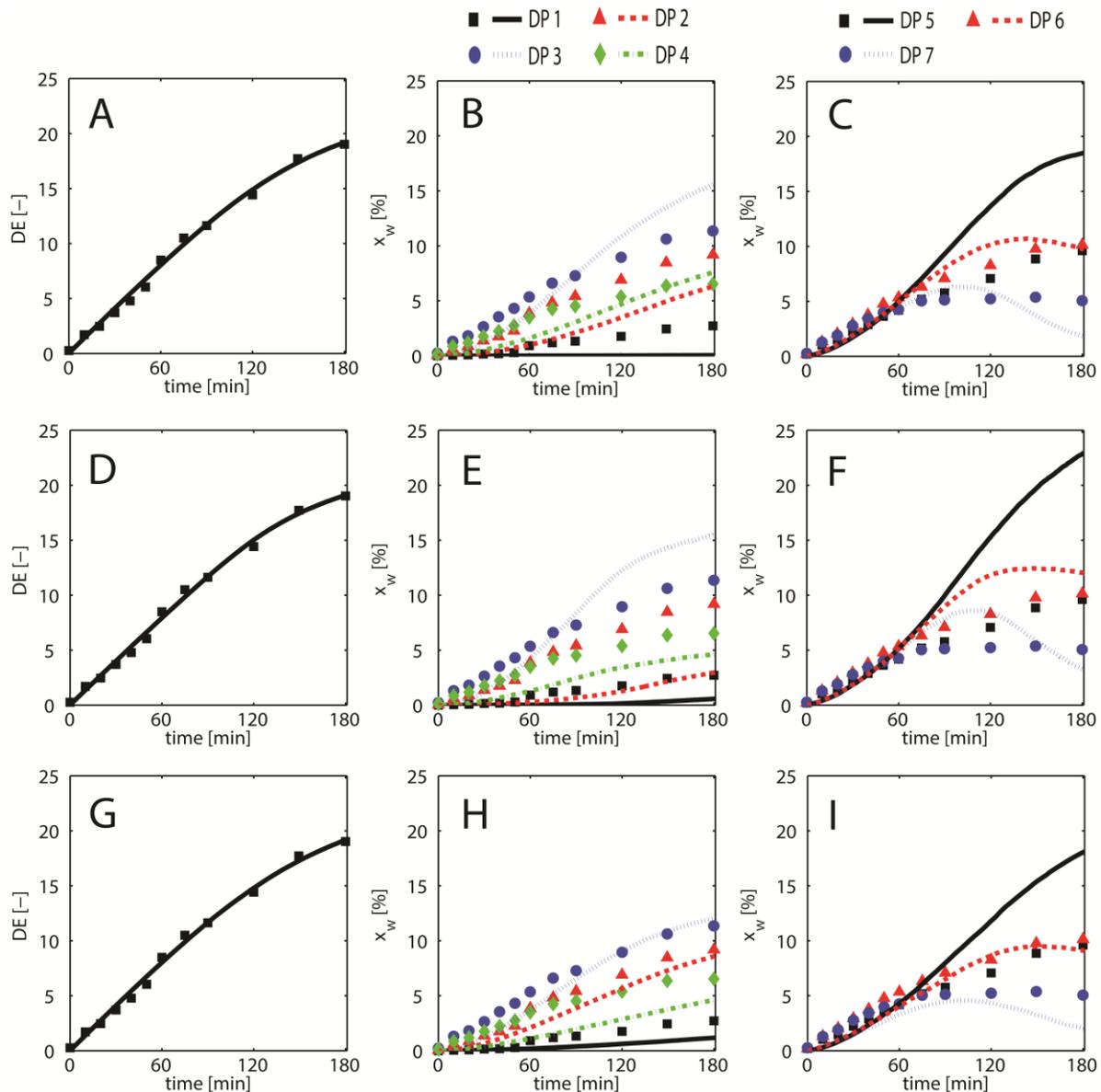
Figure 6.3 Data collected by Baks et al. (2008) during wheat starch hydrolysis by BLA at 60 w/w% dry matter at 90°C compared against the simulation of the liquefaction with: (A-C) subsite map of Kandra et al. (2006) for 100°C; (D-F) subsite map of Kandra et al. (2006) for 80°C; (G-I) new subsite map for 80°C from chapter 4.

The graphical representation of the model results (Figure 6.3) and the SRSS values (Table 6.6) both confirm that our new subsite map allows for more accurate predictions of the oligosaccharides than the remaining two subsite maps. With the subsite maps of Kandra et al. (2006), the final mixtures contain almost no glucose, the concentrations of DP 2 and 4 are too low and the concentrations of DP 3, 5, 6 and 7 are too high. The subsite map at 100°C results in better predictions than that of 80°C, because of the higher energy values of subsite -5 and lower of subsite +3 (Table 6.5). These are also some of the changes we suggested for our subsite map at 80°C. The lack

of glucose is the result of the very negative binding energy value (-17.5 kJ·mol^{-1}) assigned to the subsite +2 at 100°C.

Table 6.6 The residuals (SRSS) of the experimental hydrolysis data with 60 w/w% wheat starch and the model computations using the three subsite maps: two subsite maps proposed by Kandra et al. (2006) for 80 and 100°C and the new subsite map at 80°C that we proposed in chapter 4.

Subsite map	DP$_1$	DP$_2$	DP$_3$	DP$_4$	DP$_5$	DP$_6$	DP$_7$	Total
Kandra 100°C	11.0	24.7	18.8	13.3	31.9	9.7	11.1	120.5
Kandra 80°C	9.8	37.4	20.5	19.4	41.6	16.1	14.0	158.8
New subsite map 80°C	7.2	11.4	12.0	21.6	26.0	11.2	14.8	104.3

As we showed before, small differences might apply to subsite maps at different temperatures, hence some room for improvement of the subsite map at 90°C remains. DP 4 and 5 are strongly influenced by the energy of subsite -5. Lowering the binding energy of subsite -5 may further improve the predictions. However, this detracts in no way from the fact that our subsite map, although based on data obtained at a lower temperature and with lower resolution, simulates the concentrations of products more accurately.

6.3.3. Starch hydrolysis by BAA at 5 w/w% of substrate

The method we described for predicting the product concentrations over time using the subsite map can also be used to assess the subsite maps found in literature. We used data collected by Besselink et al. (2008) for α-amylase from *Bacillus amyloliquefaciens* (BAA) and the available subsite maps of this enzyme (Allen and Thoma (1976), Mótyán et al. (2011)). Besselink et al. (2008) already showed that the energies assigned to subsites in the subsite map of Allen and Thoma (1976) (Table 6.7) could not entirely predict the composition of the starch hydrolysis products. Figure 6.4 shows the computations using the subsite map of Allen and Thoma (1976) along with the outputs of the model with the subsite map given by Mótyán et al. (2011). Mótyán et al. (2011) suggested a different combination of subsites, with an extra subsite on the non-reducing end side and lack of the barrier site in subsite +4 (Table 6.7).

