Unterschrift des Betreuers

TECHNISCHE UNIVERSITÄT WIEN
Vienna University of Technology

DIPLOMARBEIT

Novel GH family 3 enzymes from the thermophilic organism *Rhodothermus marinus*

Ausgeführt an der Faculty of Engineering am
Department of Biotechnology
der Lund University

unter der Anleitung von
Prof. Dr. Eva Nordberg Karlsson

eingereicht am Institut für
Verfahrenstechnik, Umwelttechnik
und Technische Biowissenschaften
der Technischen Universität Wien

zur Beurteilung durch
Prof. Dr. Christoph Herwig

von
Kulwant Kandra, BSc

Rainergasse 8/2/6
1040 Vienna
Austria

_____ _____

Datum Kulwant Kandra

The intellect is a bird; on account of its actions,
it is sometimes high, and sometimes low.
Sri Guru Granth Sahib

If we knew what it was we were doing,
it would not be called research, would it?
Albert Einstein

Acknowledgements

I would like to thank sincerely Prof. Eva Nordberg-Karlsson and Prof. Christoph Herwig for their great supervision during my master thesis. I want to thank you Eva for the chance you gave me to work at your lab, for the time you gave me when I had questions and your support, when I came to you with my ideas. I always enjoyed and learnt a lot during our discussions. I very much appreciate your warm and friendly welcoming nature, which made it easy for me to integrate into your research group, where I made good friends.

Thanks Christoph, for your support during my time in Sweden and your help for organizing and planning the project, helping me with all arrangements before my semester abroad, but also after I was back from Lund. Thanks for your time you always gave me, responding to my mails whenever I had questions or cheering me up when I was sometimes desperate with my experiments and for your patience you had with me. I am very grateful for all your great inputs, which helped me with my thesis. In addition, I also want to thank for all your motivation and understanding.

Zubaida, I want to especially thank you for being such a good lab-supervisor. You always helped me whenever I needed you. I had a great time with you and it means to me a lot that I was not only a student doing my master thesis, but also with time we got good friends, who had a nice time during the dark and cold days in Lund.

I also want to thank all the people from the Biotechnology department and special thanks to the members of the Fermentation group. Christina for the nice talks we had, Peter, Tania and Javier for their time they gave me when I had questions regarding some assays, Emanuel for his stories about Sweden and the world as well as Reza for the nice time we had, making jokes. Thanks all to you for making my stay in Sweden unforgettable with all the fika time. Also thanks to you Carl, for your help with LC-MS measurements and Patrick for your inputs on my report.

Finally, I want to thank my family, without their support I probably would not have come this far. You all are the reason for my success, because you always stood behind me. Thanks Vinnie, for your open ear, when I had concerns about my work. You have been a great support during my master thesis. Also thank you Kevin, for your jokes and encouragement, when I felt homesick. Thanks mum for your delicious energetic food during my sleepless nights throughout my studies and dad for your fresh bags with mixed nuts keeping my mind fresh as well as special thanks for your support always during my studies.

Abstract

The genes encoding a β-glucosidase (BGL) and a β-N-acetylglucosaminidase (NAG) from the thermohalophilic bacterium *Rhodothermus marinus* were successfully expressed in *Escherichia coli*. These recombinant enzymes, which belong to the glycoside hydrolase family 3 (GH3), were characterized regarding various parameters. Multiple sequence alignment of these two GH3 enzymes with GH3 members from other organisms allowed prediction of the catalytic nucleophiles, which are positioned at D283 for NAG and D306 for BGL. Identifying catalytic acid/base residues proved to be difficult for BGL, whereas for NAG, D207 together with H209 were predicted. Application of LC-MS provided the molecular mass for BGL and NAG with 87.2 kDa, respectively 104.5 kDa. Both enzymes acted very differently regarding substrate specificity. BGL was classified as a broad-specificity β-glucosidase revealing activity towards many substrates and among the various tested it showed highest activity towards pNPX and xylobiose, followed by pNPG and cellooligosaccharides (cellotriose and -tetraose). NAG on the other hand turned out to be an enzyme with narrow substrate specificity, only exhibiting activity for pNPGlcNAc and chitooligosaccharides (CH_2 to CH_5). Furthermore, kinetic parameters were determined for NAG, where a catalytic efficiency towards pNPGlcNAc of k_{cat}/K_m = 162.30 s^{-1}/mM was obtained. A low K_m (0.07 mM) was measured, leading to conclude a high affinity towards the applied substrate. The primarily assay of optimum temperature measurements allowed to measure an activation energy for NAG of 26.9 kJ/mol, between 50°C to 90°C and revealing a maximum activity at 90°C. After 30 minutes, total inactivation was detected at 90°C and a residual activity of 50% was left after 3 hours of incubation at 80°C, with an apparent half-life of 96 minutes for NAG. Higher residual activity was measured at 70°C compared to 80°C for BGL, with a residual activity of 80% remaining after 180 minutes of incubation and an apparent half-life of 79 minutes (at 70°C) was determined. Another approach than the classical one-parameter-at-a-time investigation was conducted (Design of Experiments), where the effect of temperature and pH was studied in regard of enzyme activity, leading to determine a range of both parameters, in which the enzymes show optimum values. The simultaneous variation of temperature and pH on applied aryl substrates, revealed an optimum enzyme activity at pH 5.4, at all specific investigated temperatures, for both enzymes. Analyzing the effects on xylobiose showed a shift with temperature of BGL´s optimal range. Response Surface Modeling allowed to build a model for NAG, where the consumption of chitobiose was used as a depended variable. NAG exhibited optimal values with increasing pH and decreasing temperature (pH = 6.2 and 65°C-70°C). The evaluated model provided evidence of an influence by both investigated parameters. This kind of study leads to a better understanding of the enzyme´s stability in a more effective way as well as it forms the basis for faster optimum activity determination, which can be easily applied to any enzyme.

Contents

1 Introduction

1.1 Background & Motivation

At a very early stage in history scientists revealed to be very interested in the performance and study of enzymes. Today, enzymes are used in a numerous range of applications, supported by the fact that they have remarkable benefits compared to chemical catalysts [1]. They originate from renewable resources, show ability for biodegradation and provide extreme selectivity in reactant as well as product stereochemistry [1]. Due to the aim to work under environment-friendly conditions and world´s industrialization, this results in the increased interest in enzymes [1].

To illustrate, what an impact enzymes have in the global market, a study for the industrial application of enzymes was carried out by the Business Communications Company Inc [2]. It showed that the global market in 2009 computed an amount of $3.1 billion for industrial enzymes [2]. For 2011 that market was estimated to reach $3.9 billion, growing with a compounded annual growth rate of 9.1% to achieve $6 billion by the year 2016 [2]. The largest sector out of these industrial enzymes, make the food and beverage enzymes, which is followed by the technical enzyme market [2].

The above mentioned results of the global market report reveal a boost in the section of industrial enzyme applications. It should be considered that economically enzymes compete highly with chemical processes, although enzymes show more benefits in regard of environmental point of view compared to chemical treatments [1]. Reasons behind this are the significant lower costs and the traditionally established chemical processes [1]. In industry, enzyme application (e.g. for synthesizing particular chemicals, pharmaceuticals and more) is often hindered through the arising costs during processing, which suffer from low selectivity and formation of undesired byproducts [3]. Furthermore, these obstacles are accompanied by the fact that mostly mesophilic enzymes are applied, which proved to be not highly effective as well as not qualified for the harsh conditions used in industrial processes and this is due to the lack of enzyme stability [3]. Therefore, another option to the mesophilic counterparts provides the research and application of enzymes from extremophilic organisms.

Extremophiles are grouped due to their specific conditions in which they live [4]. A subgroup of the extremophiles, the thermophiles gained great interest, which started already back in the 1960´s, trying to understand their functions [5]. Enzymes from these organisms are extremely thermostable [4], leading to the maintenance of activity at high temperatures [6]. Besides the latter mentioned ability, thermostable enzymes have various additional benefits (e.g. better catalysts, reveal higher reaction rates, higher process yields, reduced

contamination risk and more) [7], which make them interesting for industrial applications. Most of the isolated genes from thermophiles can be successfully cloned, expressed and produced using mesophilic hosts [5]. These beneficial properties will be discussed later on in more detail.

The isolation of thermophilic microorganisms from diverse exotic ecological zones (e.g. hot springs, deep sea vents, submarine hydrothermal areas) of the earth and the access of their enzymes, lead to progresses in the area of enzymatic bioprocess procedures [5], [6], [7]. Finding enzymes with enhanced catalytic properties, the marine environment revealed to have great potential regarding to discover new glycoside hydrolases [8]. These isolated enzymes show a very interesting field for further research, since they can be easily applied in bioprocesses, which is due to their ability to carry out reactions on simple substrates and not needing any external co-factors [5], [8].

Since thermostable glycosidases have several future industrial application perspectives (e.g. during bioethanol production or recycling of particular compounds), detailed investigations in regard of biochemical and structural properties are essential. Each individual application demands for specific requirements from the enzyme to fulfill during the carried out process, such as optimum temperature, pH, stability towards heat, substrate specificity and more. Enzymes originating from extremophilic organisms can be used as models in the following way that the gained information from characterization and the understanding of which factors lead to their stabilization property, can be further applied to engineer more effective enzymes. Furthermore, various methods in use to characterize an enzyme are traditionally established, but that does not mean that they do not provide correct and useful results. Nevertheless, methods exist, which are more efficient than classically applied methods. Such as, when enzyme activity needs to be optimized in regard of specific parameters, Design of Experiments can be used. That allows to change the parameters simultaneously, leading to reduce number of experiments (saving costs) and therefore reveals to be efficient, rather than the often applied procedure of varying one parameter at a time.

The motivation of this master thesis was based on the several advantages enzymes provide from extremophilic organisms, which therefore lead to focus on the detailed characterization of two novel glycoside hydrolase enzymes isolated from a marine thermohalophilic organism *Rhodothermus marinus.* For expression and recombinant protein production, the well explored bacterial organism *Escherichia coli* was used as host. Besides characterizing these glycosidases biochemically and analyzing the primary structure, optimization of enzyme activity regarding temperature and pH was carried out.

1.2 Extremophiles

Extremophiles are a group of organisms living in environments, which are defined as extreme for human life [4]. These extremophiles can be classified respective to their extreme living conditions, which are represented in the following Figure 1.

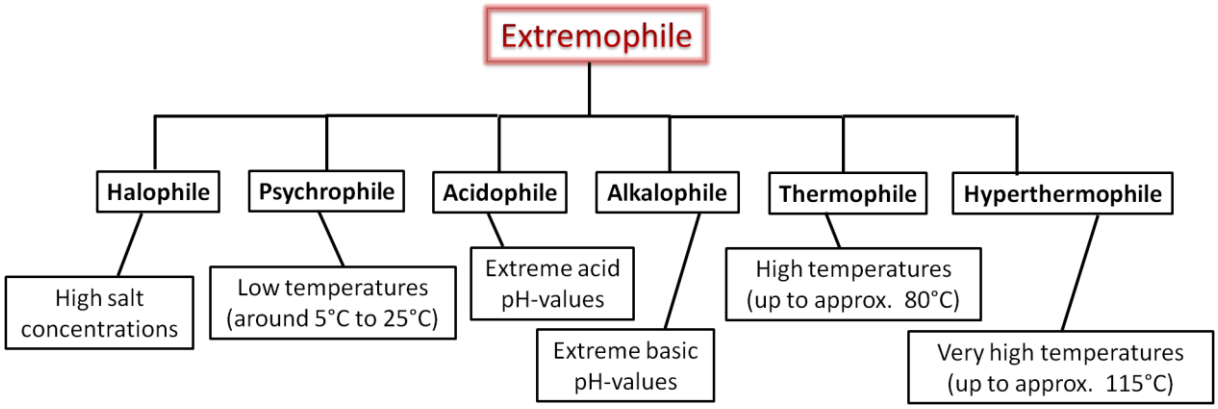

Figure 1: Representing the main subunits of extremophiles. Below each named subunit, the specific living conditions are given, which were obtained from C. Vielle et al. [9] and M. de Champdoré et al. [4].

In addition, some extremophilic organisms exist, which show ability of living in presence of high metal concentrations, the metallophiles or when high levels of radiation exist (radiophiles) [4] .

Extremophiles were discovered as microorganisms producing valuable biocatalysts, which offer to be potential tools for various sustainable developments in industry, due to their stability under extreme conditions [4], [10]. Several enzymes from extremophiles already found industrial applications, in most cases these involve their biomolecules (primarily enzymes and other proteins, lipids and different small molecules) [11]. Most successful and well-known representative application of a extremophilic organism, is the Taq-DNA polymerase (*Thermus aquaticus*) [11]. Other examples of effective appliance of extremozymes are for instance proteases originating from thermophiles used in detergents and brewing [3]. Furthermore, proteases from psychrophiles have found applications during cheese maturation as well as the enzymes from alkalophiles, the cellulases are applied for degradation of polymers in detergents [3]. These are only some examples for industrial applications of enzymes, which were isolated from extremophiles.

1.3 Thermohalophiles

The thermohalophiles describe unique organisms among extremophiles, which grow at high temperatures (up to 80°C) as well as salt concentrations [12]. This kind of property was adapted by these organisms under environmental stresses [12]. The living in high salt habitats is based on the ability of halophiles to maintain constant ion concentration in the cell due to the accumulation of potassium content and removing sodium ions from the cell [4]. Furthermore, higher acidity was detected in halophilic proteins than in mesophiles, when comparing their amino acidic composition [4]. Reasons behind this are the negative charges on the surface that enable a higher surface hydration [4]. Common environments for thermophiles are hot springs or deep sea hydrothermal vents [7]. The enzymes, which live in conditions of high temperature are described as very thermostable as well showing resistance in regard of denaturation and proteolysis [7]. These organisms produce proteins, the so called chaperonis, which are the helping hand after denaturation through refolding and restoring protein functions [7]. Another characteristics of thermophiles to survive at elevated temperatures is the fact that the cell membrane is composed by saturated fatty acids, which support the cell by providing a hydrophobic environment and also retaining the cells rigidity [7]. Due to the properties of thermohalophilic enzymes, these may find application during sugar processing or after chemical treatments in hydrolysis processes [12].

1.4 *Rhodothermus marinus* - A habitant of the marine environment

Over three-fourths of the earth are covered with marine habitats, which compromise a big variety of different microorganisms living under extreme conditions in regard of temperature, pressure, salt concentration, light and nutrient availability [13], [14]. The enzymes from these organisms show valuable biochemical properties, which allow them to adapt among these environmental conditions, so that they can live and thrive [14]. Such enzymes form unique biocatalysts.

Rhodothermus marinus is one of these marine organisms, which lives in the presence of high temperature and high salt concentrations. It is a thermohalophilic bacteria isolated from alkaline submarine hot springs in Iceland [15]. The cells of this marine habitant are rod-shaped, which have a diameter of around 0.5 µm and are 2-2.5 µm long [16]. Furthermore, no spores, flagella as well lipid granules were determined, however the formation of a slime capsule was observed when the bacteria was grown on carbohydrate-rich media [16]. The investigation of this marine bacteria´s colonies revealed to be convex and contain a carotenoid pigment, which leads to their reddish color [15]. Additionally, cells of this strain are oxidase negative, but catalase positive as well as a

performance of nitrate reduction was not followed and the fermentation of the sugars is not carried out anaerobically [16]. *R. marinus* as an aerobe and showing to be moderate halophilic, grows optimally at 65°C and pH 7, with approximately 2% (w/v) of NaCl [15]. A big variety of common sugars are utilized by this bacteria, such as glucose, galactose, lactose and many more [16].

Investigating the phylum of the marine thermohalophilic bacteria exhibits that it belongs to the Bacteroidetes [17]. The groups of this phylum include bacteria, which demonstrate to be very important organic matter degraders under aerobic conditions [18]. In fact, various strains were detected among marine isolates, where mainly cellulose degraders were found [18]. Nolan and colleagues studied the phylogenetic neighborhood of *R. marinus*, which is supposed to be a member of the class Sphingobacteria, with other members of the same class [16]. As the closet relative of *R. marinus* resulted to be the organism called *Salinibacter ruber*, which is an extreme halophile [16]. It is important to keep in mind, that the class of Sphingobacteria is not adequately defined [16].

Besides the various mentioned properties of *R. marinus*, this bacteria is known for its thermostable, polysaccharide hydrolyzing enzymes, which show high potential in industrial applications [17]. Several isolated enzymes were optimally active at temperatures even higher as the optimum temperatures for growth, such as a cellulase (100°C) responsible in beta-glucan hydrolysis or a chitinase (70°C), which functions during chitin hydrolysis [17].

1.5 The host - *Escherichia coli*

The gram-negative bacteria *Escherichia coli* is one of the most widely used mesophilic hosts for recombinant protein production. Its genetics is well understood and it is able to grow fast with high density on inexpensive substrates [19]. Since, thermophiles are often hard to grow, leading to difficulty in protein production, *E. coli* is commonly used as host for the production of thermophilic enzymes [9]. It was shown that most of the proteins of thermophiles are correctly folded and expressed at lower temperatures, they also do not lose their thermostability as well are not hydrolyzed by host proteases [10]. Reasons behind the correct fold of the produced proteins might be the fact that crystal structures of the recombinant proteins are similar with those of the mesophilic homologues [9].

1.6 Thermostable enzymes and their advantages regarding industrial applications

In industry, enzyme costs play a significant factor during enzyme catalyzed processes [20]. Various ways exist to reduce enzyme costs, such as through strain improvement and optimization of the fermentation procedure, which would lead to an increase of enzyme yield [21]. Furthermore, to evolve processes, where the enzyme can be recovered or

re-used [21]. Another option is to use thermostable enzymes, since the parameter temperature is the most likely physical factor, where optimization can be performed [21]. Thermostable enzymes are defined as enzymes which have the ability to be stable as well active at elevated temperature that is even higher than the optimal temperature needed for the growth of the specific microorganism [7].

The research and discovery of novel thermostable enzymes is booming in the last years, due to their beneficial properties and potential for a variety of industrial applications compared to their mesophilic partners. Working at elevated temperatures decreases substrate viscosity leading to higher reaction rates and it allows to obtain higher process yields, which are based on the rising solubility of substrates as well as products [7]. There is also reduced risk of contamination and the higher processing temperatures permit greater tolerance to organic solvents [7]. The property of thermostability offers the possibility for extended storage at room temperature [5], without loss of activity. Further advantages of performing reactions at higher temperatures are that the solubility of raw materials increases and makes them more accessible for enzymatic attack [21]. Carrying out reactions with enzymes at elevated temperatures, the time of conversion as well as enzyme amount can be reduced [22]. Therefore, thermostable enzymes are better catalysts for several industrial applications (e.g. starch industry, cellulase degradation, etc.) [5]. In addition, investigating and understanding the enzyme´s stability as well as activity helps for further developments in protein engineering [7].

1.7 Factors describing and affecting enzyme stability - comparison to mesophilic homologues

The enzyme´s stability can be defined through the fact, that although disruptive forces act on the enzyme, it keeps its active structural conformation [22]. In general, enzymes from thermophiles do not differentiate that much compared to mesophilic enzymes, except their phylogeny difference and the temperature range in which these enzymes show to be stable and highly active [9]. Investigation of the homologous sequences, structures as well as their catalytic mechanisms, resulted in a high similarity between thermophilic and mesophilic enzymes [9].