Table 6.7 The binding energies [kJ·mol^{-1}] of subsites of *Bacillus amyloliquefaciens* α-amylase (BAA) developed by Allen and Thoma (1976) and Mótyán et al. (2011) along with the new subsite map at 50°C that we propose.

Subsite map	Binding energy kJ/mol										
	-7	-6	-5	-4	-3	-2	-1	+1	+2	+3	+4
Subsite map of Allen & Thoma	-	-4.5	-10.2	-0.7	-4.2	-9.5	13.8	-14.4	-7.2	-4.0	5.3
Subsite map of Motyan	-4.0	-8.1	-13.2	-4.1	-7.6	-13.2	0.0	0.0	-10.7	-7.5	-
Proposed subsite map	-	-4.5	-10.2	-0.7	-4.3	-9.5	0.0	0.0	-4.0	-4.0	2.0

The predictions based on the subsite map developed by Mótyán et al. (2011), were more inaccurate for all oligosaccharides, except for DP 3 and 4 (Figure 6.4 F-H), than the predictions using the subsite map of Allen and Thoma (1976) (Figure 6.4 B-D).

In three simple steps, using the available subsite maps as initial guesses, we were able to find the subsites responsible for the inaccurate predictions and assign new energy values that allowed for better descriptions of the data. First, we looked at which oligosaccharides are predicted inaccurately and whether they are over- or underestimated by the model. Figure 6.4 shows that the new subsite map should allow for higher production of glucose, maltose and maltotetraose as compared to both available subsite maps, but should also balance the concentrations of maltotriose. Maltopentaose, maltohexaose and maltoheptaose were predicted quite well by the subsite map of Allen and Thoma (1976). Next, we compared the outputs using the available subsite maps and identified the subsites that will influence each of the oligosaccharides. Finally, we combined that information and, after just one attempt, proposed a new subsite map (Table 6.7).

We retained the positive energy value of subsite +4 in the new subsite map. This positive value along with the negative binding energies assigned to subsites +2 and +3 allowed for improvement in predictions of DP 1, 3 and 4 (Figure 6.4 J, K). The energies assigned to subsites -4, -5 and -6 are crucial for the correct quantities of DP 4 through 7 and remained unchanged. The small changes overall were sufficient for the model to show a similar trend to the experimental data, and in some cases close to quantitatively predict the hydrolysis products (DP 1, 3 and 7). The total SRSS values were now much lower (Table 6.8). The residuals of all of the oligosaccharides were lower than with the

subsite map of Mótyán et al. (2011). However, DP 5 and DP 6 were described less satisfactorily compared to the subsite map of Allen and Thoma (1976).

Table 6.8 The residuals (SRSS) of the experimental data of wheat starch hydrolysis by BAA and the model computations using the three subsite maps: the subsite map proposed by Allen and Thoma (1976), Mótyán et al. (2011) and the new subsite map we proposed for BAA.

Subsite map	DP_1	DP_2	DP_3	DP_4	DP_5	DP_6	DP_7	Total
Allen&Thoma	26.3	64.6	64.3	42.5	9.3	19.5	15.9	242.3
Motyan	30.4	99.2	25.2	24.5	42.6	43.5	123.9	389.2
New subsite map	16.5	72.4	10.8	23.7	25.4	44.6	15.6	209.1
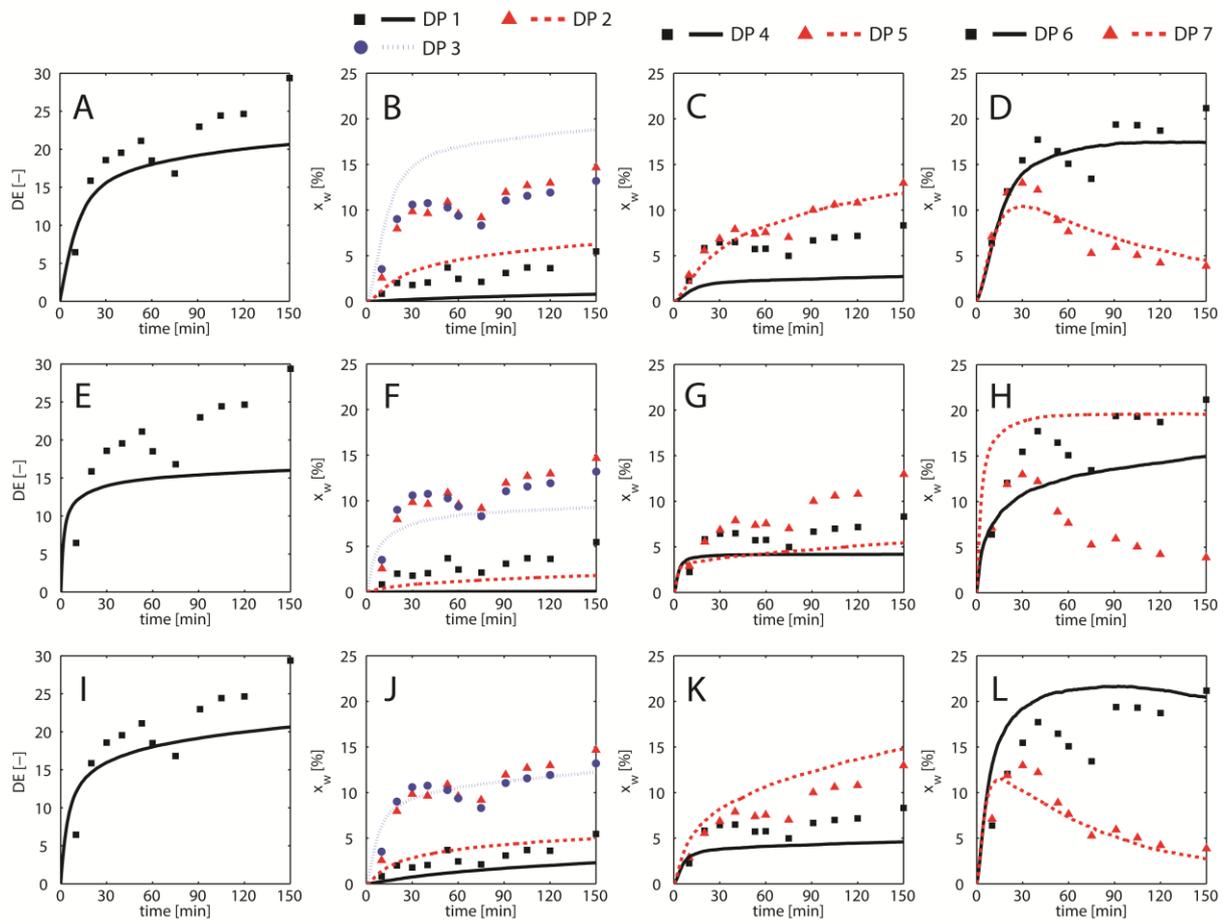
Figure 6.4 The predictions of the liquefaction model (lines) plotted against the experimental data (points) from wheat starch hydrolysis by BAA at 50°C (Besselink et al. 2008). The simulations used: (A-D) Subsite map of BAA from Allen and Thoma (1976); (E-H) Subsite map of BAA from Mótyán et al. (2011); (I-L) A new subsite map used in the model. All subsite maps are given in Table 6.7.

This new subsite map was not fully optimized to predict all products accurately. However, this case study demonstrates that one can obtain a new subsite map for

another α-amylase relatively easily by applying the assumptions and reasoning we developed for BLA.

The most accurate way of finding the energy values of a subsite map would be through multivariate analysis and simultaneous fitting of all the experimental data by minimizing the sum of squares. Despite our attempts, this was not possible to achieve on a regular PC, as too many parameters and data values needed to be simultaneously taken into account. The approach we finally chose, based on insight in the function and influence of the different parameters, while not exactly giving the globally best fit with the experiments, did lead us to conclusions others were not able to draw. One of these findings was the identification of the subsites that truly have an impact on the carbohydrate composition. We demonstrated that our model, method of obtaining the parameters and gathered insight from two hydrolases can easily be extended to new sets of experimental data and even other enzymes.

6.4.   Factors affecting saccharification at high substrate concentration

Next to collecting experimental data to validate our models, we studied the hydrolysis process itself and looked into factors that might affect the hydrolysis reaction (Figure 6.5). Increasing the concentration of the substrate can have a substantial effect on the progress of the reaction, as crowding of molecules and increased viscosity affects the mobility of the enzyme and its activity. The concentrated system as a whole is more prone to local differences caused by impaired mixing, including local low water concentration or increased glucose concentration. This local increase in the concentration of the reaction products may increase local product inhibition and may lead to reversed reactions. Over time, the rate of the reaction decreases not only because the substrate depletes or the enzyme is inactivated, but also due to the accumulation of branches, which are hydrolysed at a much lower rate than the α-1,4-bonds.

With a number of simple tests we show the influence of some of the discussed factors on the concentration of products during saccharification by glucoamylase from *A. niger*. Some of these tests included commercial maltodextrins as substrates, since they are better soluble in water. Commercial maltodextrins were used as a model system for comparison with our own samples.
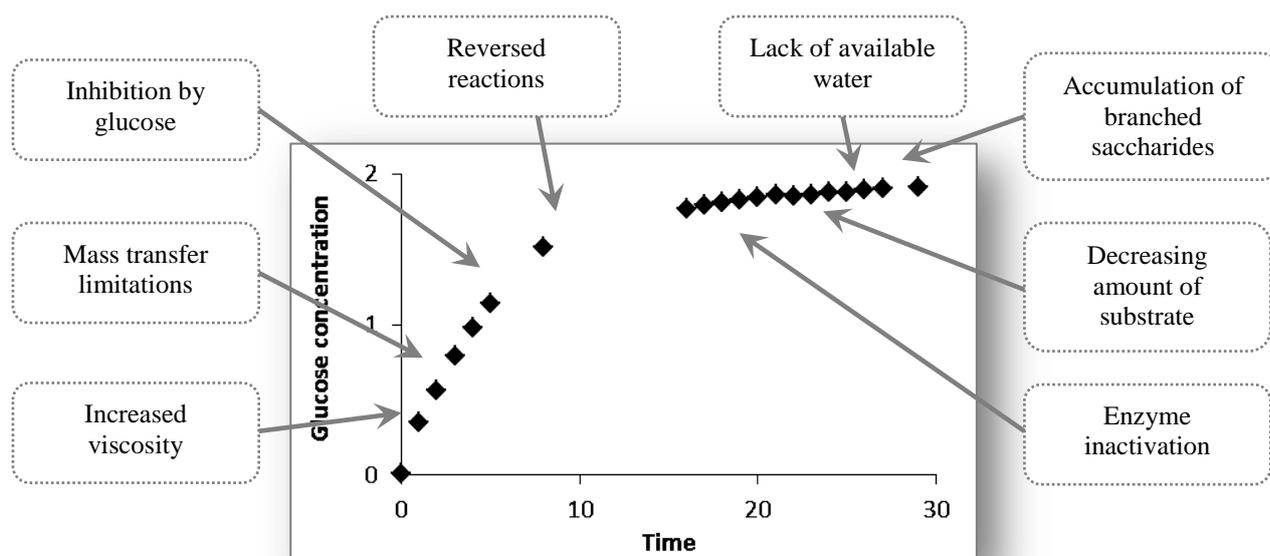
Figure 6.5. A summary of factors influencing the hydrolysis of maltodextrins by glucoamylase.

The influence of stirring speed on glucose production during hydrolysis of a 70 w/w% maltodextrins solution shows that mixing influences the production of glucose (Figure 6.6). However, with less vigorous mixing more glucose is produced in the reaction. This effect was also noticed earlier by Miranda et al. (1991), who explained that mixing affects the apparent viscosity of the solution, which in turn influences the reaction rate. The substrate composition will also play a role in this effect, because higher contents of amylopectin, that exhibits strong pseudoplastic behaviour (Miranda et al. 1991), will increase the shear-thickening behaviour of the starch suspensions.
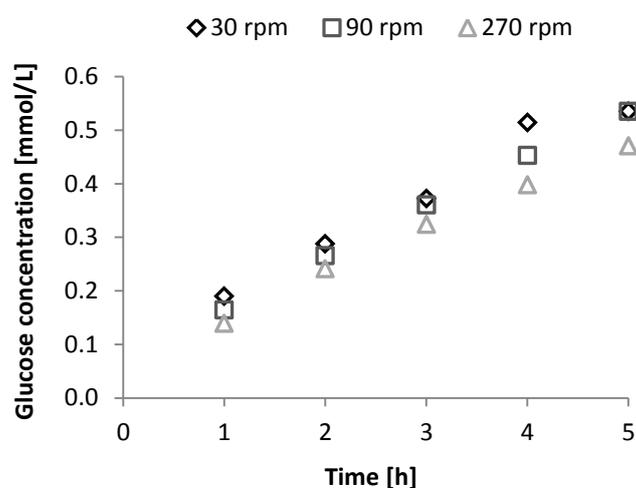
Figure 6.6 The influence of stirring speed on the production of glucose during hydrolysis of 70 w/w% commercial maltodextrins by glucoamylase.