An assessment for thermostability is the fact that enzymes from thermophiles reveal to be more rigid than their mesophilic counterparts at mesophilic temperatures and this rigidity is suggested to explain often the cause of the inactivation of thermophilic enzymes at low temperature [9]. Although, a thermophilic enzyme is inactive at low temperatures, this property disappears when the enzyme moves towards the temperature, where it shows maximal activity, which is known as the value of temperature optimum. [9] At that specific

value the enzyme performs to be flexible enough and exhibits full activity [9]. It is important to mention that the determination of temperature optimum describes only enzyme´s relative stability and it should be considered that this value varies with the applied substrate concentration, pH as well with time of incubation used for the particular assay [21]. To further describe the thermal stability of an enzyme, the melting temperature as well as the half-life are valuable descriptors.

Yeoman et al. discussed in their recent review the fact that the main effects, which lead to keep the active conformation of the enzyme are hydrophobic effects and the hydrogen bonding. Furthermore, in the native form of the enzyme, the mentioned effects and other forces for stabilization compensate disruptive forces. The responsible factor describing this kind of compensation is known as the free energy change of the protein ΔG. Comparing proteins of thermophiles to their mesophilic counterparts it has been detected that thermostability character is implied by the increase in ΔG, which is probably induced by salt or disulfide bridges, ion pairings as well as through an increase in interactions between aromatic residues, and this in particular in the active site [22].

Analysis of several crystal structures of proteins from extremophilic organisms suggested that ion pairs play an important role during protein stabilization and that these networks of ionic interaction are able to work in a much longer range compared to hydrophobic interactions [3]. Examples exist to prove these assumptions, such as the endocellulase Cel12A from the thermophilic bacterium *R. marinus* revealed an unambiguous increase regarding ion pairings as well as aromatic residues in the active site of the enzyme, relative to its mesophilic homologues [22]. Another example, where the higher amount of ion pairs on the protein surface leads to improve the thermostability, was the comparison of thermostable β-glucosidases to their mesophilic homologues [3].

Further factors contributing to the enzyme´s stability have been discussed in detail by Demirjian and colleagues. They mentioned that under extreme conditions, such as temperature, pH and pressure labile amino acids are subject of covalent modification and that the enzyme stability increases through these labile amino acids in the hydrophobic core. Comparing them to their mesophilic homologues, the appearance of labile amino acids (such as cysteine, asparagine and aspartic acid) is significantly lowered in the more stable proteins from thermophiles. In addition, very often enzymes from extremophiles show to have an unusual character at the solvent-exposed area, which leads to the increase of stability. An example for such a property gives an α-amylase from *A. acidocaldarius*, a halophilic enzyme, which contained a very highly negative surface charge that allowed an increase of solubility, making the enzyme more flexible at high salt concentrations [3].

1.8 Glycoside hydrolases (GH)

Glycoside hydrolases, also referred to as glycosidases or carbohydrases, are a widespread group, which are able to catalyze the hydrolysis of O-, N- and S- glycosidic bonds [23]. Since many decades these enzymes have been the interest of several biochemical investigations and can be found in almost all living organisms [24]. To date, 133 GH families exist in the CaZy (**C**arbohydrate-**A**ctive en**ZY**mes) database [25], [26] which can be classified in the following different ways.

Exo -/endo-acting GH

Depending on the enzymatic attack, either the glycoside hydrolase acts at the end of the applied substrate (exo-acting) or shows to be an endo-acting enzyme, which cleaves in the middle of the substrate chain [27]. It should be considered that the occurrence of an intermediate or the appliance of an unfitted substrate is possible and can make the determination of the enzyme´s substrate cleavage difficult [27].

Sequence-based classification

The heading already describes the way of classification, where glycoside hydrolases are classified based on their amino acid sequence similarity. Each of the GH families compromise proteins, which are related by their sequence, corollary and fold [23]. Therefore, this kind of classification makes it possible to obtain valuable predictions as the catalytic composition, the molecular mechanism and the geometry around the glycosidic bond are conserved for the majority of GH families [23].

Mechanistic classification

Two main mechanisms were proposed (Koshland, 1953) [28], which are carried out by the glycoside hydrolases. Basically, either a change of net retention or an inversion is observed of the anomeric configuration [27]. The understanding of the stereochemistry leads to further follow the acting of the enzyme [27]. More detailed information about the catalytic mechanism of glycosidases is going to be discussed in the next section.

In addition, it is also possible to classify these enzymes through their substrate specificity. However, it should be taken in account that if an enzyme shows substrate specificity towards several substrates, classification results to be difficult and not appropriate enough [27].

Glycoside hydrolases are an enzyme group with a broad range of various application areas. Some of these are mentioned as examples in the following:

- disease research (e.g. Alzheimer) [29]
- development of biofuels (production of butanol or ethanol) [30]
- preparation of synthetically demanding oligosaccharide structures [31]
- and more

1.8.1 Hydrolysis mechanism

As already mentioned glycosidases can be divided into two groups based on the mechanism they follow, either retaining or inverting. The main difference between these two mechanisms is the change (inverting mechanism) or no change (retaining mechanism) of the anomeric center (see Figure 2, red circles).

Vuong and Wilson described very properly the way how glycoside hydrolases act during these mechanisms. In Figure 2a, the inverting glycosidases are shown, where the catalytic acid is responsible for donating a proton to the anomeric carbon, whereas the catalytic base makes it possible to increase the nucleophilicity via removing a proton from the water molecule and therefore enabling its attack on the anomeric center. For glycosidases, which act as retaining glycoside hydrolases (see Figure 2b), two steps during this mechanisms can be followed. The first step allows donation of a proton to the glycosyl oxygen atom and composing an intermediate with the nucleophile (glycosylation step). Whereas in the second step, the deprotonated acid/base residue is in charge to behave as a general base, activating the water molecule that performs the nucleophilic attack on the glycosyl-enzyme intermediate from the previous step, together with two inversion steps resulting in the retention of the anomeric center [29].

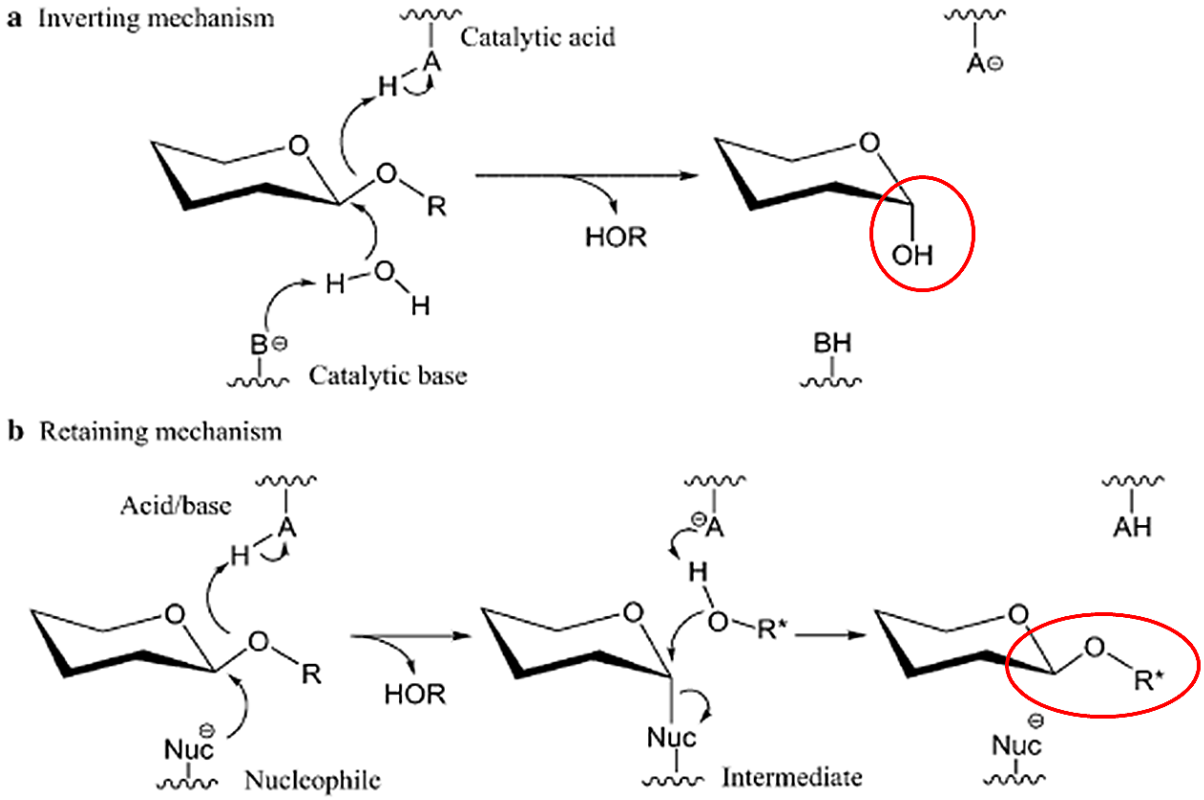

Figure 2: Representing the two mechanisms carried out by glycoside hydrolases (Figure a: inverting, Figure b: retaining). The abbreviations mean the following: catalytic acid residue (AH), catalytic base residue (B^-), nucleophile (Nuc), carbohydrate derivate (R) and finally exogenous nucleophile (HOR*). Red circles indicate the change (a) and no change (b) of the anomeric center. The figure was taken and adapted from T. Vuong et al. and therefore all the listed abbreviations originate from the mentioned source [29].

1.8.2 Glycoside hydrolase family 3 (GH3)

The glycoside hydrolase family 3 is very large, compromising to date over 5000 enzymes in the CaZy database, from which only 223 were characterized [25], [26]. These enzymes occur in bacteria, fungi as well plants and carry out several important roles, such as:

- degradation of cellulosic biomass [23]
- assimilation of glycosides [32]
- pathogen defense [32]
- recycling of bacterial cell wall elements [32]

To date, various activities are known of this family, at which some are α-L-arabinofuranosidase, β-D-glucosidase, β-D-N-acetyl-glucosaminidase as well as β-D-xylosidase [33]. The glycoside hydrolases of family 3 cleave their substrates from the exo-site [23] and are classified to follow the retaining mechanism [23], [34]. Furthermore, GH3 family enzymes do not only cleave simple sugars, but it was also detected that they remove glycoside units from heterogeneous molecules (e.g. xyloglucan or

arylglucosides) [32]. In addition, these glycosidases often prove to exhibit dual or broad substrate specificity [32]. Such as the β-glucosidases from thermophilic organisms *C. thermocellum* or *Thermotoga sp.* revealed broad substrate specificity [35], [36]. Bifunctional glycosidases from GH3 family have also been characterized for example originating from *A. tumefaciens* [37].

1.8.2.1 β-glucosidase

β-glucosidases (EC 3.2.1.21) are enzymes, which can be found in all kingdoms of live, ranging from bacteria to highly evolved mammals [38]. These enzymes are responsible for hydrolyzing glycosidic bonds at the non-reducing end and thereby releasing glycoside residues from glycosides and oligosaccharides [39]. To date, β-glucosidases are grouped into six different glycoside hydrolase families (GH1, GH3, GH5, GH9, GH30 and GH116) [25], [26].

Functions

The β-glucosidase enzyme gained much interest due to its involvement in various biological functions [40]. A very fundamental role is carried out by this enzyme in regard of cellulose and other carbohydrates degradation, which is important for nutrient up-take [41]. Furthermore, it functions during biogenesis of several functional molecules from glycoside precursors, for example flavonoids [42]. β-glucosidases are also known for chemical defense against pathogen attack [40].

Substrate specificity

β-glucosidases can be divided into three groups based on their substrate specificity. One group belongs to aryl-β-glucosidases, which show high affinity towards aryl-β-glucosides [43], [44]. Another one, the cellobiases, responsible for hydrolyzing only oligosaccharides and the third type are the broad-specificity β-glucosidases revealing activity towards many substrate types [43], [44]. As already mentioned above, it is not uncommon that enzymes from the GH3 family are multifunctional, in the sense of exhibiting several activities [32]. In the case of β-glucosidase, mostly the combination of β-glucosidase and β-xylosidase [45], [46], but also β-glucosidase/β-xylosidase/α-arabinosidase activity [47], was detected.

The role of β-glucosidase during lignocellulosic biomass degradation

The production of ethanol from renewable biomaterial gained much interest in the last years due to the shortage of fossil fuels and the concern about the environment (emission of greenhouse gasses and air pollution) [48]. Therefore, the potential research of converting renewable biomass into chemicals and biofuels leads to produce environmental-friendly products. Lignocellulosic biomass compromises the carbohydrate polymers cellulose and

hemicellulose and is the most abundant renewable polymer in nature [49]. These two carbohydrate polymers are the main building blocks of plant cell walls [20]. Cellulose is known to be an unbranched polymer, which is composed of β-(1,4)-linked D-glucose residues [48]. Whereas hemicellulose is a more complex and highly branched polymer that is mainly built of xylan and other different polysaccharides (glucans, xyloglucans, callose, mannans and glucomannans) [22].

For the degradation of cellulose three enzymes work together. Endoglucanases and cellobiohydrolases are responsible in the first step of degrading native cellulose and producing cellodextrins [42]. These products are then hydrolyzed by β-glucosidase, which makes it possible to not only terminate the last step of cellulolysis of lignocellulosics obtaining glucose, but also helps in reducing the product inhibition (cellobiose) on the first two mentioned enzymes employed for cellulose breakdown [42]. In the case of hemicellulose, degradation takes place by applying xylanases and β-xylosidases, at which the latter mentioned enzyme is in charge of catalyzing the hydrolysis of xylobiose and xylooligosaccharides to xylose [49].

In regard of bioethanol production, one of the key bottleneck is still the step of cellulosic biomass degradation to fermentable sugars (e.g. glucose, xylose) [49]. Although, different technologies have been developed for the biomass treatment and also several enzymes have been found and are commercially applied for the degradation process, limitations exist [49]. These are mainly applied to the used enzymes, which make the difference regarding the expenses incurred [49]. This is based on being subject to restrictions such as their stability, showing sensitivity towards some byproducts or they do not hydrolyze efficiently [49]. One of the major problems often is that applied enzymes are not thermostable enough during the saccharification step, where steam is applied to make the raw material more easily accessible for the enzymatic hydrolysis [44]. Hence, researchers are in search of novel β-glucosidases, which show ability of thermostability, are highly active or reveal to be multifunctional [49], because these properties would lead to decrease in enzyme costs.

Application

Besides the appliance of β-glucosidase for the degradation of cellulosic material, it has also been used for several other potential applications. Diu et al. mentioned their application for the release of flavor compounds in juices and wines as well the release of phenolic compounds, which reveal antioxidant activity from fruit and vegetables [44]. Furthermore, it is possible to synthesize oligosaccharides and glycosides of interest with the help of

β-glucosidases, since these are able to carry out reverse hydrolysis and transglycosylation processes [39].

1.8.2.2 β-N-acetylglucosaminidase

The enzyme β-N-acetylglucosaminidase is also well-known under the name β-N-acetyl-D-hexosaminidase. With the EC-number 3.2.1.52, β-N-acetylglucosaminidase is until now not only a member of the glycoside hydrolase family 3, but it is as well grouped to the families GH18, GH20, GH84 and GH116 [25], [26]. β-N-acetylglucosaminidases are a widespread group of enzymes, occurring in different organisms, which range from bacteria to humans [50]. Their role is to catalyze the removal of N-acetyl-D-glucosamine (GlcNAc) or N-acetyl-D-galactosamine (GalNAc) at the non-reducing end from various glycoconjugates and saccharides [50], [51]. In general, β-N-acetylglucosaminidases are relatively specific regarding their substrate specificity and investigations showed that β-N-acetylglucosaminidases in glycoside hydrolase family 3 are more specific for the gluco-configuration rather than galactosyl substrates [52].

β-N-acetylglucosaminidase was found to have various different functions. Such as their ability to work during defense-related processes or in plants metabolism, storing glycoproteins [53]. Furthermore, it is responsible in processes regarding bacterial growth, insect and fungal pathogenesis and in lysosomal degradation of glycolipids as well glycoproteins, where a damage of it, leads to the Tay-Sachs and Sandhoff disease in humans [31].

Besides the mentioned functions, in which β-N-acetylglucosaminidase is involved in, a key role during chitin degradation is attributed to this enzyme, where it performs an essential vegetative function [31]. Chitin is an insoluble homopolymer composed of β-(1,4)-linked N-acetylglucosamine residues and after cellulose the second most abundant natural biopolymer, which particularly occurs in the marine environment [10]. The polymer chitin can be found in the cell wall of fungi, algae and yeast but also in the exoskeleton of crustaceans.

In Figure 3 the beginning of the chitin catabolic pathway is given, which represents the biomass degradation carried out by the combined action of chitinase (EC 3.2.1.14) and β-N-acetylglucosaminidase. The given pathway was studied in the family Vibrionaceae and since the two mentioned enzymes are detected in many other organisms, the here illustrated stage of the pathway is potentially existing in a wide taxonomic range [54]. The subsequent part of the pathway is further given in the chitin derivates degradation pathway, which will not be discussed here, because this would go into too much detail in this context.

The breakdown of the chitin polymer is initiated by endo-acting chitinases, producing a smaller subunit, a chitodextrin, which contains more than two GlcNAc monomers. This subunit is further degraded into smaller soluble di- and oligosaccharides (N, N´-diacetylchitobiose, chitotriose), which are then finally hydrolyzed to GlcNAc, by the action of β-N-acetylglucosaminidase. The degradation products, in the most cases the GlcNAc monomers are used as carbon and nitrogen source by the cells [55].

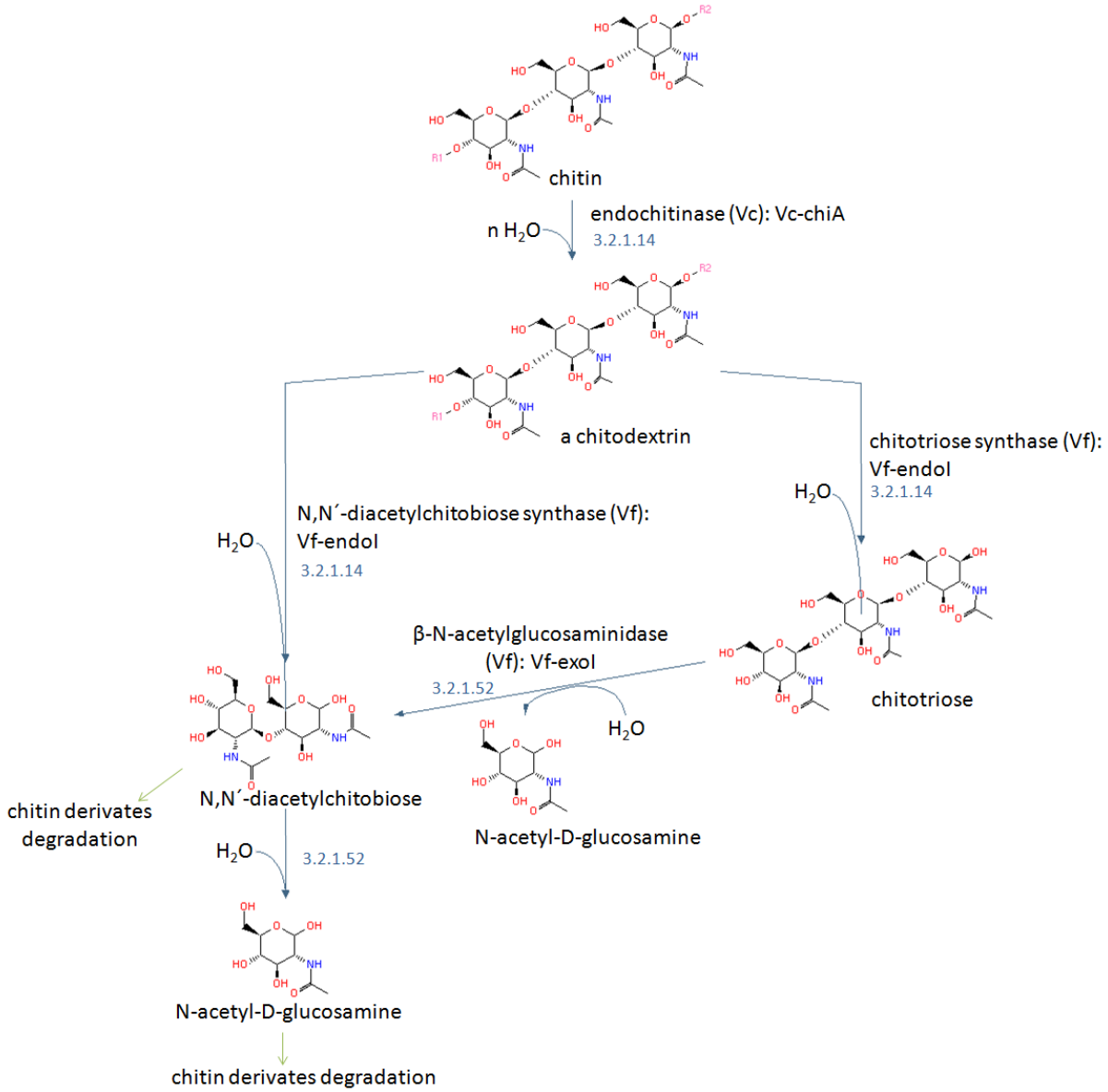

Figure 3: Chitin degradation pathway obtained and adapted from MetaCyc [54]. Chitinases (EC 3.2.1.14) and β-N-acetylglucosaminidase (EC 3.2.1.52) are working together in the degradation of chitin and its subunits, to produce N-acetyl-D-glucosamine. The represented enzyme activities in this pathway have been proven experimentally and besides the enzyme name, the gene name is also given. Although many taxa are known to possess this pathway, the two shown here are *Vibrio cholerae* (Vc) and *Vibrio furnissii* (Vf).