A lower rate of hydrolysis after several hours of reaction signifies that either the reaction is approaching its final stages (most substrate has been converted into products) or that the enzyme was affected during the reaction through inactivation or inhibition. Adding a second dosage of enzyme after several hours of hydrolysis when the reaction rate becomes steady, demonstrates whether available substrate is still present in the reaction mixture and whether the first enzyme dosage underwent inactivation.
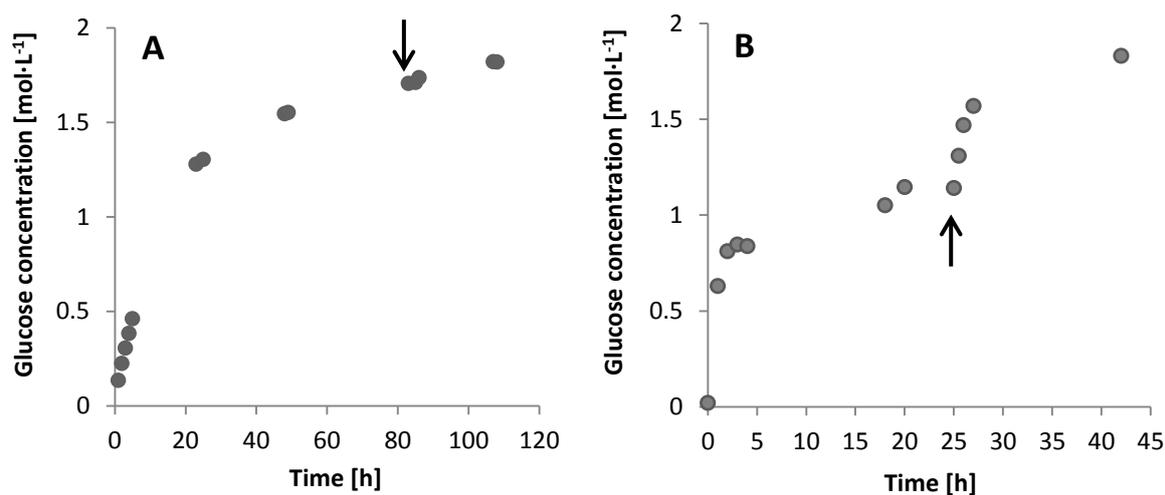
Figure 6.7 The changes in the concentration of glucose during hydrolysis of 70 w/w% commercial maltodextrins (A) or 50 w/w% liquefied starch (B) by glucoamylase from *A. niger*. A second dosage of enzyme, indicated by the arrow, has been added after several hours of hydrolysis.

After saccharification of commercial maltodextrins for 80 hours the reaction reaches 40% conversion (Figure 6.7 A). At that stage the majority of the substrate is still available for the enzyme. Activity measurements show that the glucoamylase retains more than 50% of its activity even after 48 hours of reaction (not shown). The addition of glucoamylase to the maltodextrin barely increases the glucose concentration, demonstrating that the high concentration of reaction products inhibits the enzyme and prevents further hydrolysis or that the equilibrium of the reaction is reached.

When the saccharification of liquefied starch (Figure 6.7 B) passes the 18 hours of reaction time, the reaction rate is also strongly reduced, and the concentration of glucose increases only slowly. A fresh dosage of enzyme at 25 hours leads to a surge in the reaction and boosts the production of glucose. The conversion degree increases from 46% at 25 hours to nearly 60% at 42 hours. Despite the initial rapid increase of the reaction rate, after 42 hours the curve of glucose concentration converges to a plateau again. Even if the hydrolysis takes place at optimal conditions, after prolonged

incubation the inactivation of the enzyme is inevitable. If fresh enzyme is supplied at a stage with sufficiently high conversion, the high concentration of reducing sugars will influence the enzymatic activity resulting in promoting reversed reactions or inhibition by reaction products.

Both enzymes used during hydrolysis (BLA and *A. niger* glucoamylase) cleave the branch points with more difficulty than the α-1,4-bonds. This means that as the hydrolysis reaction proceeds, the proportion of α-1,6-glycosidic bonds to α-1,4-glycosidic bonds will increase. This accumulation of branches will also affect the reaction rate.
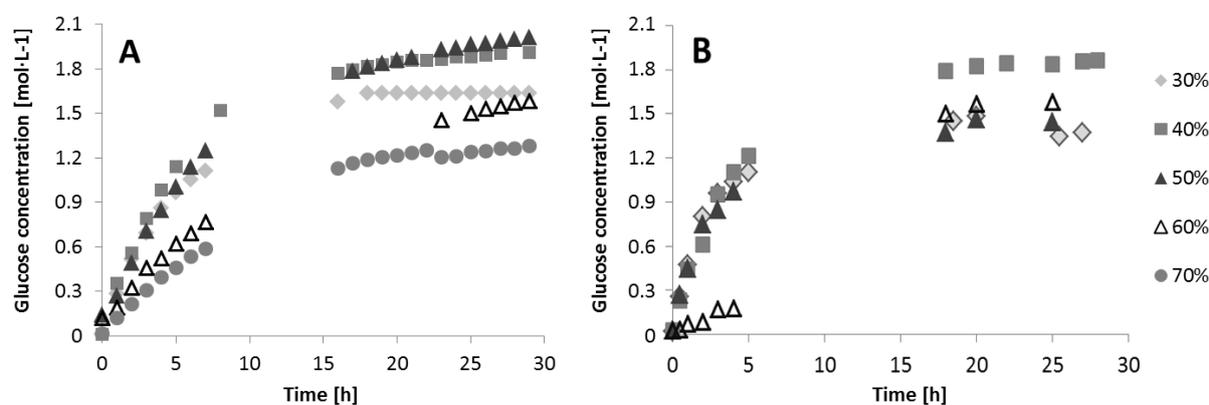
Figure 6.8 The changes in glucose concentration during hydrolysis of (A) commercial maltodextrins (30–70 w/w%) (B) wheat starch liquefacts (30–60 w/w%) by glucoamylase from *A. niger,* the same data as in chapter 5.

Apart from hydrolysing wheat starch, we also saccharified commercial maltodextrins at elevated dry matter contents (30–70 w/w%). The conditions of both reactions were as described in chapter 5. The production of glucose in time during the hydrolysis of both substrates is shown in Figure 6.8. At 30 w/w% of substrate, the hydrolysis reaches a high degree of conversion (88% for commercial maltodextrins and around 75% for wheat starch liquefacts), meaning that most of the substrate has been converted to glucose. As we increase the substrate concentration to 40 w/w%, also the concentration of glucose increases. The conversion degree is still high, reaching just above 76% for both substrates. The reaction rate in the beginning of the reaction at both concentrations is similar, and the higher concentration of substrate decides that more glucose produced at 40 w/w%.

At 50 w/w% of dry matter the reaction rate in the first hours is slower compared to the reaction at 40 w/w% (Figure 6.8). Even so, during the hydrolysis of maltodextrins, the final glucose concentration at 50 w/w% exceeds the glucose content at 40 w/w% (degree of conversion of 65%; Figure 6.8 A). In contrast, in the wheat starch samples at 50 w/w% the low reaction rate in the beginning of the reaction also leads to a lower production of glucose in the end (47% of conversion).

From 50 to 70 w/w%, increasing the concentration of substrate decreases the reaction rate both in the hydrolysis of commercial maltodextrins and in the hydrolysis of wheat starch liquefacts. In the maltodextrins samples (Figure 6.8 A) this lower initial rate leads to a decline in the production of glucose within the given time and enzyme dosage. The final concentration of glucose at 60 w/w% reaches around the same values as at 30 w/w% for both substrates (Figure 6.8 A and B), but this means that only 43% of substrate is converted to glucose. For wheat starch liquefacts this also shows that more glucose is produced at 60 w/w% than at 50 w/w% and a similar degree of conversion is reached despite the difference in substrate concentration.

At low starch contents (10-30 w/w%), the amount of generated glucose increases proportionally when the substrate concentration is increased (Cepeda et al. 2001). This is still true for the increase of substrate concentration from 30 to 40 w/w% (Figure 6.8). However, as the concentration of substrate increases further, the increase in the amount of produced glucose is no longer proportional to the increase in the substrate concentration.

At 50 w/w%, both during hydrolysis of commercial maltodextrins, and during hydrolysis of starch liquefacts, a change in the system occurs that affects the activity of the enzyme and influences the reaction rate. We propose that the difference between the reactions using the two substrates at 50 w/w% is caused by impaired mixing of the liquefacts compared to the mixing of commercial maltodextrins. Miranda et al. (1991) suggest that an increased substrate concentration might not necessarily lead to substrate inhibition as such, but rather impose mass transfer limitations or a change in thermodynamic activity. The increasing apparent viscosity of the solutions, as an additional obstacle for the enzyme, will influence the reaction rate at low moisture contents. Finally, also increased product condensation at higher substrate concentrations (van der Veen et al. 2005) will add to the lower production of glucose.

Designing a process of starch hydrolysis at high dry matter contents is a challenge, as more than just the right dosage of enzyme needs to be considered. Starting from choosing the right method and equipment for gelatinization, which will allow both for complete gelatinization and the transport of a highly viscous and adhesive starch slurry, all the way to choosing the right reactors for hydrolysis. We demonstrated that mixing has an impact on the reaction already on a small scale. Choosing the right mixing speed and stirrers will become even more crucial at increased reaction volumes.

Adding the α-amylase before gelatinization is not advised, as the enzyme will undergo inactivation by the application of shear, and therefore adding the enzyme in the final stage of gelatinization would be more optimal. The advantages would be complete gelatinization, which might not be achieved if the enzyme is added from the beginning, better mixing of the enzyme and substrate and finally easier transport of the mixture after partial liquefaction.

During the saccharification at high substrate concentrations, adding the enzyme in steps might be more beneficial for reaching the high degree of conversion, than adding the whole enzyme dosage at the start. This will not however prevent the occurrence of reversed reactions. Perhaps the already existing methods that deal with the reversed reactions, including removing the reaction products, will be sufficient also at the increased substrate concentrations.

6.5. Potential of stochastic modelling for other types of enzymes

We demonstrated in this thesis the possibilities of stochastic modelling for wheat starch hydrolysed by an α-amylase and a glucoamylase. Our aim was attained, as we designed models capable of following all of the products of the complete starch hydrolysis reaction, even at changing reaction conditions. We showed that mechanistic modelling also has potential for practical considerations, as it takes into account the properties of the system, the structure of the substrates and enzymes, and allows to explore the nuances of enzymatic hydrolysis.

We have come across subsite maps available for other hydrolases including α-glucosidase (Kita et al. 1991), β-glucosidase (Yazaki et al. (1997); Tahara et al. (1998)), xylanase (Biely et al. (1983); Meagher et al. (1988); Pollet et al. (2010)), galactanase (Bonnin et al. 1997) and polygalactouronase (Kluskens et al. 2005). With parameters, i.e.

subsite maps, already available in literature for different hydrolases, mechanistic models based on different enzyme-substrate couples can be designed without great effort. Such stochastic models were already developed for hydrolysis of cellulose (Levine et al. (2010); Kumar and Murthy (2013)), and even hydrolysis by proteases (Wierenga et al. 2014).

Mechanistic models help to extend our understanding of complex systems, for which the typical Michaelis-Menten kinetics are too simple. A mechanistic model accounts for more than only the substrate or enzyme concentration – it takes into account the structure of the substrate and the enzyme, which renders it more accurate, and better suitable for extrapolation to new conditions. Mechanistic models allow for the prediction of more than just the concentration of one product or substrate and are less limited to specific conditions of the reaction.

As the intensification of processes is a trend in the industry, reactions at high substrate concentrations will become inevitable. When stepping outside of the simple reaction kinetics or even representing simultaneous action of multiple enzymes, models will also need to account for more complexity, and this is where mechanistic models surpass the simple kinetic models.

6.6. Practical use of the model and the experimental findings

The starch hydrolysis models we developed can have several potential applications. As we mentioned in chapter 1, different mixtures of carbohydrates can have the same dextrose equivalent, which implies that the DE, representing just the overall degree of conversion, is insufficient to fully characterize hydrolysed starch. However, using only the DE values in combination with the subsite map of a particular enzyme, the full product profiles can be predicted for particular reaction conditions using our model. In this way, our model can be used to demonstrate the action patterns of enzymes without even conducting detailed experiments.

The models can also be used in experimental design, to tailor the hydrolysis products to specific needs. Using our model can help choosing the right enzyme, depending on the required characteristics of the final products, or perhaps even in engineering enzymes by directed mutagenesis. The simulations can also provide the information on when the reaction should be terminated, e.g., when presence of

particular oligosaccharides in the mixture is or is not required. The model allows to extrapolate the data, therefore if we are able to do experiments only at initial stages of reaction, the model can be used to demonstrate how the products will develop in the following stages of the reaction, as the DE increases.

The combination of the liquefaction and the saccharification model can be used to demonstrate the effects of the extent of liquefaction on the product profiles during saccharification. This way, the outcomes one would want to achieve during saccharification can be predicted without conducting any experiments. Using the model to indicate when the liquefaction should to be terminated for the required result during the saccharification to be achieved, can then significantly reduce the number of analyses.

Another application of the model could be testing whether different parameters chosen for the liquefaction affect the product combinations during saccharification. For example, replacing the subsite map of BLA with a subsite map of another enzyme can be used to determine which of the enzymes better fulfils the requirements of a particular process. The model results can even indicate, whether it is necessary to use a debranching enzyme, based on the quantities of branched molecules left after model hydrolysis. The choice of the suitable enzyme combination can therefore be made even before actually purchasing the enzymes.