Each year a great steady-state amount of chitin accumulates in the marine environment [56]. To maintain the ecosystem of this environment, it is of great importance to recycle this abundant biomass [55]. Therefore, the appliance of enzymes from thermophiles would lead to more effective degradation due to the advantages thermostable enzymes have compared to their mesophilic partners. To date, numerous β-N-acetylglucosaminidases have been isolated and characterized, but very few of these have thermostable character (e.g. from thermophilic bacteria *Streptomyces thermoviolaceus* [57] or *Bacillus stearothermophilus* [58]). Not only during the process of chitin degradation thermostable β-N-acetylglucosaminidase has potential and is interesting for further research, there are also many other areas for its useful application. Such as for the preparation of synthetically demanding oligosaccharide structures [31] or of pharmaceutically relevant chitooligosaccharides and N-acetyl-D-glucosamine [59]. Other application areas do also include the use of β-N-acetylglucosaminidase as biocontrol agent of phytopathogenic fungi or as a working assistance for investigation of complex glycosylation chains on glycolipids and -proteins [31].

1.9 Bioinformatics analysis - structural information about GH3 family enzymes

To date, only few enzymes of the GH3 family are understood based on structural knowledge, although a high number of sequences are known. The very first enzyme of the GH3 family, which was characterized the best, both biochemically as well as structurally, is the β-D-glucan exohydrolase (HvExoI) from barley (*Hordeum vulgare*) [60]. Various further structurally characterized enzymes from the GH3 family used HvExoI as a model to understand them regarding domain architecture, catalytic mechanism and more. The GH3 enzymes, which have been investigated based on structure revealed to be composed of two domains, although some exist to exhibit a three domain architecture. One of the first three domain glycosidase of family 3 was the β-glucosidase originating from the bacteria *Thermotoga neapolitana* (TnBgl3B) [61], which was then followed by the structural analysis of other three domain glycosidase hydrolase family 3 enzymes, such as the exo-1,3⁄1,4-β-glucanase from *Pseudoalteromonas sp.* BB1 (ExoP) [62] or another β-glucosidase from *Aspergillus aculeatus* (AaBGL1) [63]. Furthermore, also some β-N-acetylglucosaminidases have been structurally investigated, such as from *Vibrio cholerae* [64] or *Bacillus subtilis* [65], both of them are composed of two domains. Until now no three domain β-N-acetylglucosaminidases have been examined.

GH3 family enzymes reveal to be similar regarding their composition of the first two domains, which is not uncommon. Since, the above mentioned enzymes from GH3 family exhibit a TIM barrel domain, which represents the first domain (N-terminal domain) and a

α/β-sandwich domain as the second domain (C-terminal domain). For detailed understanding and further discussion of glycoside hydrolase family 3 structure analysis, the structures of β-glucosidase of *T. neapolitana* Figure 4 [61] and the β-N-acetylglucosaminidase of *B. subtilis* Figure 5 [65] are shown as ribbon models compared to HvExoI from barely [60], which are well investigated representatives of this subgroup of glycosidases. The domains of TnBgl3B are connected by linkers (Figure 4A) and the two domains of HvExoI superimpose very well with the first two domains of TnBgl3B [61]. Furthermore, investigations showed that a great variation exists between TnBgl3B and HvExoI in these domains regarding the loop structure, which is believed to may result in influencing the substrate specificity of these two GH3 family enzymes [61]. The active site of the GH3 enzyme from barley is found in a pocket, located at the interface of the first two domains and Pozzo et al. reported that the location of the mentioned active site of HvExoI is identical with TnBgl3B [60], [61]. Since, HvExoI revealed broad substrate specificity, it has been discussed that this can probably be traced back to the geometry of the active site, which shows to be shaped as a shallow coin slot, leading to the binding of a wide range of substrates, where the overall shape of the substrate is not considered as a determinant for the formation of the enzyme-substrate complex [66].

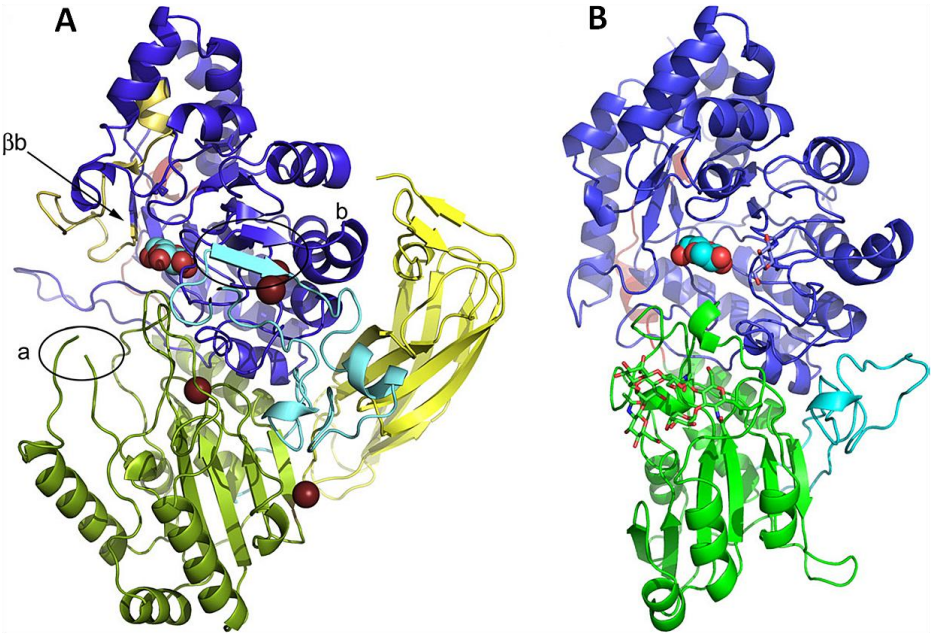

Figure 4: Overall structure of the three domain TnBgl3B from *T. neapolitana* (A) compared to the two domain HvExoI from barely (B). (A) The domains are colored differently, at which domain 1 is presented as blue colored strands, domain 2 in green and domain 3 in yellow. The red colors represent the linker between domain 1 - domain 2, whereas the linker of domain 2 - domain 3 are given in cyan. Box "a" shows the extended loop region in domain 2, which extends to the enzyme´s active site. Whereas Box "b" labels the short anchoring β-interaction between domain 2 and 3 (residues 568-571) as well domain 1 (residues 178-180). (B) The overall structure of HvExoI is colored the same as in figure 4A and the binding of the glucose moiety is illustrated as spheres. These structures were taken and adapted from Pozzo et al. [61] to illustrate the GH3 enzyme´s three dimensional architecture.

Among glycoside hydrolases from family 3 the catalytic nucleophile is highly conserved, always being present in the first domain [61], [63], [66]. Whereas the catalytic acid/base residue reveals not high conservation and this is particularly true for distantly related GH3 family members [66]. Often glutamic acid serves as the acid/base residue, which is true for several β-glucosidases [60], [61], [63], but was never identified for the β-N-acetylglucosaminidase of *B. subtilis* [65].

Litzinger and colleagues proved structurally and with kinetic data the catalytic acid/base residues of the two-domain β-N-acetylglucosaminidase from *B. subtilis*. Residue H234 of the Asp-His dyad acts as the acid/base catalyst, which is supported by Asp232 (Figure 5D, orange sticks). It is possible to characterize β-N-acetylglucosaminidases of the family 3 through the highly conserved sequence pattern KH(F/I)PG(H/L)GX(4)D(S/T)H that is located at the N-terminal domain. Furthermore, the structural investigation showed that the determined catalytic acid/base residue (glutamic acid) of HvExoI located on a short helix on the C-terminal, which is closely associated with the active site area of the N-terminal domain (Figure 5A,B), superimposes with H234 of the dyad of *B. subtilis* β-N-acetylglucosaminidase. Some β-N-acetylglucosaminidases of family 3 do not have a C-terminal domain at all, which leads to the need of other residues for the catalytic acid/base mechanism. Moreover, it has been hypothesized that the Asp-His dyad in β-N-acetylglucosaminidases may orient in a proper way during substrate binding, which leads to high substrate specificity. It is also believed that this general acid/base catalyst may not be necessary, when hydrolysis of natural substrates takes place. In addition, the results that a histidine residue takes part in an Asp-His dyad is found to be unique among glycosidases. Since, such an involvement often occurs among enzymes that cleave phosphodiester bonds, for example ribonucelases [65].

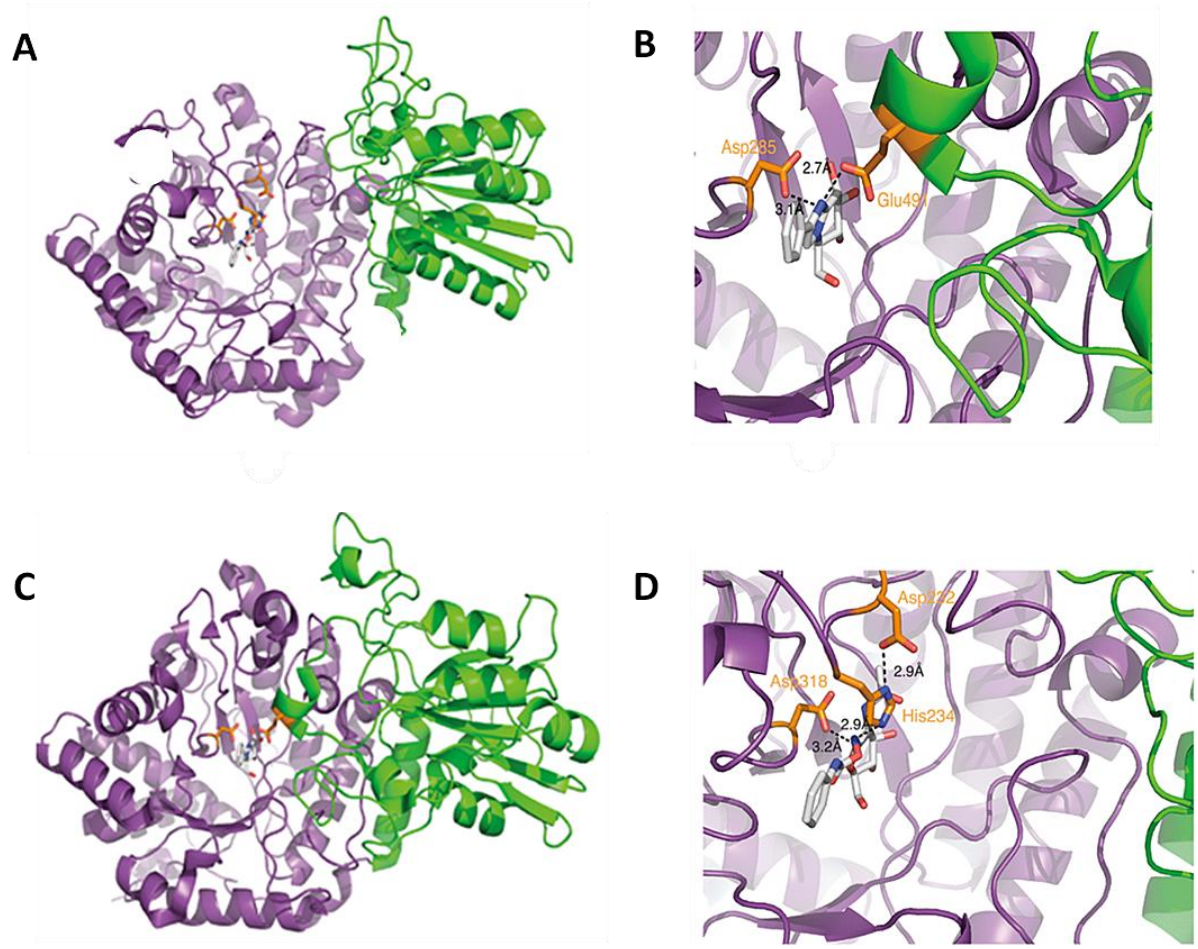

Figure 5: Representing ribbon models of β-D-glucan exohydrolase (HvExoI) from barley (*Hordeum vulgare*) (A, B) and β-N-acetylglucosaminidases from *Bacillus subtilis* (B, C). The gray sticks illustrate the bound inhibitor in the active site, whereas the orange colored sticks represent the catalytic nucleophile as well as acid/base residues. The N-terminal and C-terminal domain are colored in magenta and green, respectively. (A) The top of the TIM-barrel domain is shown, where the inhibitor glucophenylimidazole is bound in the active site that is carrying the catalytic nucleophile (orange sticks on the left site, A285). C-terminal domain is associated closely with the TIM-barrel domain also participating in the active site. (B) Shows the ribbon model of a short helix from the C-terminal domain, which gets closer to the bound inhibitor and is carrying the catalytic acid/base residue (orange, right side, G491). (C) The same as in (A), the only difference is that the N-terminal domain is bound to O-(2-acetamido-2-deoxy-D-glucopyranosylidene) amino N-phenylcarbamate (PUGNAc) in the active site, carrying the catalytic nucleophile (orange, left side, A318). Compared to HvExoI, the C-terminal domain located in a greater distance from the C-terminal domain and is not contributing to the active site as it was the case in HvExoI. (D) The closer view represents the Asp-His dyad, which work as the catalytic acid/base residue (Asp232 and His234). The catalytic nucleophile as well as the acid/base residue H234 are H-bonding the PUGNAc inhibitor. The ribbon models, showing the overall structure as well as details about residues acting during catalytic mechanism, were taken and adapted from Litzinger et al. [65].

As already mentioned above, besides the common two-domain architecture of GH3 enzymes, some were detected to reveal also a third domain. Pozzo and colleagues showed in the results of crystallization experiments of TnBgl3B, the appearance of a third domain, known as a Fn(III)-like domain and is present in several extracellular glycoside hydrolases originating from bacteria [61]. Little is known about this domain, although some suggestions

have been made. Such as acting during substrate binding or working as spacers between the other domains [61], [67]. However, clear evidence for its function has rarely been proven.

1.10 Design of Experiments

Whenever an enzyme is characterized regarding optimal conditions in order to obtain high activity, which is influenced by different factors (e.g. pH, temperature), this is done in most cases by varying one factor at a time. Such a procedure reveals to be inefficient, not necessarily providing optimal conditions. In fact this is true, if the investigated factors exhibit to interact with each other. This leads to the wrong interpretation, believing that the optimum has been reached due to the fact that the variation of one factor at a time does not cause any additional improvement in the results [68]. In Figure 6A, an example for the described procedure is given, which does not allow the determination of the real optimum [68]. The solution for this purpose gives the simultaneous variation of all relevant factors, which is an approach known as **D**esign **o**f **E**xperiments (DoE). Figure 6B, provides such an example, where two factors can be investigated at the same time by applying a small experimental design and the distribution of the experiments as a rectangular allows to identify a direction, which will lead to better results [68].

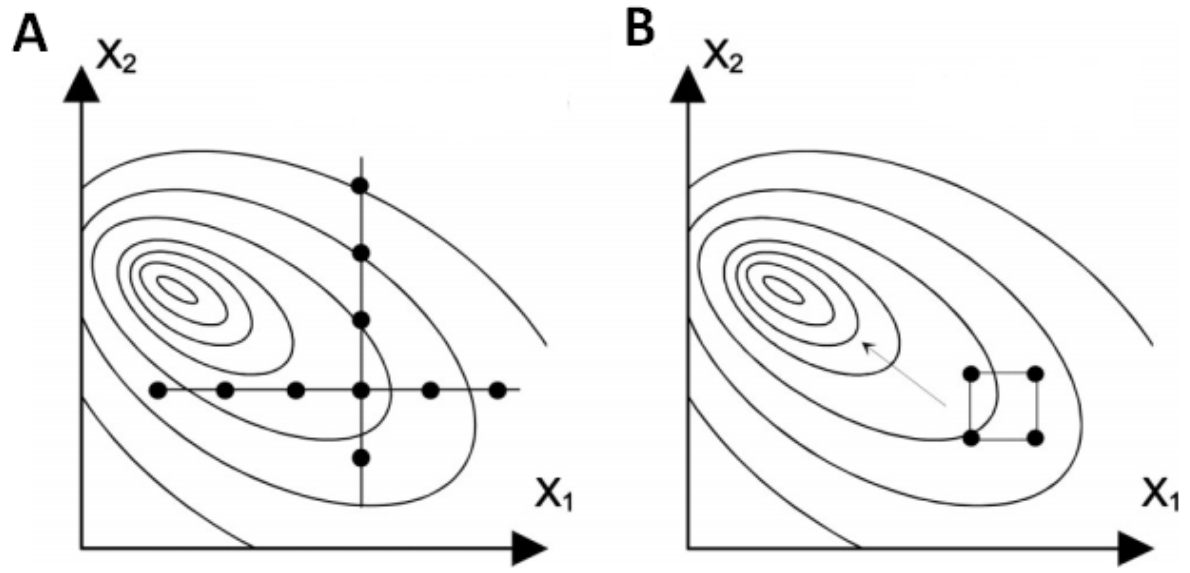

Figure 6: Illustration of the difference between varying one factor at a time (A) and simultaneous variation (B) during optimization study (X_1, X_2 are factors). In figure A, the factor X_1 is kept constant and X_2 is varied. After determining the optimum value, in the next step X_1 is varied, which leads to a quasi-optimum due to the variation of one factor at a time. Figure B, shows the process of simultaneous variation of both factors that allows determining the true optimum. The presented figures were taken and adapted from Mandenius et al. [69].

DoE provides a versatile tool for statistical design and evaluation of experimental data in a predefined experimental range. This kind of statistical experimental planning allows to investigate as well correlate defined input factors of a process and the obtained experimental outcomes, known as responses (Figure 7) [69]. DoE can be used for various problems, which are referred to as experimental objectives, mainly screening, optimization and robustness testing [68]. Besides the advantage of DoE to extract maximal amount of information through carrying out minimal amount of experiments, it also prevents from experimental biases as well as allows to reveal interactions between factors and how they influence the responses [69].

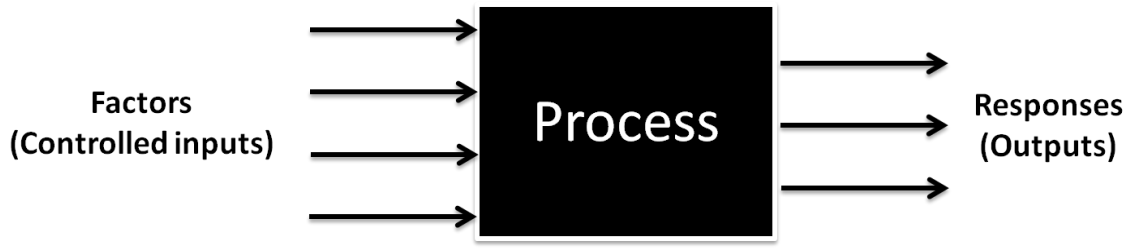

Figure 7: Schematically showing the black box process model characteristics of DoE. The applied factors and obtained responses of the specific process are mathematically related to each other. Adapted and modified from Mandenius et al. [69].

The application of DoE is accompanied by carrying out several steps. Initially the experimental objective has to be selected and the factors defined as well as responses specified. A model describes the mathematical correlation of the input factors and the output responses. Three polynomial models exist, which are frequently used in DoE. These are the linear, interaction and quadratic regression models [68]. The selection of the model is already carried out, when deciding for an experimental objective. Experimental runs can be positioned in the experimental design region, which is the region to be investigated. In this case also the choice of experimental design is based on the requirements of the specific experimental objective. Screening is an efficient approach to identify factors influencing the responses the most, in only few experiments, but also to determine ranges for the factors to be tested and investigated [68], [69]. This kind of experimental objective is often used as an initial step, before continuing to apply optimization or robustness testing.

In the optimization task, designs of the composite family (**C**entral **C**omposite **F**ace-centered design, CCF or **C**entral **C**omposite **C**ircumscribed design, CCC) are applied and based on quadratic polynomial models [68]. As an example, a CCF design with two factors is shown in Figure 8, where in the middle three center points are positioned for investigation of

experimental error and to avoid systematic error the experiments should be carried out in a randomized order [68]. Besides the aim to determine the optimum, the importance of the specific factors is also analyzed by answering the question, if there exists a positive or negative relation between the input factor of the process and the outcome response [68]. To represent the results of the fitted model obtained from applying optimization as the experimental objective, the best way is to use response surface plots, which is also well established as the approach known as **R**esponse **S**urface **M**odeling (RSM) [68].

The third experimental objective is nothing else than to test the robustness of a product or method, before it is going to be released [68]. For more details about the latter mentioned objective, the interested reader is referred to the following literature [68].

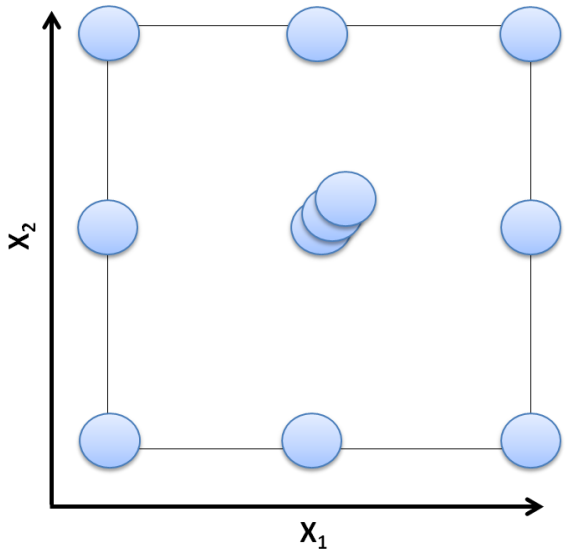

Figure 8: Illustration of a CCF design, which shows the to investigate experimental space in regard of two factors X_1, X_2. Each of the represented circle stands for an individual experimental run at specific conditions. In the middle of the design, the triplicate center points are positioned, which are carried out to estimate experimental error.

Multiple **L**inear **R**egression (MLR) is used to fit the obtained model data. Two important parameters exist to describe the model validity, which are the goodness of fit value R^2 and the goodness of prediction value Q^2 [68]. R^2 is known as the fraction of the variation of the responses, which can be explained by the model and it spans a value between 0 to 1 (from no variance explained by the model to variance explained by the model) [68], [69]. As the description of R^2 already states, it is responsible for revealing how well the experimental runs are reproduced in the mathematical model [69]. Whereas Q^2, which also spans values between 0 to 1, describes the fraction of the variation of the responses that is predicted by the model [68], [69]. According to Mandenius and colleagues, good models are described

through values which are above 0.75 for R^2 as well 0.60 for Q^2 and obtained values below 0.25 are considered to be unreliable [69].

1.11 Goals

The thesis deals with the investigation of two novel glycoside hydrolase family 3 enzymes from the marine bacteria *Rhodothermus marinus*. Production and purification of these enzymes were the first objectives. In order to gain insight of the enzymes unknown properties, the main focus was set to characterize them and therefore being able to gain knowledge for possible future applications. Another aim was to detect optimal working conditions for these analyzed enzymes, through optimizing enzyme activity by varying pH and temperature simultaneously and to understand their influence.

1.12 Roadmap and structure of thesis

Two different GH3 family enzymes were investigated in this thesis, which are divided into two parts. The results of bioinformatics analysis and subcloning experiment are given as a separate section, before continuing with the individual parts of the analyzed enzymes. Part I, describes the results of the first characterized enzyme β-glucosidase (BGL), which is then followed by discussing these results and giving a short conclusion. The same is valid for the second enzyme of interest, the β-N-acetylglucosaminidase (NAG), which can be found in Part II.

Next, in the overall discussion, all the obtained results of both enzymes are compared. Finally, the achievements are summarized in the final overall conclusion as well recommendations for further improvements of characterization methods are given, in the following outlook section.

2 Materials and Methods

2.1 Composition of buffers and solutions

2.1.1 Protein production and purification

Lysogeny Broth (LB) media

Bacto-tryptone	10 g
Bacto-yeast extract	5 g
NaCl	10 g
dH_2O	to 1 l

Binding buffer

Imidazole	10 - 20 mM
NaCl	750 mM
TRIS	20 mM
pH adjusted with 4 M HCl to 7.5	
dH_2O	to 1 l

Elution buffer

Imidazole	250 - 300 mM
NaCl	750 mM
TRIS	20 mM
pH adjusted with 4 M HCl to 7.5	
dH_2O	to 1 l

2.1.2 SDS-PAGE

5x Electrode buffer

TRIS	15 g
Glycine	72 g
SDS	5 g
pH adjusted with 4 M HCl to 8.3	
dH_2O	to 1 l

Sample buffer

0.5 M TRIS (pH 6.8)	3.3 ml
10% w/v SDS	8 ml
Glycerol	4 ml
0.2% w/v Bromphenol blue	1 ml
DTT	2.62 g
Milli-Q-H_2O	to 20 ml

L1 stock solution

Acrylamide	29 parts
BIS	1 part

L2 stock solution

0.75 M TRIS	22.7 g
0.2% w/v SDS	0.5 g
pH adjusted with 4 M HCl to 8.8	
dH_2O	to 250 ml

L3 stock solution

0.25 M TRIS	30 g
0.2% w/v SDS	0.2 g
pH adjusted with 4 M HCl to 8.8	
dH_2O	to 100 ml

Staining solution

0.2% Coomassie	1 g
40% MeOH	400 ml
10% HAc	100 ml
Milli-Q-H_2O	to 1 l

Destaining solution

40% MeOH	400 ml
10% HAc	100 ml
Milli-Q-H_2O	to 1 l

Running gel (10%) - 4 gels

L1	7.5 ml
L2	15 ml
dH_2O	6.5 ml
APS[a]	400 µl
TEMED	15 µl

[a] Concentration of 10% (in dH_2O)

Stacking gel (3%) - 4 gels

L1	0.8 ml
L3	8.0 ml
dH_2O	1.2 ml
APS[a]	200 µl
TEMED	15 µl

[a] Concentration of 10% (in dH_2O)

2.1.3 Affinity gel electrophoresis

6.5% native gel - (2 gels)

Water	4.5 ml
Substrate[a]	5 ml
Buffer (5x)	3 ml
Acrylamide	2.5 ml
Riboflavin[b]	10 µl
TEMED	10 µl
APS[c]	75 µl

[a] For control gels, replace the substrate with water
[b] 10 mg/ml solution in dH_2O
[c] 100 mg/ml solution in dH_2O

5x Running buffer (pH 8.7)

TRIS base	30.29 g/l
Boric acid	7.73 g/l

5x Loading buffer (8.7)

TRIS	0.66 g
Boric acid	0.15 g
Bromphenol blue	20 mg
87% Glycerol	10 ml
dH_2O	10 ml

2.1.4 Biological buffers

100 mM citrate phosphate buffer (pH 2.6 to 7.6) prepared according [70].

100 mM potassium phosphate buffer (pH 5.8 to 8) as given in [71].

200 mM glycine-NaOH buffer (pH 8.6 to 10.6) preparation followed as in [72].

Table 1: Chemicals

Chemical	Company
1,3-β-D-Laminaribiose	Megazyme
1,3-β-D-Laminaritriose	Megazyme
1,4-β-D-Cellobiose	Megazyme
1,4-β-D-Cellotetraose	Megazyme
1,4-β-D-Cellotriose	Megazyme
1,4-β-D-Xylobiose	Megazyme
2-Amino-2-hydroxymethyl-propane-1,3-diol (TRIS)	Saveen Werner AB
4-Nitrophenyl N-acetyl-β-D-glucosaminide	Sigma-Aldrich
4-Nitrophenyl α-L-rhamnopyranoside	Sigma-Aldrich
4-Nitrophenyl β-D-cellobioside	Sigma-Aldrich
4-Nitrophenyl β-D-glucopyranoside	Sigma-Aldrich
4-Nitrophenyl β-D-xylopyranoside	Sigma-Aldrich
Acetic acid	Merck
Agarose standard	Saveen Werner AB
Ampicillin sodium	Duchefa Biochemie
Bacto tryptone	Duchefa Biochemie
Bacto yeast extract	Duchefa Biochemie
Barely β-glucan	Sigma-Aldrich
Bicinchoninic acid (BCA) solution	Sigma-Aldrich

Table 2: continued

Chemical	Company
Citric acid	Merck
Copper (II) sulfate solution (4% w/v)	Sigma-Aldrich
Di-N-Acetyl-Chitobiose	Megazyme
Dithiothereitol (DTT)	Saveen Werner AB
Ethanol	VWR
Ethylenediaminetetraacetic acid (EDTA)	Sigma-Aldrich
Glycine	Merck
Hydrogen chloride	Merck
Hydroxyethyl Cellulose	Sigma-Aldrich
Imidazole	Scharlau
Kanamycin sulfate	Sigma-Aldrich
Lichenan	Megazyme
L-Rhamnose monohydrate	Sigma-Aldrich
Methanol	VWR
N,N,N',N'-tetramethylethylenediamine (TEMED)	Merck
N-Acetyl-D-glucosamine	Sigma
Oatspelt xylan	Sigma-Aldrich
penta-N-Acetyl-Chitopentaose	Megazyme
Potassium dihydrogen phosphate	Merck
Potassium hydrogen phosphate	Merck

Table 3: continued

Chemical	Company
Sodium hydroxide	Merck
Sodium phosphate dibasic dihydrate	Applichem
Tetra-N-Acetyl-Chitotetraose	Megazyme
Tri-N-Acetyl-Chitotriose	Megazyme
Xyloglucan	Megazyme

2.2 Genes and strain

The genes Rmar_0925 (plasmid pSO262a) and Rmar_1080 (plasmid pSO257a) which have been used, were cloned without signal peptide sequences and ligated into the vector pJOE3075. All plasmids contain an ampicillin resistance as a selection marker as well as His_6-tag at the C-terminus. Expression for the production of recombinant proteins took place in *Escherichia coli* (strain: C43-DE3), where rhamnose (T7Rha promoter) was used as inducing agent.

2.3 Subcloning experiment

Rmar_2069 (plasmid: pSO265a) was another gene of interest not containing His_6-tag, which is needed for further protein purification. Therefore a subcloning procedure for His-tag insertion had to be carried out. Firstly digestion of pET21b, pET28b as well as of the vector of interest pJOE3075 was performed. For this purpose the vectors were digested each with 2 µl NdeI (Thermo Scientific, Fermentas, USA) and HindIII (Thermo Scientific, Fermentas, USA) at 37°C (Termaks, Norway) for 3.5 hours. The successfulness of the restriction digest was checked up on 1% agarose gel and a 1kb ladder (Generuler 1kb DNA ladder, Thermo Scientific, Fermentas, USA) was used. The electrophoresis was carried out in three steps. Firstly the gel was run at 100 V for 10 minutes, followed by 20 minutes at 120 V and finally for 25 minutes at 130 V.

The technique of subcloning was not continued, because of problems with the vector pJOE3075. For more detailed information refer to results (section 3.1).

2.4 Sequence analysis

Primary structure was investigated by BLASTP and Pfam database [73]. Domain architecture was analysed by using Interpro (EMBL-EBI) [74] as well NCBI-conserved domain platform [75], [76], [77]. Selection of various protein sequences for **M**ultiple **S**equence **A**lignment (MSA) were based on protein existance, showing evidence at protein level as well on known crystal structures, which were retrieved from UniProt [78]. The alignment was generated through applying ClustalW2 [79], [80].

2.5 Expression of recombinant protein

Pre-culture was prepared, where 50-100 µl cells from glycerol stock were added to LB-media containing 100 mg/ml ampicillin as selection marker. Parameters of incubation were 37°C and 180 rpm (IKA KS400 icontrol, Germany).

For the production of recombinant protein, a shake flask cultivation of 500 ml LB-media (100 mg/ml ampicillin concentration) was inoculated with 5 ml of pre-culture (OD_{620} > 2.5). After reaching an OD_{620} value greater than 0.5, expression of the protein was induced with 10% rhamnose (0.1 % final concentration). For further expression level check, samples of 1 to 1.5 ml were taken for at least four hours after induction. These collected samples were centrifuged for 2 minutes at 13 000 rpm (VWR Galaxy 14D, USA) and supernatant was discarded (only saved of the fourth hour sample). The obtained cell pellets as well as one sample with supernatant were saved at -20°C. The incubation temperature of the culture was reduced to 35°C for a continued overnight growth. Total process time of the batch cultivation was approximately 24 hours.

Cells were harvested through centrifugation at 4°C and 10 000 rpm (RC 5C- Sorvall Instruments, DuPont, USA) and kept at -20°C for further purification procedure.

2.6 Cell rupture via ultrasonication

Firstly the harvested cells were resuspended in binding buffer. In the second step, the cell suspension was sonicated for 2 minutes on ice (UP400S, Dr. Hielscher GmbH, Germany), with an amplitude of 60% and a 0.5 cycle. After each ultrasonication step the cells were incubated for 2 minutes on ice. These steps were repeated 3 to 5 times.

The cell extract could be separated through centrifugation at 10 000 rpm for 20 minutes and 4°C (RC 5C- Sorvall Instruments, DuPont, USA). The obtained supernatant containing the target enzyme was saved at +4°C. In order to increase protein yield, the remaining cell pellet was treated with a chemical lysing reagent (Bugbuster Protein Extraction Reagent, Novagen, Merck, Germany). Afterwards the cell suspension was diluted with binding buffer and another centrifugation step applying the same conditions as mentioned before, was carried out.

2.7 Protein purification

Immobilized **M**etal **A**ffinity **C**hromatography (IMAC) was performed, where histidine residues on the protein bound to the transition metal ion Cu^{2+} and thereby purified the target protein [81].

All buffers (binding and elution buffer), solutions as well as the supernatant containing protein were filtered through 0.45 µm filter (acrodisc syringe filters with supor membrane, 25 mm filtration area, Pall Corporation, USA) and de-gassed prior applying them on the chromatography system ÄKTAprime (GE Healthcare, Sweden). In general the usual applied concentration of imidazole is 20 mM for binding buffer and 250 mM for elution buffer. The

concentration of imidazole only varied for one produced batch of enzyme (10 mM binding buffer and 300 mM elution buffer).

A 5 mL HiTrap FF crude column (GE Healthcare, Sweden) was applied for the protein purification, where firstly 1 M copper-(II)-sulfate was attached as a specific ligand for the histidine-tagged enzyme. After equilibration with approximately 30 ml of binding buffer, the column was loaded with sample. Proteins which do not bind to the column were washed with binding buffer until a straight baseline of the followed UV-measurement was obtained. The elution took place by using a linear gradient of elution buffer and 1 ml fractions were collected. During the elution step a flow rate of 1 ml/min was chosen and the pressure limit was 0.5 MPa. The proteins which were His-tagged could be detected by measuring the UV-absorbance at 280 nm. After separation, the column was treated with 0.5 M EDTA to remove copper. 20% ethanol and Milli-Q-H_2O was used for system wash.

Initial protein concentration of the collected fractions was measured at 280 nm by spectrophotometer NanoDrop (Thermo scientific NanoDrop 1000 spectrometer, USA) and fractions with high concentration were further analyzed.

2.8 Separation of imidazole

The excess amount of imidazole was tried to remove via a falcon tube with a cut-off membrane of 30 kDa. Therefore fractions of purified protein were pooled with 5 ml of cooled 20 mM citrate phosphate buffer (pH 7) and centrifuged for 5 minutes at 4°C and 4000 rpm (Sigma 3-16PK, Germany). After the centrifugation, following steps (1-3) were repeated 3 times:

1. sample was mixed through pipetting to avoid clogging of the membrane surface
2. sample was filled up to 15 ml with buffer
3. sample centrifuged for 10 min using the same conditions as mentioned before

2.9 Analytical methods

2.9.1 SDS-PAGE

For checking the expression as well as purification level **S**odium **D**odecyl **S**ulfate **P**olyacrylamide **G**el **E**lectrophoresis (SDS-PAGE) was carried out according to Laemmli [82].

Preparation of samples collected after induction

The collected samples were lysed with a chemical lysing reagent (Bugbuster Protein Extraction Reagent, Novagen, Merck, Germany), where the added amount varied from 50 µl for the 0 hour sample to 150 µl for 4 hour and overnight sample. After addition the

incubation was carried out for 20 minutes at room temperature. Samples were diluted with phosphate buffer (pH 7.5) to achieve concentrations of around 4 µg/µl for equal sample loading on the gel. After lysing and dilution of the samples a centrifugation step followed lasting for 10 minutes at 13 000 rpm (VWR Galaxy 14D, USA) and room temperature. The obtained cell pellet was again diluted with approximately the same amount as before.

Gel preparation and denaturation

All samples including from the purification step were mixed with sample buffer. Depending on the amount of cell pellet, collected samples after induction were diluted 1:1.6-1.75 (rarely 1:3.5, sample:sample buffer). Whereas purified fractions were mixed in a 1:1.66 (rarely 1:3.3) dilution. Protein samples were boiled at 100°C for 20 minutes to denaturize and then cooled on ice. The prepared 10% gels were loaded per lane with 20 µl of sample and 6 µl of protein marker (1 kDa Precision Plus Protein Standard All Blue-Biorad, USA). The separation was carried out through firstly applying a voltage of 100 V for 15 minutes and afterwards increasing to 200 V for 45 minutes. Gels were stained in Coomassie blue overnight and destained until bands were visible.