Finally, the models are another way of characterizing hydrolytic enzymes. On top of providing the optimal temperature and pH, at which the reaction should take place, the information on the expected product composition from a particular substrate can be added. Since the type of the models we use are not limited to one enzyme-substrate combination, the potential for characterizing other hydrolases becomes endless.

The conceptual process design for gelatinization of starch in the shearing device and subsequent liquefaction in a batch reactor that we adapted from van der Veen et al. (2006) was used in our experiments at high substrate concentrations. We confirmed what the authors expected: proper starch gelatinization without adding enzyme can be achieved with a thermo-mechanical treatment in the shear cell. When the two processes (i.e. gelatinization and liquefaction) were separated, as also demonstrated for extruders (Grafelman and Meagher (1995); Baks et al. (2008)), the gelatinization was extensive and facilitated the liquefaction at low moisture contents. As we could also successfully follow the liquefaction reaction with the hydrolysis by glucoamylase, we support the

conclusions of van der Veen et al. (2006) about the validity of the design for hydrolysis at high dry matter contents.

The use of this process design in practice would require the addition of new type of equipment, as drastic changes in a process (i.e. doubling the substrate concentration) generally require more adaptations. However, we believe that sustainability of the process, after largely reduced water and energy consumption, along with the higher productivity that can be achieved at high substrate contents could compensate for the costs of the investment in the new equipment.

6.7. References

Allen JD, Thoma JA. 1976. Subsite mapping of enzymes - application of depolymerase computer model to two alpha-amylases. Biochemical Journal 159(1):121-131.

Baks T, Kappen FHJ, Janssen AEM, Boom RM. 2008. Towards an optimal process for gelatinisation and hydrolysis of highly concentrated starch–water mixtures with alpha-amylase from B. licheniformis. Journal of Cereal Science 47(2):214-225.

Besselink T, Baks T, Janssen AE, Boom RM. 2008. A stochastic model for predicting dextrose equivalent and saccharide composition during hydrolysis of starch by alpha-amylase. Biotechnology and Bioengineering 100(4):684-97.

Biely P, Vršansicá M, Gorbacheva IV. 1983. The active site of an acidic endo-1,4-β-xylanase of Aspergillus niger. Biochimica et Biophysica Acta (BBA) - Protein Structure and Molecular Enzymology 743(1):155-161.

Bonnin E, Vigouroux J, Thibault JF. 1997. Kinetic parameters of hydrolysis and transglycosylation catalyzed by an exo-β-(1,4)-galactanase. Enzyme and Microbial Technology 20(7):516-522.

Cepeda E, Hermosa M, Ballesteros A. 2001. Optimization of maltodextrin hydrolysis by glucoamylase in a batch reactor. Biotechnology and Bioengineering 76(1):70-76.

Grafelman DD, Meagher MM. 1995. Liquefaction of starch by a single-screw extruder and post-extrusion static-mixer reactor. Journal of Food Engineering 24(4):529-542.

Kandra L, Remenyik J, Gyémánt G, Lipták A. 2006. Effect of temperature on subsite map of Bacillus licheniformis α-amylase. Acta Biologica Hungarica 57(3):367-375.

Kita A, Matsui H, Somoto A, Kimura A, Takata M, Chiba S. 1991. Substrate-specificity and subsite affinities of crystalline alpha-glucosidase from Aspergillus niger. Agricultural and Biological Chemistry 55(9):2327-2335.

Kluskens LD, van Alebeek G-JWM, Walther J, Voragen AGJ, de Vos WM, van der Oost J. 2005. Characterization and mode of action of an exopolygalacturonase from the hyperthermophilic bacterium Thermotoga maritima. FEBS Journal 272:5464-5473.

Kumar D, Murthy G. 2013. Stochastic molecular model of enzymatic hydrolysis of cellulose for ethanol production. Biotechnology for Biofuels 6(1):1-20.

Levine S, Fox J, Blanch H, Clark D. 2010. A mechanistic model of the enzymatic hydrolysis of cellulose. Biotechnology and Bioengineering 107:37 - 51.

Meagher MM, Tao BY, Chow JM, Reilly PJ. 1988. Kinetics and subsite mapping of a d-xylobiose- and d-xylose-producing Aspergillus niger endo-(1→4)-β-d-xylanase. Carbohydrate Research 173(2):273-283.

Miranda M, Murado MA, Sanroman A, Lema JM. 1991. Mass transfer control of enzymatic hydrolysis of polysaccharides by glucoamylase. Enzyme and Microbial Technology 13(2):142-147.

Mótyán JA, Gyémánt G, Harangi J, Bagossi P. 2011. Computer-aided subsite mapping of α-amylases. Carbohydrate Research 346(3):410-415.

Pollet A, Lagaert S, Eneyskaya E, Kulminskaya A, Delcour JA, Courtin CM. 2010. Mutagenesis and subsite mapping underpin the importance for substrate specificity of the aglycon subsites of glycoside hydrolase family 11 xylanases. Biochimica et Biophysica Acta (BBA) - Proteins and Proteomics 1804(4):977-985.

Tahara N, Yano H, Yoshinaga F. 1998. Subsite structure of exo-1,4-β-glucosidase from Acetobacter xylinum BPR2001. Journal of Fermentation and Bioengineering 85(6):595-597.

Tran PL, Lee JS, Park KH. 2014. Experimental evidence for a 9-binding subsite of Bacillus licheniformis thermostable α-amylase. FEBS Letters 588(4):620-624.

van der Veen ME, van der Goot AJ, Boom RM. 2005. Production of glucose syrups in highly concentrated systems. Biotechnology Progress 21(2):598-602.

van der Veen ME, Veelaert S, Van der Goot AJ, Boom RM. 2006. Starch hydrolysis under low water conditions: A conceptual process design. Journal of Food Engineering 75(2):178-186.

Wierenga PA, Butre CI, Stoychev IG, Gruppen H. 2014. A simulation model to describe the hydrolysis of proteins by specific and a-specific proteases. Wageningen: Wageningen University, PhD thesis.

Yazaki T, Ohnishi M, Rokushika S, Okada G. 1997. Subsite structure of the β-glucosidase from Aspergillus niger, evaluated by steady-state kinetics with cello-oligosaccharides as substrates. Carbohydrate Research 298(1–2):51-57.

Summary

Starch, used for energy storage in many plants, is one of the most abundant polysaccharides. It is isolated from wheat, potatoes, corn and rice, and modified to yield ingredients for many industrial processes. Apart from the use in the food industry, starch and its derivatives are used in paper industry, in the production of adhesives and even in the textile industry.