2.9.2 BCA

The determination of the total protein content was based on the principle of the **B**icin**c**honinic **A**cid assay (BCA) [83]. Collected fractions showing high protein concentration at pre-measurements with NanoDrop (NanoDrop 1000 spectrometer, Thermo scientific, USA) were diluted 1:5 with 20 mM Tris-HCl (pH = 7.5). A standard series in a range of 0.2 (rarely 0.05) to 1 mg/ml of **B**ovine **S**erum **A**lbumin (BSA) concentrations was prepared for the quantification of protein content. The preparation of the working reagent was carried out by adding one part of 4% (w/v) $CuSO_4$ to 50 parts BCA-solution. Protein samples of 50 µl volume were mixed with 1 ml working reagent. All samples including the standard series were incubated for 40 minutes at 37°C (QBD2, Grant, England). Afterwards the samples were cooled down to room temperature (10-15 minutes) and the absorbance was read at 562 nm (WPA Biowave II, Biochrom, UK).

2.9.3 LC-MS

In addition to theoretical molecular mass determination of the proteins computed with ExPASy tool [84] (SIB-Bioinformatic Resource Portal), the analytical method of mass spectroscopy was applied for more exact determination. Protein samples were injected on the HPLC column (Discovery Bio Wide Pore C5 - 3 µm pore size, L x I.D. 10 cm x 2.1 mm - Supelco, Sigma-Aldrich, USA) in a 1:10 dilution of Milli-Q-H_2O. In order that metabolites can bind to the column via hydrophobic interactions, Milli-Q-H_2O (A) and acetonitrile (B), both

containing 0.1% formic acid were used as eluents. For ionization of the molecules **E**lectron**s**pray **I**onization (ESI) was used as method in the LC-MS system (PE SCIEX Api Qstar Pulsar, Applied Biosystems, USA). A detailed description of the applied gradient program is given in Table 4. Investigation of generated spectra was performed with software Analyst QS (Version 1.1, Applied Biosystems, USA).

Table 4: Gradient program

Step	Total time [min]	Flow rate [µL/min]	A [%]	B [%]
0	0	400	50	50
1	5	400	50	50
2	6	400	20	80
3	12	400	20	80

2.9.4 Enzyme assay

Enzyme activities could be determined by measuring the hydrolysis of specific substrates (aryl and natural substrates).

The method of colorimetry was applied on aryl substrates (**p**ara-**n**itro**p**henyl-substrates, (pNP)), where the change of color was read as absorbance at 405 nm. **β**-**gl**ucosidase (BGL) activity was obtained by following the release of p-nitrophenol from 4-Nitrophenyl β-D-xylopyranoside and for β-**N-a**cetyl**g**lucosaminidase (NAG) by applying 4-Nitrophenyl N-acetyl-β-D-glucosaminide.

High **P**erformance **A**nion **E**xchange **C**hromatography with **P**ulsed **A**mperometric **D**etection (HPAEC-PAD) was performed as method to follow hydrolysis product formation of applied natural substrates. As natural substrates chitobiose was employed for NAG and xylobiose for BGL. The integration of obtained product peak area led to direct information of generated product amount.

The applied enzyme assays can be divided into two groups:

1. <u>Continuous assay</u>

The formation of product respectively the take up of substrate from the enzyme is monitored continuously in real time (UV-1650 PC spectrophotometer, Shimadzu, Japan). For determining the activity of the enzyme towards the used substrate, a molar extinction

coefficient is needed (see Table 5) [85]. Therefore the law of Lambert-Beer (formula 1) was applied to quantify the pNP release during reaction, where a linear as well as quantitative relation between concentration and absorbance was followed. The usage of linear regression led to determination of the extinction coefficients from the slope.

$$A = \varepsilon \cdot c \cdot d$$

Formula 1: Lambert-Beer Law

A ... Absorbance

ε ... Molar extinction coefficient [$M^{-1}\,cm^{-1}$]

c ... Concentration [M]

d ... Thickness [1 cm]

Table 5: Molar extinction coefficients (ε) of pNP in 20 mM citrate phosphate buffer pH 5.6 at 405 nm [85]

Temperature [°C]	ε_{pNP} [ml/µmol*cm]
50	1.3058
60	1.4526
70	1.7849
80	2.4639
90	2.9527

2. <u>Discontinuous assay</u>

In this assay the incubated enzyme reactions were terminated after fixed time intervals, by addition of strong base and therefore leading to the same final pH-value.

For applied aryl substrates each minute (1 to 3 minutes) an aliquot (100 - 200 µl) of sample was always stopped with 1 ml of 0.1 M NaOH. Product formation could be measured by reading the absorbance at 405 nm (WPA Biowave II, Biochrom, UK). Determination of molar extinction coefficient was performed through preparing a series of different pNP-concentrations (0.05 to 0.6 mM) in dH_2O and incubating the samples for 6.5 minutes at room temperature. After incubation same volume of stopping solution was added and

absorbance (duplicates) was measured, as mentioned above. An extinction coefficient of 18 072 M^{-1} cm^{-1} was obtained by applying formula 1.

In the case of employed natural substrates, reactions of 500 - 1000 µl (containing 20 µl enzyme), were run for a particular time for each enzyme. Sampled reactions of 15 µl volume were terminated by addition of 985 µl high pH water (pH 9-10). The amount of product was determined using standards of specific applied substrates with HPAEC-PAD. Details of the HPAEC-PAD specifications are given in Table 6 .

Table 6: Specifications of used HPAEC-PAD

Flow rate	0.5 ml/min
Injection volume	5-10 µl
Detector	conductivity and electrochemical
System	ICS-5000 DIONEX DC (Thermo Scientific, USA)
Column	PA200, PA20
Software	Chromeleon 7.2

In both cases (continuous and discontinuous assay) it was worked with selected temperature and pH ranges. These applied ranges and the used concentrations of enzyme, buffer as well as substrate solutions are specified in the particular analysis procedures. All other information given in these two described assays remains same, such as reaction volume, assay time, determined pNP-coefficients or volume of added stopping solution, if not stated different.

Determination of activity

The released color over time is in a proportional relationship with the measured enzyme activity. One unit of enzyme activity is defined as the amount which catalyzes the hydrolysis of 1 µmol pNP-substrate in 1 minute.

Application of formula 1 helped to determine the amount of product (µM), which was plotted over time (min). Initial velocity V_0 (µM/min) could be obtained from the slope by applying linear regression. V_0 was multiplied with the volume of reaction and the volumetric activity (U/ml) was calculated through division of the volume of enzyme, which was used for

the reaction. The enzyme concentration was referred to the volumetric activity, to obtain specific activity (U/mg).

In order to evaluate the specific activity of applied natural substrates, which were separated by HPAEC-PAD, product amount (µM) was measured though using standards for each substrate. Therefore with the information of how much µmol per liter product was formed, the initial velocity and finally the specific activity could be obtained, as already mentioned above.
The determination of residual activity was carried out by the following formula 2, where activity (volumetric or specific activity) was measured initially (0 min) and related to measured activity at x minutes of incubation at specific temperature.

$$\text{Residual activity } [\%] = \frac{\text{Activity}_0}{\text{Activity}_x} \times 100$$

Formula 2: Determination of residual activity

2.9.5 Substrate specificity

The substrate specificity of NAG was determined using p-Nitrophenyl-β-D-glucuronide (pNPG), p-Nitrophenyl-β-D-xylopyranoside (pNPX), p-Nitrophenyl-β-L-arabinopyranoside (pNPA), p-Nitrophenyl-β-D-cellobioside (pNPC), p-Nitrophenyl-α-D-mannopyranoside (pNPM), p-Nitrophenyl-β-D-galactopyranoside (pNPGal), p-Nitrophenyl-α-L-rhamno-pyranoside (pNPR) and p-Nitrophenyl-N-acetyl-β-D-glucosaminide (pNPGlcNAc). Color formation of pNP-substrates was followed continuously at 60°C and 70°C, by using an final enzyme concentration of 0.5 µM in 100 mM citrate phosphate buffer (pH = 5.6) and mixed with an end concentration of 0.74 µM substrate.

BGL activity (enzyme concentration 0.8 mg/ml) applying 1 mM of 1,4-β-D-cellobiose, 1,4-β-D-cellotriose, 1,4-β-D-cellotetraose, 1,4-β-D-xylobiose, 1,3-β-D-laminaribiose, 1,3-β-D-laminaritriose, xyloglucan and lichenan was tested at 60°C**.** Whereas the specificity for NAG (enzyme concentration 1.09 mg/ml) towards natural substrates was analyzed using 1 mM of chitooligosaccharides (di-N-acetyl-chitobiose to penta-N-acetyl-chitopentaose) at 70°C. Quantification of used natural substrates was not performed, but activity of the enzymes towards them was determined through monitoring production profile until 1 hour. The methods of HPAEC-PAD for separating natural substrates are given in 2.10.1 and 2.10.2.

2.9.6 Thermal stability

Determination of activation energy

The influence of temperature on the enzymes was measured through continuous assay. Besides determining the activation energy, the maximal specific activity ranging between 50°C to 90°C was also investigated for β-N-acetylglucosaminidase in this assay (temperature optimum). Therefore 773.7 µl of 20 mM citrate phosphate buffer (pH 5.6) was pre-incubated for 3 minutes with 200 µl of 3.68 µM 4-Nitrophenyl N-acetyl-β-D-glucosaminide (end concentration 0.74 mM). The absorbance of the 1 ml reactions were followed for 3 minutes at 405 nm (UV-1650 PC spectrophotometer, Shimadzu, Japan) after addition of enzyme (end concentration 0.3 µM).

The equation of Arrhenius (see formula 3) leads to obtain the activation energy, after taking the natural logarithm. The plotting of ln k against 1/T provides a slope of $-E_a/T$.

$$k = A * e^{\frac{-E_a}{RT}}$$

Formula 3: Arrhenius equation

k ... Reaction rate [M/s]

A ... Pre-exponential factor (independent of temperature)

E_a ... Activation energy [kJ/mol]

R ... Gas constant [J/mol*K]

T ... Temperature [K]

Determination of half-life

Analyzing the enzymes, based on their thermal inactivation, BGL was incubated at 70°C and 80°C, whereas for NAG the temperatures of 80°C as well 90°C were applied (HLC Thermo Shaker MKR 13, Ditabis, Germany). Aliquots were sampled every 30 minutes between 0 to 180 minutes and each sample was incubated on ice for approximately 5-6 hours. Afterwards a centrifugation step of 1 minute at 13 000 rpm and 4°C (Sigma 1-14K, Germany) was followed to obtain the protein containing supernatant.

In order to determine specific activity, discontinuous assay was carried out with 0.1 M citrate phosphate buffer (pH 5.6) and 5 mM pNP-substrate (end concentration 1 mM). Reactions of 400 µl volume with 0.8 mg/ml β-glucosidase (V_{enzmye} = 25 µl) were run at 60°C (HLC Thermo Shaker MKR 13, Ditabis, Germany), whereas for β-N-acetylglucosaminidase an enzyme concentration of 1.09 mg/ml (V_{enzmye} = 20 µl) was used in 500 µl, performing analysis at 70°C

(HLC Thermo Shaker MKR 13, Ditabis, Germany). The continuation of discontinuous assay remained same as mentioned in 2.9.4.

The thermal inactivation was investigated through calculating the residual activity (see formula 2). In addition for further understanding of enzyme stability, the half-life was calculated. Therefore the natural logarithm of the residual activity (without unit) was plotted against time (SigmaPlot, version 12.5). The obtained slope was inserted in the following equation (formula 4)

$$t_{\frac{1}{2}} = \frac{\ln 2}{k}$$

Formula 4: Half-life

k ... Reaction rate [min^{-1}]

$\mathbf{t_{1/2}}$ ... Half-life [min]

2.9.7 Kinetic study

The investigated enzyme NAG was more specified in its catalyzing reactions through analyzing enzyme kinetics. Reactions were carried out in 0.1 M citrate phosphate buffer pH 5.6 and pNP-substrate ranging from 0.05 to 0.4 mM concentration. The added enzyme end concentration was 0.06 µM. Absorbance was monitored continuously (5 minutes) at 405 nm and 70°C (UV-1650 PC spectrophotometer, Shimadzu, Japan).

For describing the kinetic parameters, the equation of Michaelis-Menten [86] was applied (see formula 5)

$$V_0 = \frac{V_{max}[S]}{K_m + [S]}$$

Formula 5: Michaelis-Menten Kinetics

The reciprocal of formula 5 provides the linear equation (formula 6) below, which is known as Lineweaver-Burk relation [87].

$$\frac{1}{V_0} = \frac{K_M}{V_{max}} \cdot \frac{1}{[S]} + \frac{1}{V_{max}}$$

Formula 6: Lineweaver-Burk

V_0 ... Initial velocity [µmol/(min*mg)]

V_{max} ... Maximal velocity [µmol/(min*mg)]

K_m ... Michaelis constant [mM]

[S] ... Substrate concentration [mM]

Plotting 1/V against 1/[S] obtains the intercept of $1/V_{max}$ and a linear slope of K_m/V_{max}. In addition to understand how many molecules of substrate can convert one molecule of enzyme into product per unit of time, the turnover number k_{cat} (formula 7) was defined [88], [89].

$$k_{cat} = \frac{V_{max}}{e_t}$$

Formula 7: Turnover number

k_{cat} ... Turnover number [s^{-1}]

e_t ... Initial enzyme concentration [mM]

The calculated data was also fitted with the curve fitting tool from SigmaPlot (Version 12.5). Therefore it was possible to compare the evaluated kinetic parameters to those from the provided Michaelis-Menten model.

2.9.8 Affinity gel electrophoresis

The experiment of affinity gel electrophoresis was carried out at the Danish Technical University. Native-PAGE gels (6.5%) were used for detecting the binding of enzymes towards polymeric substrates, through retarding the movement of the proteins that interact with them. Substrates **o**atspelt **x**ylan (OSX), **h**ydroxy**e**thyl **c**ellulose (HEC), **b**arley **b**eta-**g**lucan (β-glucan) and **x**ylo**g**lucan (X-glucan) were prepared in a 0.1% concentration. The gel-solutions were de-gassed in sonication water bath for about 10 minutes. Gel casting

could take place after adding riboflavin, APS and TEMED to the solution and solidification lasted for around one hour. All enzyme samples were diluted using 5x loading buffer to reach an end concentration 0.5 mg/ml. A volume of 4 µl protein ladder (Native Mark, Invitrogen, Life Technologies, USA) was mixed with the same volume of protein sample. The performed conditions for running the gels were 90V, 3.5 hours and 4°C. Gels were washed 5 minutes with dH_2O, stained overnight in SimplyBlue SafeStain (Invitrogen, Life Technologies, USA) or Coomassie Blue and destained until bands were visible.

2.10 Optimization study

The determination of optimal pH and temperature involves the study of both parameters at a time, leading to detect a range, in which enzymes show higher activity. Therefore, such a study helps to understand the parameters interactions and their influence of enzymatic activity (Design of Experiments). The different experimental setups for each of the two enzymes are given in the following and enzyme activity was calculated as described in 2.9.4.

2.10.1 Experimental setup for BGL

The activity of β-glucosidase was investigated on xylobiose by measuring the amount of produced xylose through HPAEC-PAD. The enzyme was diluted 20 times in 20 mM citrate phosphate buffer of pH 5, resulting in a concentration of 0.014 mg/ml. Enzyme samples were incubated at three different temperatures 55°C, 60°C, 65°C and shaken at 400 rpm (HLC Thermo Shaker MKR 13, Ditabis, Germany), within a pH-range from 3.5 to 7.8 (using 20 mM citrate phosphate and potassium phosphate buffer). Reactions were performed at specific time points (0, 10, 15, 20, 30 minutes) in duplicates. The applied method for the separation of the disaccharide is given below.

HPAEC-PAD method for BGL

I. Pre-equilibration with 16 mM NaOH for 2.5 minutes
II. 0-23 minutes run with 16 mM NaOH
III. 23-25 minutes run with 200 mM NaOH

The determined activity of BGL by applying xylobiose in the above mentioned range, helped to further follow the influence of temperature and pH on pNPX. BGL specific activity was determined at 60°C and 65°C (HLC Thermo Shaker MKR 13, Ditabis, Germany) from pH 5.4 to 7.3 (100 mM citrate phosphate buffer), in duplicates. Enzyme reactions of 800 µl (containing 20 µl of 0.8 mg/ml enzyme) with 1 mM end concentration of pNPX were carried out under standard conditions in a discontinuous assay.

2.10.2 Experimental setup for NAG

Screening

The parameters (pH and T) were selected to be investigated in a broad range between pH 4.3-7.8 and 60°C to 80°C (HLC Thermo Shaker MKR 13, Ditabis, Germany) in a discontinuous assay, aiming to find optimum activity for NAG. 100 mM citrate phosphate buffer (pH 4.3-7.4) and 100 mM potassium phosphate buffer (pH 7.8), were used. A final concentration of 1 mM pNPGlcNAc was applied in 200 µl reaction volume, where 5 µl of 0.68 mg/ml enzyme were added. Experiment was continued as usual under standard conditions.

Design of experiments - DoE

The application of statistical experimental planning helps to identify optimal dataset of employed factors and to describe the relationships among them and obtained response. Therefore, screening analysis on aryl substrates allowed to narrow the analyzed range of pH and temperature, which was then further investigated on the natural substrate, chitobiose. For this purpose **R**esponse **S**urface **M**odeling (RSM) was employed, choosing a **C**entral **C**omposite **F**ace-centered design (CCF) as the experimental design, which involves 11 runs. These runs refer to four cube points, four axial points carried out in duplicates and three central points for the estimation of experimental error. The design of experiment is given in Figure 9, conducing as the experimental base and each point represents an enzyme reaction run at specific parameters. As response (dependent variable), the integrated relative peak area [%] of consumed reactant chitobiose was used, after 40 minutes of enzyme incubation. The MODDE 9.1 software (Umetrics, Sweden) was used to analyze the obtained experimental data.

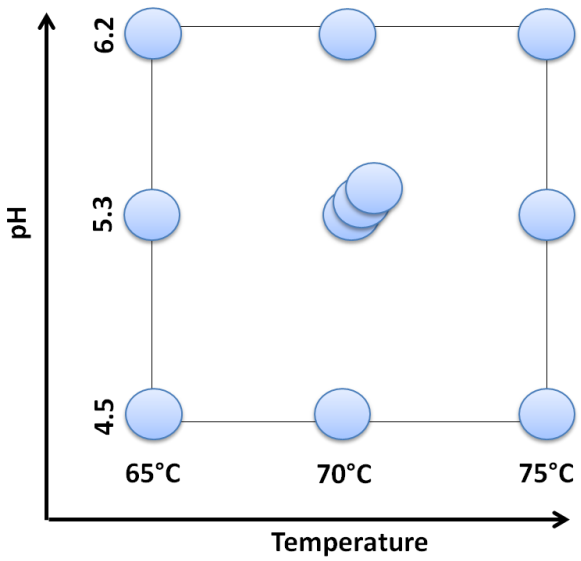

Figure 9: Illustration of experimental design space revealing the variation of temperature and pH, at which the x-axis indicates the different applied temperatures, whereas y-axis describes the used pH-values. Each circle represents an individual enzyme reaction, which was run at specific time points. In the middle of the design, the triplicate center points are positioned (70°C and pH 5.3).

The enzyme reactions were run at the specified range in Figure 9. NAG was diluted 20 times in a 20 mM citrate phosphate buffer of pH 6 and was added to the reaction vial with an 0.055 mg/ml enzyme concentration, containing 1 mM chitobiose as substrate. For the different pH range (4.5 - 6.2) a 20 mM citrate phosphate buffer was used. The method, which was applied to measure the amount of produced N-acetylglucosamine by HPAEC-PAD is described below. Hydrolysis product formation was monitored after 0, 10, 20, 30 and 40 minutes.