The most common modification of starch is its enzymatic hydrolysis into maltodextrins or glucose. The products of these enzymatic conversions are used in the production of confectionery, beverages, bakery products or are fermented to yield new types of products. Two enzymes are commonly used for the hydrolysis of starch on industrial scale: α-amylase from *Bacillus licheniformis* (BLA) and glucoamylase from *Aspergillus niger*.

The aim of the work described in this thesis was to gain insight into the enzymatic hydrolysis of starch and the mechanism of the relevant enzymes used in the process. The aim was approached by designing models that allow the prediction of all of the hydrolysis products based on the quantification of the interaction between substrate and the subsites of the active centres of the enzymes.

Chapter 2 describes a re-evaluation and extension of an existing model of starch hydrolysis by α-amylases towards predicting all products of starch hydrolysis. The hydrolysis of amylopectin cannot be entirely random, because not all of the bonds are equally accessible for the enzyme. This was incorporated into the model. Comparing the model with collected experimental data on starch hydrolysis, showed that the prediction of the precise composition during the reaction was not yet satisfactory. The parameters of the model taken from literature did not allow for more satisfying predictions, and therefore new parameters were required.

In **chapter 3**, the subsite map, which reflects the action pattern of the enzyme, was shown to influence the production of the oligosaccharides in the model. The larger, polymeric molecules were less affected by the values assigned to the subsite map, but more by the inhibition in the model. This inhibition is based on the inability of the enzyme to hydrolyse the branch points of amylopectin and the bonds adjacent to the branch points. The predictions of the model improved substantially after creating a new subsite map and finding new branch factors (inhibition): all hydrolysis products present

in the mixture of hydrolysed wheat starch could be predicted adequately as a function of time.

With the potential of the model demonstrated using one set of experimental conditions, the predictive potential of the model towards a higher temperature could be explored. **Chapter 4** discusses the differences and the similarities in the product composition of the experimental data from starch hydrolysis at two temperatures, 50°C and 80°C. The differences between the two sets of data were insubstantial, which indicated that the parameters used for the higher temperature should be similar to those for the lower temperature. Indeed, the final subsite map and the inhibition parameter were rather similar for the two temperatures.

In **chapters 3** and **4** it was also demonstrated that not all of the subsites in the subsite map are equally important in making predictions. The binding energy values assigned to the outer subsites, rather than the inner ones, affect the carbohydrate composition during hydrolysis.

In **chapter 5** the complete wheat starch hydrolysis by α-amylase from *Bacillus licheniformis* and glucoamylase from *Aspergillus niger* was explored at elevated dry matter contents. Increasing the concentration of the substrates required a change in the methods used for gelatinization of starch and an addition of a shearing-heating step. The experimental data from the liquefaction step were again used in the model as described in the previous chapters. The liquefaction products found by the model were subsequently used as substrates in the saccharification of maltodextrins by glucoamylase. The core of chapter 5 was a new model for the saccharification of liquefied starch into glucose. Based on the predictions of the model, subsite maps were found that described the formation of all the products during the saccharification process.

Finally in **chapter 6**, the potential of the derived model was discussed. The modelling approach along with the parameters that were found, were shown to describe independently collected data as well. The model could also be extended towards other enzymes. The discussion was concluded with examples on where and how the modelling approach could be applied.

S

Samenvatting

In planten dient zetmeel als energieopslag. Zetmeel is een van de meest voorkomende polysacchariden. Zetmeel wordt geisoleerd uit graan, aardappelen, mais en rijst, en wordt gemodificeerd voor de toepassing in verscheidene industriële processen. Naast de toepassing in levensmiddelen, worden zetmeel en hiervan afgeleide producten gebruikt in de papierindustrie, bijvoorbeeld bij de productie van lijm, en zelfs in de textielindustrie.

De meest voorkomende modificatie van zetmeel is de enzymatische hydrolyse. Hierbij worden maltodextrines of glucose gevormd. Deze eindproducten worden gebruikt in de productie van snoepgoed, frisdrank en bakkerijproducten. Soms worden ze gefermenteerd, wat leidt tot nieuwe productsoorten. Voor de enzymatische hydrolyse van zetmeel op grote schaal worden voornamelijk α-amylase van *Bacillus licheniformis* (BLA) en glucoamylase van *Aspergillus niger* gebruikt.

Het doel van het onderzoek wat beschreven is in dit proefschrift was inzicht te verkrijgen in de enzymatische hydrolyse van zetmeel en het mechanisme van de enzymen die gebruikt worden in dit proces. Dit doel werd bereikt door een model te ontwerpen wat de concentratie van alle hydrolyse producten kan voorspellen. Dit model is gebaseerd op de interactie tussen substraten en de subsites in het actieve centrum van het enzym.

In **hoofdstuk 2** wordt een bestaand model voor de enzymatische zetmeelhydrolyse opnieuw bestudeerd en verder uitgebreid. Met deze uitbreiding kunnen alle producten van de zetmeelhydrolyse voorspeld worden. De hydrolyse van amylopectine is niet willekeurig, maar volgt juist een bepaald patroon. Dit wordt veroorzaakt door de mate waarin een binding beschikbaar is voor het enzym. Dit is dan ook geïmplementeerd in het model. De vergelijking van het model met de experimentele data van de zetmeelhydrolyse laat zien dat de voorspellingen van de productcompositie nog niet helemaal bevredigend zijn. De parameters van het model, die uit de literatuur naar voren kwamen, zijn niet accuraat genoeg om goede voorspellingen te maken. Er zijn nieuwe parameters nodig om de vorming van hydrolyseproducten te voorspellen.

Hoofdstuk 3, laat zien dat de 'subsite map' – die de werking van het enzym weergeeft, invloed heeft op de productie van de oligosachariden in het model. De waarden van de subsite map hebben weinig invloed op de grote moleculen. Deze zijn met name afhankelijk van de inhibitiefactoren in het model. Deze inhibitie wordt

veroorzaakt door het onvermogen van α-amylase om de 1-6 bindingen van amylopectine te hydrolyseren. Ook de bindingen die dichtbij zo'n vertakking liggen kunnen niet goed afgebroken worden door α-amylase. De voorspellingen van het model verbeterde aanzienlijk na de ontwikkeling van een nieuwe subsite map en nieuwe inhibitie factoren. De concentratie van alle hydrolyse producten in het reactiemengsel konden nu op elk tijdstip bepaald worden.

In **hoofdstuk 4** wordt onderzocht of het model ook gebruikt kan worden om voorspellingen te doen bij een hogere temperatuur. De verschillen en de gelijkenissen in de productsamenstelling van de experimentele data van de zetmeelhydrolyse op twee temperaturen (50 en 80°C) is onderzocht. De verschillen tussen de twee temperaturen waren niet substantieel. Dit suggereert dat de parameters die zijn gebruikt voor de hogere en de lagere temperatuur wellicht ook niet zo veel zullen verschillen. Het blijkt inderdaad, dat de uiteindelijke subsite map en de inhibitie factoren vrijwel gelijk worden bij beiden temperaturen.

In **hoofdstuk 3 en 4** wordt aangetoond dat niet alle subsites van de subsite map even belangrijk zijn om voorspellingen te doen. De waarden voor de bindingsenergie van de buitenste subsites hebben meer invloed op de koolhydraatsamenstelling dan de waarden van de binnenste subsites.

In **hoofdstuk 5** is de gehele tarwezetmeelhydrolyse door zowel α-amylase van *Bacillus licheniformis* als glucoamylase van *Aspergillus niger* onderzocht bij hoge droge stofgehalten. De toename in de substraatconcentratie vereist een verandering in de procescondities voor zetmeel gelatinisering. Het is noodzakelijk een hogere temperatuur en extra afschuifkrachten te gebruiken. De experimentele data van de reactie met α-amylase zijn vergeleken met de voorspellingen die met het model gemaakt konden worden. De uitkomsten hiervan zijn vervolgens als input gebruikt in een nieuw model voor de versuikering door de glucoamylase. De kern van dit hoofdstuk is het nieuwe model dat de versuikering tot glucose beschrijft. Gebaseerd op de voorspelling van het model, is de subsite map van glucoamylase verbeterd.

Tenslotte wordt in **hoofdstuk 6** de potentie van de verschillende modellen besproken. De modelmatige aanpak zoals beschreven in deze thesis, inclusief de optimalisatie van de modelparameters is goed bruikbaar om onafhankelijk verkregen datasets te beschrijven. Dit soort modellen kunnen ook toegepast worden op andere

enzymen. De discussie wordt afgesloten met voorbeelden van waar en hoe de modelaanpak toegepast wordt.

A

Acknowledgements

As much as we like to take credit for our own accomplishments, no one would be able to obtain a PhD degree if it was not for the efforts and caring of those who helped us along the way. Not often one gets the chance to say thank you in writing to people who gave the nudge in the right direction at a certain step of the journey towards the degree of the Doctor of Philosophy. There are no wrong or right ways to say thank you, therefore I choose to do it in 'the order of appearance', as many of those who I had the pleasure to cross paths with are equally important.

I would not be able to get to this moment without my parents who always let me be who I am and allowed me to make my own choices. Thank you for taking care of me for all these years and inspiring to work hard and to never settle for anything less than what makes me happy. This thank you also extends to my brothers (for 'having a laugh'), and my grandparents for always having me, my appetite and my best interest on their minds.

I doubt that this thesis ever reaches to any of my teachers, but I could not skip them in this list, because many of them said the right thing at the right time. Thanks to them I got on the path that led me to not only to science, but to the Netherlands, and eventually also to this journey towards receiving a PhD degree.

One person was there daily, through all my ups and downs – Jorik. Over the course of the past five and a half years we helped each other through tough times. Your humour and laid back approach were the cure for my neurotic, stressful moments. Thank you for being my best friend, for your support, for cheering on and for celebrating each milestone with me. Also for doing your best to understand what this booklet is actually about.

We go through life alone, but not without a support system. My support system in the Netherlands consisted of my Dutch family van de Laar & Harders. Thank you for the interest in my work, words of encouragement and for restoring my confidence when it needed to be restored. Two special girls, Thalia and Maya, have to be mentioned by name, as without their hugs, kisses and smiles my world would be dull. The second part of the support system were all the people I got to meet and befriend. Thanks for the dinners, parties, cinema trips, city trips, football tournaments, game nights, birthdays, barbecues and conversations that gave the research into what is important in life a high impact factor.

Nicolas, Klaus and Andreas for all the fun we had during the courses, conferences, meetings and of course in the evenings after.

A big thank you to Jordi and his graphic design skills, for putting up with my wacky ideas, actually bringing one of them to life and surrounding the pages of this booklet with a creative cover.

This gratitude list will be concluded by thanking all those who have gone past the list of propositions and read this thesis, first and foremost the examining committee and once more all of my supervisors. Thank you for the time and effort you have put in, all the helpful comments and remarks and most of all for being there on the day that finalizes these four adventurous years.

About the author

Karolina Bednarska was born in Poland on the 10th of June 1986 in Dąbrowa Tarnowska. She attended Liceum Ogólnokształcące im. Adama Mickiewicza in Tarnów, with the specialization in advanced biology and chemistry.

After graduation she begun studying Applied Biotechnology at the Agricultural University in Kraków. During the course of the studies she did an internship at P.Z. Cormay (Lublin, Poland) - a company producing medical analysers for blood analysis. Towards the end of her studies she participated in the Erasmus exchange program at Wageningen University. She finalized her studies by choosing the specialization Food and Beverage Biotechnology and by defending her master thesis on 'Enzymatic isolation of pectin from apple pomace' in the Department of Food Biotechnology.

After obtaining her Master-Engineer degree, she began her research at Wageningen University, Food Process Engineering Group. During her time as a PhD student she investigated the modelling and the hydrolysis of wheat starch by enzymes that resulted in this book.

Overview of the completed training activities

Discipline specific activities

Courses

Applied Enzyme Kinetics (DTU, Copenhagen, DK)	2010
Sustainability Analysis in Food Production (VLAG, Wageningen, NL)	2011
Food and Biorefinery Enzymology (VLAG, Wageningen, NL) - poster	2011
Automated Methods of Analysis (AUA, Athens, GR)	2011
Reaction Kinetics in Food Science (VLAG, Wageningen, NL)	2012

Conferences

NPS (Annual Dutch Process Technology Symposium), Papendal (NL) - poster	2011
FoodBalt conference (Kaunas, LT) - presentation	2012
ECCE conference (The Hague, NL) – poster	2013
LGF conference (Halkidiki, GR) – presentation and poster	2013
IUFoST Conference (Montreal, CA) – presentation and poster	2014

General courses

VLAG PhD week (Baarlo, NL)	2011
PhD seminar: Science Ethics (VLAG, Delft, NL)	2012
Global Food Production in a Changing World (Sun Yat-sen, Guangzhou, CN)	2012
Scientific writing (WGS, Wageningen, NL)	2013
Career orientation (WGS, Wageningen, NL)	2013
Scientific publishing (WGS, Wageningen, NL)	2013
Career assessment (Wageningen, NL)	2014

Optional activities

Excursion to the Netherlands (Visit to DSM, annual project meeting)	2011
Excursion to Greece (Visits to companies, annual project meeting)	2011
PhD study tour to China (Visits to companies and universities)	2012
PhD study tour to Finland and Baltic countries	2012
PhD study tour to Chile	2014
Workpackage meetings (LGF project meetings)	
Food Process Engineering meetings	