HPAEC-PAD method for NAG

I. Pre-equilibration with 200 mM NaOH and 4 minutes with 5 mM NaOH and 5 mM NaAc
II. 0-20 minutes run with 5 mM NaOH and 5 mM NaAc

3 Results

3.1 Results of BGL and NAG

In this section, the joint results of performed primary structure investigation are presented as well the primarily experimental results of subcloning performance.

Initially besides analyzing Rmar_0925 (β-N-acetylglucosaminidase), the enzyme GH3-2069-C1 containing no His-tag should have also been characterized. For purification process, the gene without His-tag was subcloned without success. The used restriction enzymes did not cut correctly during the digestion and also a very low plasmid concentration was measured (Figure 34, appendix). Hence, the enzyme Rmar_2069 was replaced from an already primary characterized Rmar_1080 (β-glucosidase) enzyme.

β-glucosidase encodes 792 **a**mino **a**cids (aa) compared to the larger β-N-acetylglucosaminidase amino acid sequence of 948 aa, both excluding the His_6-tag tail at the C-terminus as well as the signal peptide sequences. Analysis of these proteins with BlastP and Pfam identified BGL and NAG as members of the glycoside hydrolase family 3 of carbohydrate active enzymes, which are classified to EC 3.2.1.21 number for β-glucosidase and EC 3.2.1.52 number for β-N-acetylglucosaminidase. BGL revealed similarity to various β-glucosidases from family 3 for example 51% identity to β-glucosidase from *Melioribacter roseus* [90] (Genbank accession no. YP_006528319.1). A number of homologies were also obtained for NAG for instance 34% identity to β-N-acetylglucosaminidase from *Cytophaga hutchinsonii* [91] (Genbank accession no. YP_676647.1).

Multiple **s**equence **a**lignment (MSA) allowed to characterize the three domain NAG as well BGL among other β-glucosidases and β-N-acetylglucosaminidases from different organisms, but same glycoside hydrolase family 3. The complete MSA is given in the appendix, Figure 39. These enzymes differentiate in their third domain, which is a fibronectin type III domain for BGL and a β-lactamase domain for NAG. BGL shared the highest similarity with β-glucosidase of TnBgl3B from *Thermotoga neapolitana* [61] and ExoI from *Hordeum vulgare* [60], in both cases around 29%. Whereas with *Aspergillus aculeatus* β-glucosidase AaBGL1 [63] BGL provided an identity of approximately 18%. The alignment of NAG with β-N-acetylglucosaminidases from *Bacillus subtilis* BsNAGZ [65] and *Alteromonas sp*. Cht60 [92] showed similarity of around 25%. Comparing BGL with NAG a score of 16% was obtained. In general the overall conservation of applied amino acid sequences was less. The MSA revealed some conserved residues in the first two domains of the proteins than compared to the third domain, where no conservation was detected, since only together with NAG and BGL, TnBgl3B as well AaBGL1 are composed of three domains.

The catalytic nucleophile aspartic acid is conserved in all GH3 family members [61] (highlighted in Figure 10A) and located for all aligned sequences in domain 1. Deduced from the MSA, D306 in BGL and D283 in NAG were determined as the nucleophiles in the catalytic reaction. The prediction of the catalytic acid/base residue is difficult. At the position where TnBgl3B and AaBGL1 have their acid/base residue [61], [63], BGL shows an asparagine residue (N537), neighbored with glutamic acid (E538). Determining catalytic acid/base of BGL compared to the acid/base residue of ExoI [60], two glutamic acids were obtained at E518 and E519 for BGL.

For NAG a totally different result for acid/base residue predication was revealed. NAG and Cht60 showed conservation at the determined catalytic acid/base residue of β-N-acetylglucosaminidase from *Bacillus subtilis* [65], which is located in the D(S/T)H motif and this motif is missing in β-glucosidases (see Figure 10D). In addition the KHFPGHG motif is highly conserved among the investigated β-N-acetylglucosaminidases, compared to β-glucosidases. Therefore multiple sequence alignment helped to predict and identify H209 together with D207 as a probable candidate for the catalytic acid/base residue in NAG.

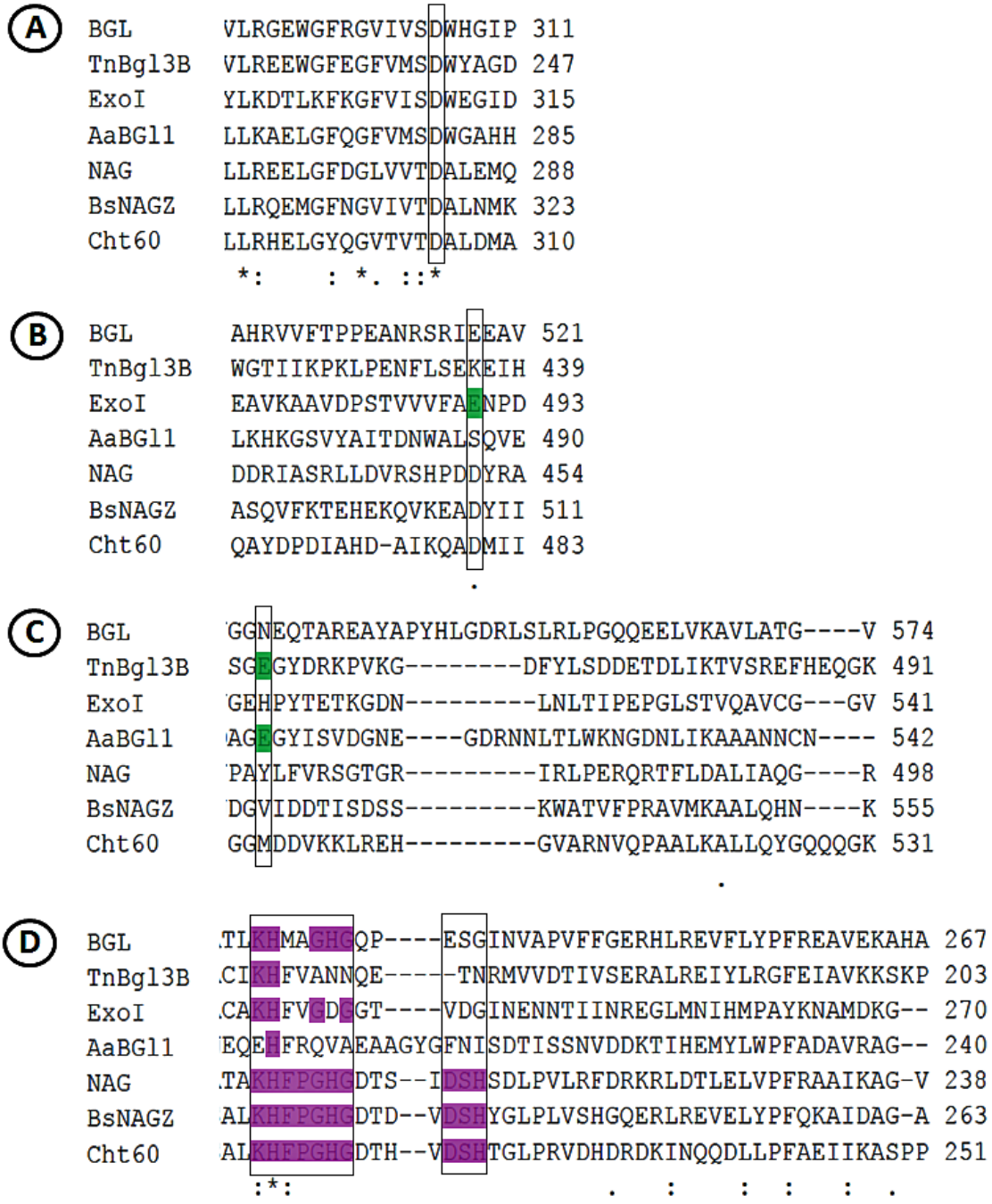

Figure 10: Partial amino acid sequence alignment of GH3 family β-glucosidases and β-N-acetylglucosaminidases obtained from ClustalW2 program [79], [80]. The amino acid sequence of the investigated β-glucosidase (BGL) and β-N-acetylglucosaminidase (NAG) from *Rhodothermus marinus* were compared to the β-glucosidase from *Thermotoga neapolitana* TnBGl3B (UniProt: Q0GC07), *Hordeum vulgare* ExoI (UniProt: Q9XEI3) and *Aspergillus aculeatus* AaBGl1 (UniProt: P48825) as well as to the β-N-acetylglucosaminidase from *Bacillus subtilis* BsNAGZ (UniProt: P40406) and *Alteromonas sp. strain O-7* Cht60 (UniProt: P48823). Sequences applied for MSA of mentioned GH3 family members from various organisms were obtained from UniProt [78]. Identical or similar amino acid residues of the aligned GH3 family members are denoted by asterisks and dots, respectively. The inverse triangle indicates the conserved residue of the catalytic acid/base, which is given in figure A. In figure B and C, the catalytic nucleophile of the aligned β-glucosidases is highlighted in green color. Whereas the catalytic acid/base residue, which is located in the D(S/T)H among aligned β-N-acetylglucosaminidases is boxed purple and is missing in investigated β-glucosidases. The highly conserved KHFPGHG pattern, which is found in aligned β-N-acetylglucosaminidases, does only occur partially in the amino acid sequences of investigated β-glucosidases.

3.2 Part I - BGL

First the experimental results of investigated β-glucosidase (Rmar_1080, BGL) are shown, which are divided in a first part of expression of the protein and the second part is based on detailed characterization. These results are then subsequently discussed and finally conclusions are drawn.

3.2.1 Results

3.2.1.1 Expression of recombinant protein

As an example for the recombinant protein production of BGL, one experimental run is given in Figure 11, which lasts for approximately 24 hours. An OD_{620}-value above 0.5 was reached after around 4 hours, where induction with rhamnose took place. The small bend between the time interval of 6 to 7 hours, results from the not high increase of biomass. Additionally it can be observed that the stationary phase was reached already after a process time of about 10 hours cultivation.

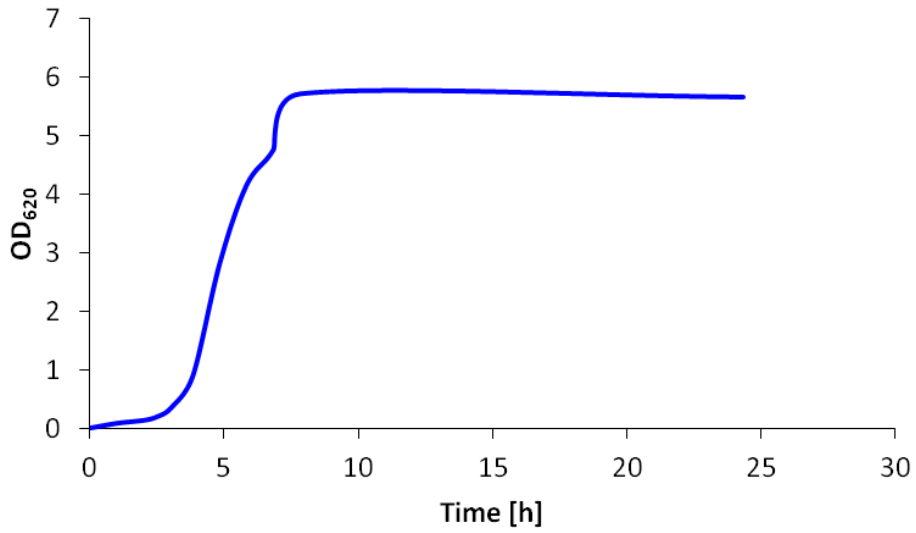

Figure 11: Recombinant protein production and expression of β-glucosidase. The shake flask cultivation was carried out at 37°C and biomass production was measured at OD_{620}. Process time lasted for approximately 24 hours.

Purification of the protein of interest could not be carried out due to column based problems. Therefore all further analysis of β-glucosidase were carried out with an already prepared batch of purified enzyme.

3.2.1.2 Characterization

3.2.1.2.1 Determination of molecular weight with LC-MS

The ExPASy tool [84] helped to predict the theoretical molecular weight of β-glucosidase (88.09 kDa), which was detected and separated through SDS-gel (86 kDa). Thereby this predetermination made it easier to analyze specific area of the obtained m/z results.

For the LC-MS experiment, a protein concentration of 80 µg/ml was used. The eluted peak between 10.35 and 10.74 minutes is given in the LC-chromatogram, Figure 12A and the corresponding MS spectra is shown in Figure 12B. A molecular mass of 87.15 kDa was determined for β-glucosidase, which is about 1.8 kDa less than the theoretical weight. The obtained m/z plot shows more noisy data than a real pattern.

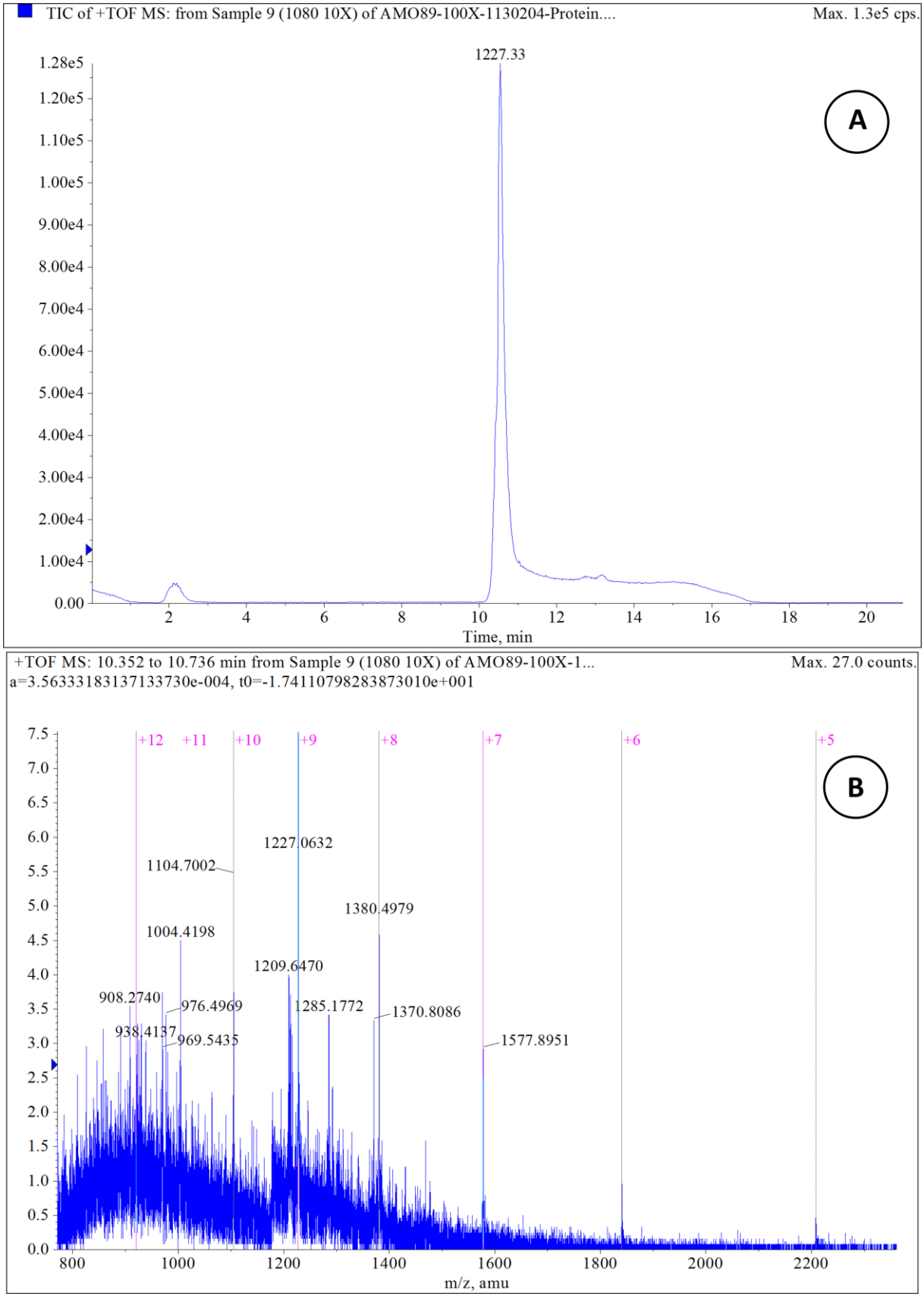

Figure 12: Molecular weight determination of β-glucosidase via LC-MS. The separated protein and eluted peak by LC shows a retention time of 10.35 to 10.74 minutes in figure A. Corresponding m/z spectra is given in figure B, where the molecular weight of β-glucosidase with 87.15 kDa was determined.

3.2.1.2.2 Substrate specificity

β-glucosidase was primary screened on six different aryl substrates, in previous work of M. Jenkins [93] and these results were used for further detailed characterization. Based on them, almost similar activity was obtained for p-Nitrophenyl-β-D-glucopyranoside (pNPG) and p-Nitrophenyl-β-L-arabinopyranoside (pNPA). Whereas highest activity was determined for p-Nitrophenyl-β-D-xylopyranoside (pNPX), which was used for all further assays. Activity screening on p-Nitrophenyl-β-D-cellobioside (pNPC), p-Nitrophenyl-α-D-mannopyranoside (pNPM) and p-Nitrophenyl-β-D-galactopyranoside (pNPGal) could not be detected [93].

For the enzyme BGL broad substrate specificity on natural substrates was detected (Table 7). The screening was carried out through monitoring the hydrolysis products of various carbohydrates by applying HPAEC-PAD. On cellotriose, cellotetraose as well on laminaritriose almost similar activity was obtained, whereas on cellobiose, xyloglucan as well as lichenan no activity could be measured. The highest activity of β-glucosidase was determined on xylobiose and laminaribiose. Further investigation with the appliance of natural substrates was initiated with xylobiose.

Table 7: Screening of β-glucosidase on various natural substrates

Substrate	Activity
1,4-β-D-Cellobiose	N.D.
1,4-β-D-Cellotriose	++
1,4-β-D-Cellotetraose	++
1,4-β-D-Xylobiose	+++
1,3-β-D-Laminaribiose	+++
1,3-β-D-Laminaritriose	++
Xyloglucan	N.D.
Lichenan	N.D.

N.D. Not detected, ++ moderate, +++ high

As an example for the separation of the applied carbohydrates, a chromatogram of cellotetraose is given in Figure 13, representing its hydrolysis products. In the previous work of M. Jenkins [93], it was discussed that β-glucosidase shows tendency to precipitate, which therefore led to not perform a dialysis step after purification [93]. Since, the content of

imidazole, which resulted in baseline disturbance during applying HPAEC-PAD, did not hinder the integration of the product peaks, analysis was continued with containing imidazole. In Figure 13, a ghost peak of unknown source was also observed, which did not influence analysis.

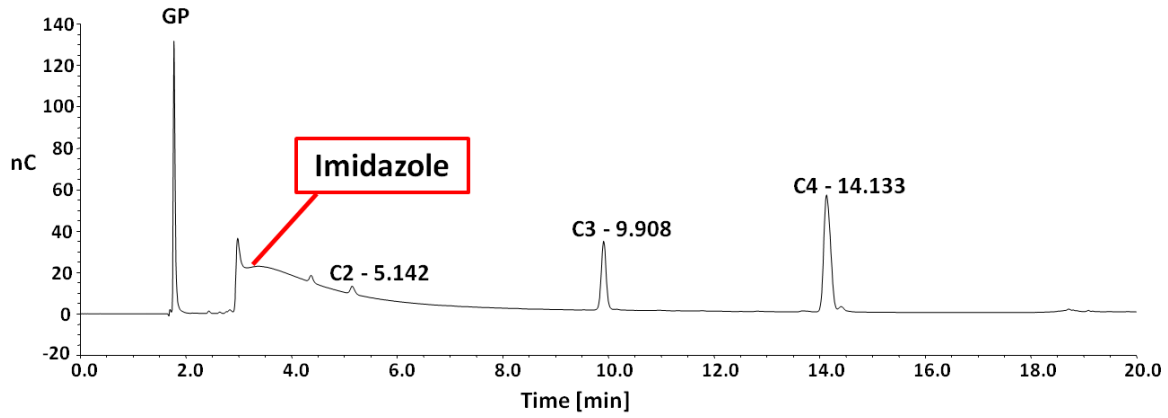

Figure 13: Example of a chromatogram representing a hydrolysis profile of β-glucosidase with cellotetraose as substrate. The formed hydrolysis products of cellotetraose (C4), which are cellotriose (C3) and cellobiose (C2) are indicated along with their retention time of elution. The product peak of cellobiose is located near the baseline disturbance, which is produced by the content of imidazole. A ghost peak (GP) can be found shortly before two minutes of applied method run.

3.2.1.2.3 Thermal stability

In order to understand how temperature influences enzymatic activity, a temperature profile was measured between 50°C to 90°C, in the above mentioned project work (M. Jenkins). This data was used for forming a view on selecting further working conditions with β-glucosidase. The highest specific activity was obtained at 60°C (temperature optimum) and the lowest at 50°C. After 60°C the activity was descending [93].

How stable the enzyme is at a specific temperature could be detected through looking at the thermal stability. Therefore β-glucosidase was incubated at 70°C and 80°C between 0 to 180 minutes. A plot of comparing the residual activity for β-glucosidase at both temperatures is given in Figure 14. Higher residual activity was measured for 70°C. The first 30 minutes of incubation result in a very rapid decrease of activity left for β-glucosidase at 80°C (29%) compared to 70°C, where residual activity is about 75%. Analyzing the trend of 70°C incubation temperature, it can be observed that the activity declines until 30 minutes and surprisingly increases with time. Whereas at 80°C the thermal inactivation occurs very fast and is followed by a slight sudden rise of residual activity at 60 minutes. However, towards the end the activity of β-glucosidase remained almost the same.

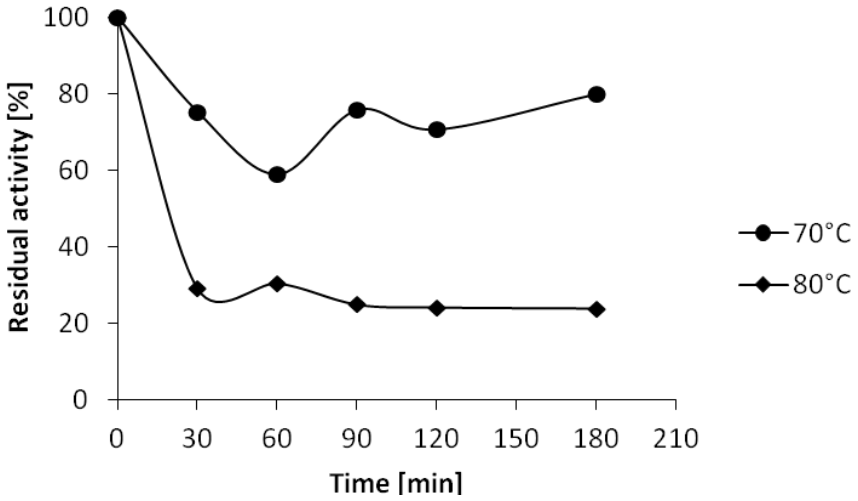

Figure 14: Thermal stability curves of β-glucosidase. Residual activity was determined at 70°C and 80°C between 0 and 180 minutes, in a time interval of 30 minutes. Inactivation of β-glucosidase occurred faster at 80°C. After 180 minutes of incubation around 80% residual activity was left at 70°C and 24% at 80°C.

For further characterization of thermal stability another parameter, the half-life was determined. An apparent half-life could only be obtained through taking the linear region of 70°C between 0 to 60 minutes of incubation time and resulting in 78.77 minutes (Figure 15).

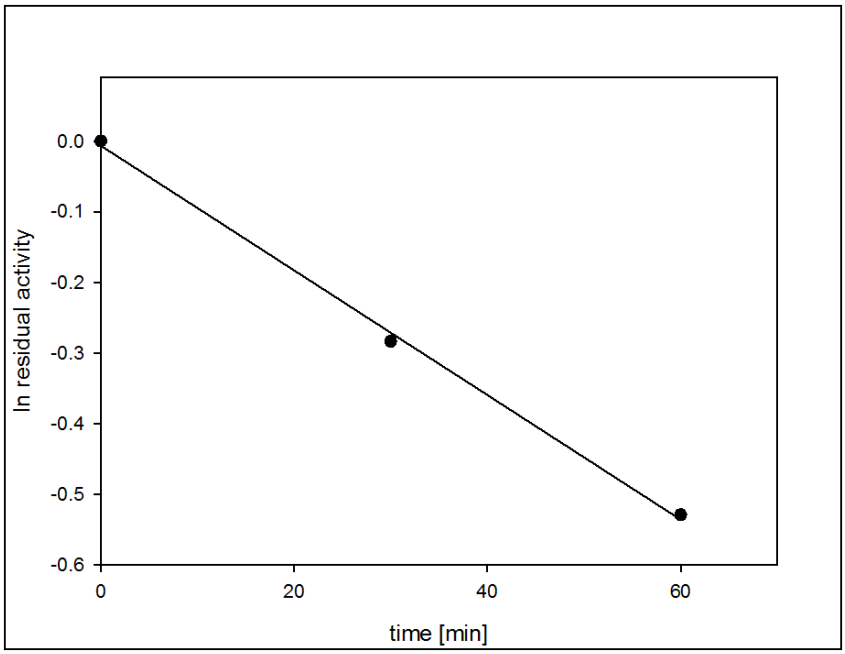

Figure 15: Determination of half-life through measuring residual activity of β-glucosidase by applying pNPX as substrate. The natural logarithm of residual activity is plotted against assay incubation time between 0 to 60 minutes at 70°C. Apparent half-life resulted in 78.77 minutes.

3.2.1.2.4 Detection of enzyme binding to polymeric substrates - Affinity gel electrophoresis

The examination of binding of β-glucosidase with four different applied substrates (barley beta glucan, hydroxyethyl cellulose, oatspelt xylan and xyloglucan) on native gels is given in the appendix, Figure 35 and Figure 36. The control is always given on the left side of the figures, whereas the analysis of substrate binding with enzyme can be found on the right side of each specific figure. Besides β-glucosidase, also other glycoside hydrolase enzymes from family 3 were analyzed.

In order to verify if β-glucosidase interacted with a polymeric substrate, this can be followed in the way that the enzyme will be slowed in its movement. All applied substrates did not bind to β-glucosidase and also no real observable substrate binding was detected from the other analyzed GH3 enzymes.

3.2.1.2.5 Effect of pH and temperature on enzyme activity - optimization study

The investigation of varying two independent variables (temperature, pH) helped to optimize a range, where the enzyme has higher activity. Based on the results of temperature optimum determination, where BGL showed highest activity at 60°C, two more temperature levels could be added (55°C, 65°C) and therefore a broader area covered. Since the pH region has not been analyzed, a pH range from 3.5 to 7.7 was examined at the mentioned temperature values on xylobiose.

BGL hydrolyzed the disaccharide xylobiose to xylose and the specific activity of the enzyme could be therefore determined by the formed product amount. Due to instability of measured product amount at pH 3.5 (especially at higher temperature 60°C as well 65°C), determination of activity was not possible (data not shown). Hence, this specific point (pH 3.5 at all applied temperatures) was not included in this investigation.

The obtained activity of the duplicate runs is plotted against the analyzed range of pH in relation to temperature (see Figure 16). At 55°C the highest activity was measured at pH 4.4 and later on the activity remains similar at pH 5.4 to pH 6. Whereas at 60°C almost similar specific activity values were obtained until pH 7, where the highest response was determined. The 65°C data showed two maximums of activity. One at pH 5.4 with a lower activity value and another at pH 7. Comparison of all temperature points provided the information that at pH 4.4 highest activity was measured for 55°C and at pH 5.4 the temperatures 60°C as well 65°C indicated similar values. This investigation revealed a shift of optimum with temperature.

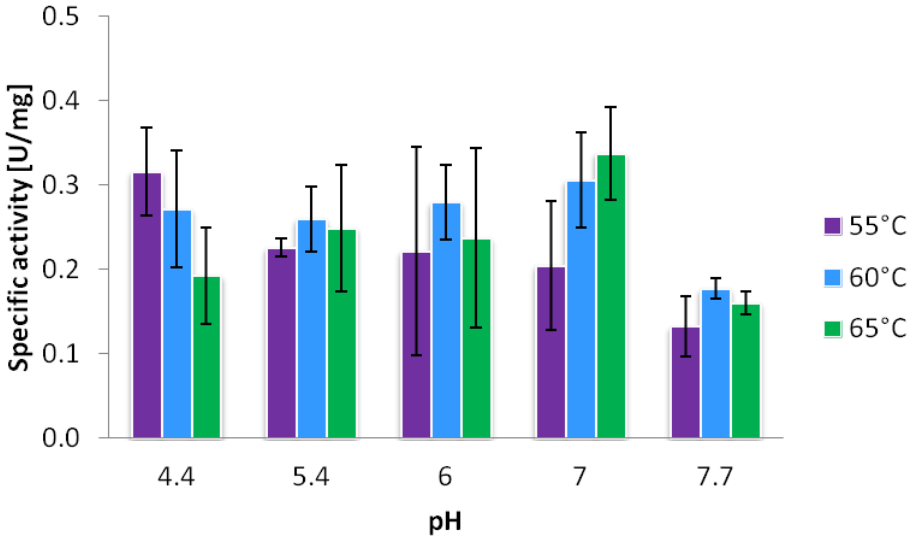

Figure 16: Optimization of specific activity of β-glucosidase with xylobiose as substrate. Determined activity is plotted against a broad range of pH-values in regard to three different applied enzyme assay incubation temperatures.

The individual enzyme reactions provided high variation between each run. In order to describe the differences between the duplicates, each run is represented in the appendix, Figure 43. Clearly the second replicate (Figure 43B) shows that at every applied temperature the enzyme intersects at pH 6 and at pH 7.7, revealing almost same response. At 55°C two maxima were obtained at pH 4.4 and 6.

The analyzed range, which was applied for BGL on xylobiose, helped to design and narrow the investigation area for high activity determination on pNPX. Stable results on the used natural substrate were obtained at 60°C and 65°C. At these temperatures, pH-value 7 revealed maximum activity. Therefore a pH range between 5.3 to 7.3 was selected to be examined at already mentioned temperatures. In Figure 17 the performance of β-glucosidase on pNPX is represented. A trend can be followed, which shows that with higher temperature the activity increases. In both cases high activity is reached at pH 5.4 and declines with increasing pH.

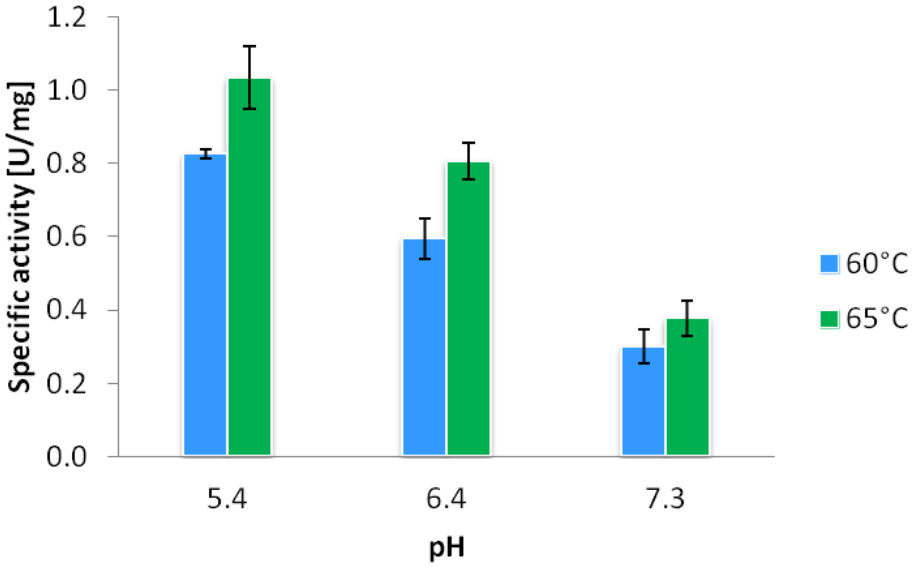

Figure 17: Determination of optimal specific activity of β-glucosidase by varying pH and temperature. The obtained specific activity is plotted against the investigated range of pH 5.4, 6.4 and 7.3 at 60°C as well 65°C, by using pNPX as substrate. For both investigated temperatures highest specific activity was determined at pH 5.4. Activity results of applied 65°C temperature point revealed always higher specific activity compared to 60°C at any tested pH value.

Finally comparing the specific activity of xylobiose and pNPX in the same range leads to the understanding that β-glucosidase performs totally opposite on both substrates. On xylobiose the activity increases from pH 5.4 to 7 (maxima), whereas on pNPX high activity was obtained at pH 5.4 and decreased with pH value. Besides optimal values were identified, the effective investigation if the applied parameters significantly influence enzyme activity, which should have been validated based on statistical methodology, was difficult. Also, studying the variables interaction and building models, by applying RSM revealed to be complicated, due to variability of obtained data from the assay on xylobiose.

3.2.2 Discussion

Nowadays thermostable β-glucosidases have found great interest in industry as well as in research area. Several GH3 β-glucosidases have already been characterized not only from archaeal and eukaryotic branch of the tree of life [25], [26], but also from various thermophile bacterial organisms such as *C. thermocellum* [35], *T. thermophilus* [94], or *T. thermosaccharolyticum* [95].

In this research a thermostable β-glucosidase (GH3 family, EC 3.2.1.21) from the marine bacteria *Rhodothermus marinus* was expressed in *E.coli* and tried to purify with the help of IMAC. While loading supernatant containing protein, the copper metal ions started to strip from the IMAC column, which led to the lost of produced protein. BGL has been produced and purified previously and no problems during the purifications step were detected, see work of M. Jenkins [93]. It is presumed that the protein sample, which contained lysing buffer was present in a concentrated form and not diluted properly, therefore supposably leading to the stripping of copper. For all further investigations, an already prepared and purified stock of BGL protein was used. As already mentioned BGL was primarily characterized (work of group member M. Jenkins), analyzing substrate specificity on various aryl substrates as well as detecting optimal temperature of BGL on the most active substrate (pNPX). Based on these results further characterization of BGL was planed.

The calculated molecular weight of BGL, based on the 792 amino acid residues and 6 histidine residues revealed a higher value (89.02 kDa) compared to the mass analyzed by SDS-PAGE (approximately 86 kDa). The obtained molecular weight of 86 kDa was similar with other family 3 β-glucosidases from thermophilic bacteria such as *Dictyoglomus turgidum* (85 kDa) [96] and *Thermotoga maritima* (81 kDa) [97], as well as with a glucosidase from thermophilic fungi *Myceliophthora thermophila* (90 kDa) [98]. In addition, the determined weight of BGL by LC-MS resulted in a molecular mass of 87.18 kDa, which is 1.84 kDa less than compared to the computed mass. Various explanations may exist to describe the molecular weight difference. One of these could be that if an amount of other proteins (e.g. *E.coli* proteins) exists, then it is possible that the produced fragments by ions overlap with the ones of the protein of interest, making analysis difficult. Another method to exactly determine the molecular mass of BGL could be through analyzing the sequence by N-terminal sequencing.

Furthermore, when searching for other thermophilic bacterial strains regarding to examine their substrate specificity, several exhibited higher affinity towards pNPG [13], [35], [36], [96] compared to pNPX. This is also true for fungal β-glucosidases [99], [100], [101], [102].

Interestingly, analyzed BGL showed highest activity on pNPX, followed by almost similar activity on pNPG and pNPA [93]. Whereas no activity could be detected on pNPM, pNPC as well pNPGal [93]. On the latter mentioned aryl substrate, thermostable β-glucosidase showed some activity, for example from *C. thermocellum* [35], an *Aeromonas sp.* [13] or *D. turigidum* [96].

In addition it could be pointed out that the characterized enzyme is a broad-specificity β-glucosidase, because it revealed activity not only on aryl substrates, but also towards natural substrates. The enzyme showed high activity especially on xylobiose as well on laminaribiose. The outcome is convenient, because the first mentioned polysaccharide can be found in pines cell wall, whereas laminaribiose in algae. Similar activity was detected on cellotriose, cellotetraose and laminaritriose, but no activity was monitored by β-glucosidase on cellobiose, lichenin and xyloglucan. Analogous results were obtained from the fungi *T. emersonii*, where no activity was detected on lichenan and xyloglucan [102]. Most characterized β-glucosidases were not active towards xylobiose, such as from a *Thermotoga sp.* [36]. Whereas one out of the numerous characterized β-glucosidases exists, which belongs to *A. oryzae*, revealing moderate activity towards xylobiose [100]. In most cases, where cellobiose was applied to examine substrate specificity, it showed high activity. Such as for the β-glucosidases of *A. fumigatus* [44] or *Thermotoga sp.* [36]. The thermophilic bacteria *C. thermocellum* [35] and fungi *A. oryzae* [100] provided almost similar results compared to BGL, where higher activity was measured on laminaribiose followed by laminaritriose and the cellooligosaccharides.

In summary, BGL provides ability to hydrolyze short chain oligosaccharides, acting from the non-reducing end. The BGL enzyme removed successfully glucose-units from laminaribiose, -triose (β-1,3), whereas for the cellooligosaccharides more than 2 glucose molecules are required, which are β-1,4 linked. In the case of xylobiose, where 2 xylose units are associated by β-1,4-linkage, were sufficient for BGL hydrolysis process. Among GH3 enzymes it is not uncommon that they exhibit several activities, which was the case for various β-glucosidases, revealing β-glucosidase as well as β-xylosidase activity [37], [49], [103]. Hence, it is believed that BGL probably is bifunctional, because it showed activity towards pNPG as well pNPX and hydrolyzed cellooligosaccharides and xylobiose. This feature of BGL should be further investigated by applying pNPG and cellooligosaccharides. Since such a characteristics makes it difficult to specify them by their natural substrates and also to classify the enzymes on the basis of the hydrolytic capacity [32].

Besides investigating substrate specificity of BGL, affinity gel electrophoresis was applied to identify the interaction of carbohydrate-binding modules with different polysaccharides of 0.1% concentration (oat spelt xylan, barley beta-glucan, xyloglucan and hydroxyethyl

cellulose). None of these used substrates bound to BGL, which was proven by no difference between control and sample run on the native gels (see Appendix, Figure 35 and Figure 36). Therefore to detect binding, further experiments have to be carried out, trying a range of different substrate concentrations as well as other substrates (e.g. lichenin, laminarin).

In most cases of characterized β-glucosidases from all domains of life, pNPG and cellobiose were chosen as substrates for applied assays, which made comparison with BGL difficult. Determined optimal temperatures of characterized β-glucosidases revealed higher optima on pNPG compared to pNPX [34], [49], [47]. Therefore it would be of interest to analyze, if BGL´s obtained optimal temperature of 60°C on pNPX [93] confirms the mentioned trend on pNPG. Optimal temperature results helped to select 70°C and 80°C as temperatures to understand more precisely the thermal stability of BGL. This assay showed that higher residual activity was measured for 70°C compared to 80°C. After 60 minutes of incubation, a residual activity of approximately 58% was measured for 70°C (see Figure 14) by contrast with the β-glucosidase from *Chrysosporium lucknowense*, where only 10% remained [34]. BGL still exhibited 29% of activity, whereas the fungal β-glucosidase of *C. lucknowense* was inactivated after 20 minutes at 75°C [34]. An apparent half-life of BGL with 79 minutes was determined, only between 0 to 60 minutes of incubation time at 70°C.

In order to find optimal working conditions regarding to detect high enzyme activity, the factors pH and temperature were varied at a time. BGL hydrolyzed xylobiose to xylose in a broad temperature and pH range, which was narrowed and examined applying pNPX with the same aim to determine an optimal range, where BGL shows high activity. If the analyzed factors had a significant influence on BGL´s activity, that should have been validated with statistical methodology. In addition combining these with RSM, which would have helped to build models and to study the factors interaction, could not be performed. Reasons behind this are the investigated range of pH and temperature using xylobiose as substrate, which revealed very variable results, making the application of RSM difficult. This variability was especially followed at 55°C, rather than 60°C and 65°C, where more stable activity data was measured. Comparing the second run at 55°C, it is evident that two maxima are present, at pH 4.4 and 6 (Figure 43B). Whereas the first run only showed highest activity at pH 4.4, for 55°C.

Although, temperature and pH effect on enzyme activity could not be refuted significantly by statistical evaluation and also no models could be built, nevertheless it was possible to find relationships between these parameters. Additionally, trends of measured data could be observed and optimal ranges determined. The obtained optimal range for BGL on xylobiose was used and tested on pNPX, to check the assumption if similar optimal range is detected

on other substrates. Interestingly, the narrowed range examined on pNPX revealed that BGL performs oppositional on the aryl substrate, showing higher activity on lower pH (pH 5.4). The similarity BGL shared with both substrates was the following, that with increasing temperature, the specific activity rose. Such an effect was also detected for the bifunctional β-glucosidase from yak rumen [49]. Furthermore, it also should be taken in account that BGL was incubated with xylobiose for a longer time period (up to 30 minutes) compared to pNPX (up to 3 minutes). This leads to the fact that enzyme activity is influenced by assay incubation time. In all, evaluating the data of both applied substrates, the conclusion could be drawn that BGL is more influenced by pH than temperature.

The study of the novel β-glucosidase determining higher activity, by varying pH and temperature at a time, is at its very beginning and needs further development. Since BGL showed to be relatively sensitive at 55°C as well as pH 3.5 (at this pH, activity determination for all measured temperature points, was not possible due to instability of measured product amount), using xylobiose as substrate, it should be considered to change the range of investigated parameters. It may be possible that BGL does not prefer to work at temperatures, which are lower than its optimum temperature, showing more rigidity. It is known, that enzymes from thermophiles improve the property of being rigid at lower temperature with moving towards the temperature, where the enzyme shows maximal activity [9]. Therefore it might be an idea to investigate the enzyme´s behavior at higher temperatures, for example 70°C. At 70°C BGL revealed to be very thermostable, because after 3 hours of incubation still approximately 80% of residual activity was left.

3.2.3 Conclusions

This work provides information about the properties of the novel β-glucosidase from the marine organism *Rhodothermus marinus*, which has been classified to family 3 of glycoside hydrolases. The various carried out characterization procedures, from molecular weight determination to catalytic performance, enzyme stability and much more, led to meaningful results. These are useful in defining specific application areas of analyzed enzyme.

The purification step revealed to be problematic, which leads to suggest non-chemical purification procedures. Based on the obtained results, β-glucosidase proved to be a very interesting enzyme. It showed broad substrate specificity and revealed the indication of being a bifunctional enzyme (β-glucosidase/β-xylosidase activity). This bifunctional character is of great interest, especially for industrial purposes. Therefore, it would be of interest to continue research in this direction. Besides detecting bifunctional property, the approximately 87 kDa marine β-glucosidase exhibited thermostability quality (up to 80°C). Since, molecular weight difference was obtained between the computed and the determined mass with LC-MS, it would be of interest to concentrate on N-terminal sequencing to explain the variation.

β-glucosidase revealed higher enzyme activity at pH 5.4 and 65°C with pNPX and with xylobiose a shift of optimum was obtained with temperature and variability detected between each run. The investigation of optimizing enzymatic activity through varying temperature and pH at a time was a new approach, to understand enzyme stability as well as the parameters influence and relationships. Such a methodology is performed very seldom during characterization of enzymes, although it has various benefits. This carried out study is in its beginning stages and forms the basis for continuing experimental analysis. Besides reducing experiments, the results of applied study can also help in further secondary characterization steps (e.g. transglycosylation assays).

3.3 Part II - NAG

Results of β-N-acetylglucosaminidase (Rmar_0925, NAG) are indicated in the following. These are firstly divided in an expression and purification part, which is followed by a part of detailed characterization of the protein. Afterwards the obtained experimental results are discussed and out of this conclusions made.

3.3.1 Results

3.3.1.1 Expression and purification of recombinant protein

The growth curves of the produced protein of interest are represented in Figure 18, where the measured OD_{620} values from the 1 ml collected samples are plotted against time. Following the course of the curves, fermentation A shows less biomass production compared to fermentation B, but both act similar in the first 3 to 4 hours of process time. Whereas fermentation A and C overlap partly during 4.5 to 5.5 hours. The steady state of fermentation C is reached after 9-10 hours and afterwards the biomass production decreases slowly. For the other two performed fermentations the production of biomass increases slowly after the exponential phase.

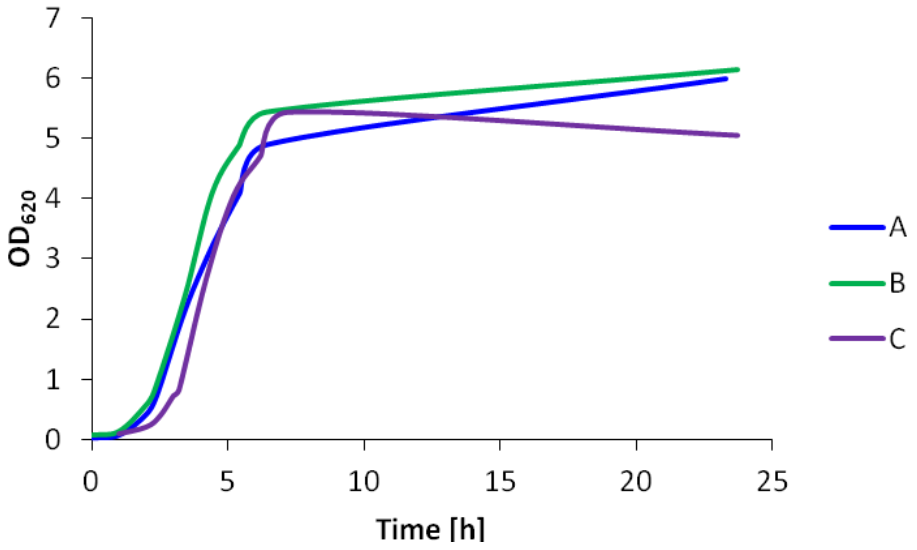

Figure 18: Recombinant protein production and expression of β-N-acetylglucosaminidase, showing three different fermentations. The shake flask cultivation was carried out at 37°C and biomass production was measured at OD_{620}. Process time lasted for fermentation A approximately 23 hours, whereas B and C were run for around 24 hours.

For verification that recombinant protein expression took place, samples from the fermentation process, which were divided in their soluble and insoluble form were loaded on SDS-gels. An example for SDS-gel quantification is given in Figure 19. The information of

theoretical weight of the protein from ExPasy tool [84] (105.17 kDa) helped to detect a band of interest approximately at 105 kDa.

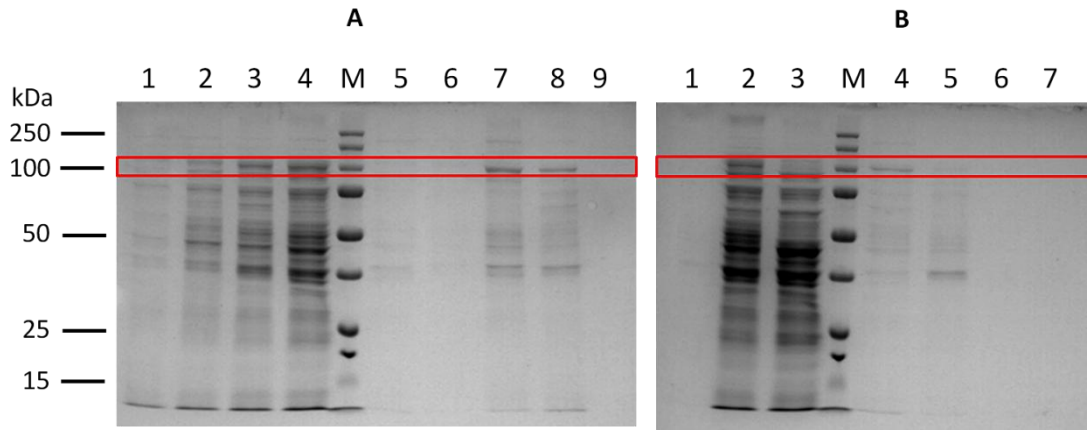

Figure 19: SDS-PAGE gels showing samples collected during recombinant protein production of β-N-acetylglucosaminidase to check expression level. In both gels (A, B) the lane M indicates the 1kDa protein ladder. Gel A shows from lane 1 to 4 every hour sample (0-3 hours) in soluble form (containing protein) and lanes 5-8 the same but for insoluble form. Whereas in lane 9 the supernatant sample of fourth hour sample is given. Gel B reveals in lane 1 again the same sample as in gel A (lane 10). In lane 2 (soluble sample) and 5 (insoluble sample) the sample after 4 hours of induction is loaded. Lane 3 (soluble sample), 6 (insoluble sample) and 7 (supernatant) indicate the overnight sample. The collected samples show at approximately 105 kDa a band, which most probably corresponds to the searched protein of interest.

The applied method of IMAC provided a very easy purification procedure, where a very sharp peak was fractioned and contained the protein of interest (see appendix, Figure 40). The successfulness of the purification step was analyzed by SDS-gels, where a band similar to the gel of the expression level was detected (see Figure 20) at around 105 kDa. The purification gels show that not only a single band of protein was obtained, but also others. It was decided that for further analysis procedures this level of purification is sufficient. SDS-gels of purified protein from other batches are given in the appendix (Figure 37 and Figure 38).

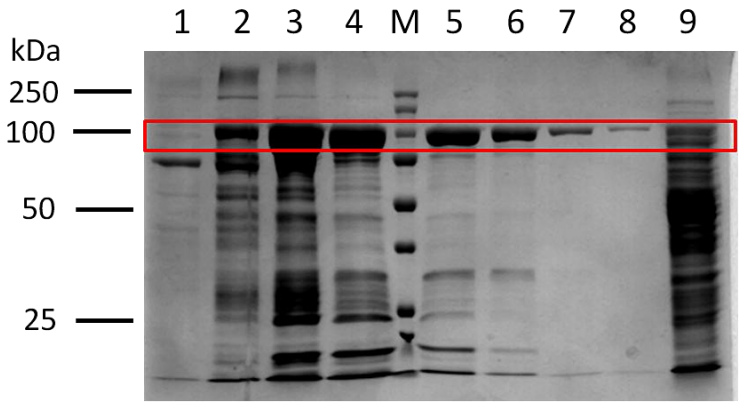

Figure 20: SDS-PAGE gel for evaluation of purification level. The collected fractions from IMAC with high protein concentration were loaded on this gel (lane 1-4 and 5-8). In lane M, the corresponding 1 kDa protein ladder can be found and lane 9 shows the flow-through sample. The samples in lane 2-4 contain together with the protein of interest at around 105 kDa additional bands than compared to the samples in lane 5-8. The latter mentioned lanes reveal samples with less purified protein amount NAG, in relation to lane 3 and 4. These are overloaded with protein sample and therefore overlap with other protein bands, resulting in to indicate higher concentration of protein.

In addition it was tried to remove the high amount of imidazole. SDS-gels were prepared to check if the purified protein was preserved (Figure 37C, appendix). The processed sample is given in lane 5, which shows a very big band at the same position, where the purified protein was detected in Figure 37A, B. After one week of imidazole separation, the protein started to precipitate and this precipitation became more with time. Therefore it was decided to work with the content of imidazole and that the purity of protein was good enough for further characterization.

3.3.1.2 Characterization

3.3.1.2.1 Determination of molecular weight with LC-MS

The relatively big protein β-N-acetylglucosaminidase has a computed theoretical weight of 105.17 kDa [84], which was detected on the SDS-gel at approximately 105 kDa. For the molecular weight determination with LC-MS a protein concentration of 109 µg/ml were applied on the HPLC column and separated. The eluted peak at a retention time of 10.39 to 11 minutes (Figure 21A) was further analyzed in the m/z spectrum (Figure 21B). A molecular weight of 104.52 kDa was obtained from the analysis, which is less than compared to the theoretical weight with a difference of around 0.65 kDa. The traceability of a pattern in Figure 21B is difficult to follow, because it is overlapped with noise.

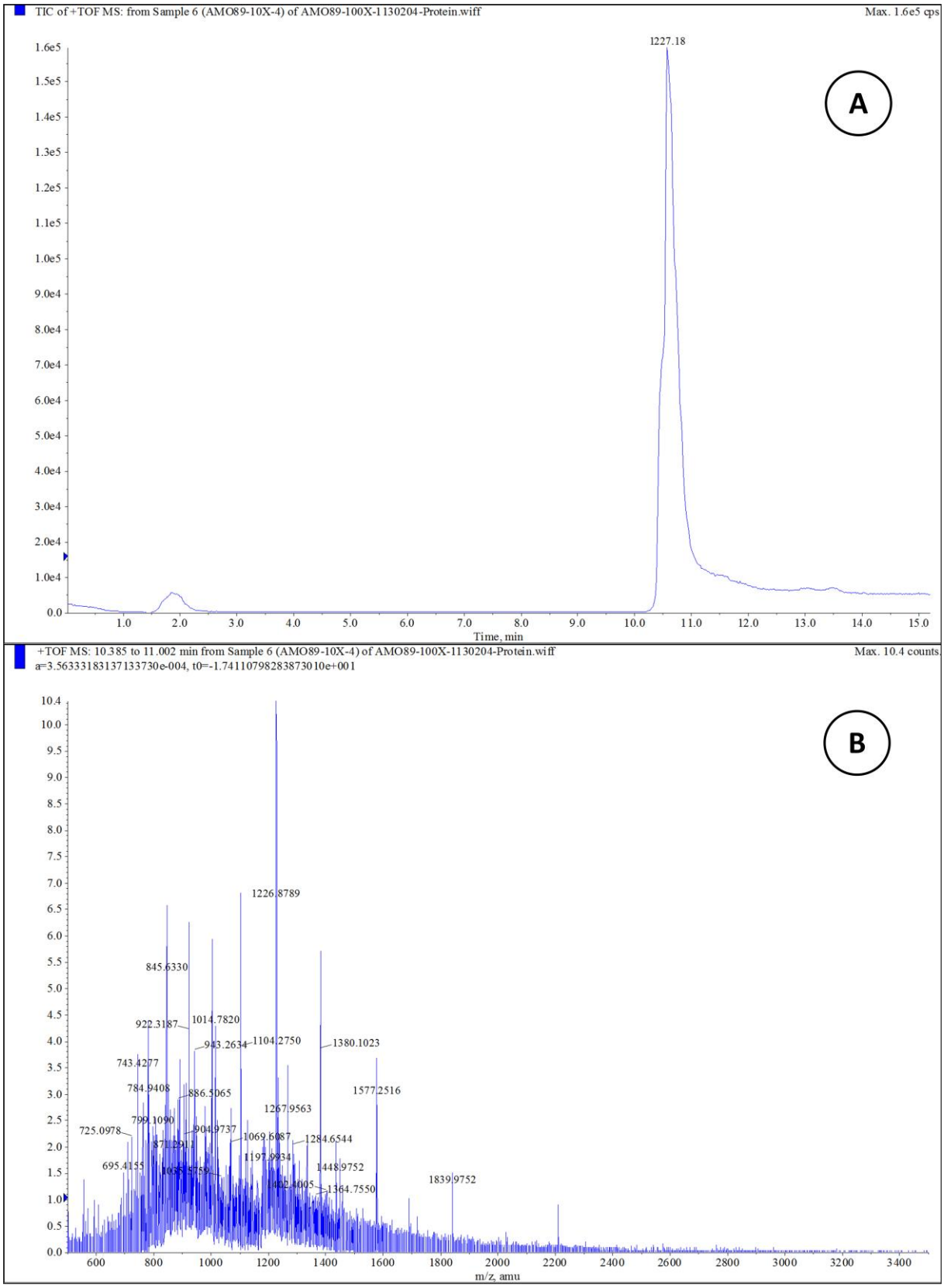

Figure 21: Exact molecular weight determination of β-N-acetylglucosaminidase. The protein of interest eluted between 10.39 to 11 minutes, which provides the LC chromatogram in figure A. A molecular mass of 104.52 kDa was determined for β-N-acetylglucosaminidase through investigation of the corresponding m/z spectra, given in figure B**.**

3.3.1.2.2 Substrate specificity

β-N-acetylglucosaminidase shows very particular specificity towards applied substrates. A wide range of synthetic substrates were screened for activity (Table 8). Color development could only be monitored with pNPGlcNAc, for all other substrates no activity was obtained.

Table 8: Screening of β-N-acetylglucosaminidase on various aryl substrates

Substrate	Activity
p-Nitrophenyl-β-D-glucuronide (pNPG)	N.D.
p-Nitrophenyl-β-D-xylopyranoside (pNPX)	N.D.
p-Nitrophenyl-β-L-arabinopyranoside (pNPA)	N.D.
p-Nitrophenyl-β-D-cellobioside (pNPC)	N.D.
p-Nitrophenyl-α-D-mannopyranoside (pNPM)	N.D.
p-Nitrophenyl-β-D-galactopyranoside (pNPGal)	N.D.
p-Nitrophenyl-α-L-rhamnopyranoside (pNPR)	N.D.
p-Nitrophenyl-N-acetyl-β-D-glucosaminide (pNPGlcNAc)	+++

Additionally to the aryl substrates, naturally based substrates the chito-sugars were also tested towards activity. β-N-acetylglucosaminidase showed to be highly active on all applied substrates and the results of the screening are given in Table 9.

Table 9: Screening of β-N-acetylglucosaminidase on various chitooligosaccharides

Substrate	Activity
di-N-Acetyl-Chitobiose	+++
tri-N-Acetyl-Chitotriose	+++
tetra-N-Acetyl-Chitotetraose	+++
penta-N-Aceytl-Chitopentaose	+++

An example chromatogram is given in Figure 22, where the hydrolysis products of the above mentioned substrates are shown. Surprisingly the larger chitooligos eluted earlier than the smaller sugars ($CH_5 \rightarrow CH_2$). In this case, a very fresh produced and purified protein sample was used for substrate specificity screening and interestingly no baseline disturbance appeared during separation. But, later on when carrying out further analysis applying same protein sample, baseline disturbance due to presence of imidazole was detected. From these 4 applied oligosaccharides, chitobiose was chosen as substrate for further characterization of β-N-acetylglucosaminidase, which hydrolyzed it to **N**-**ac**etyl**glu**cosamine (GlcNAc) monomers.

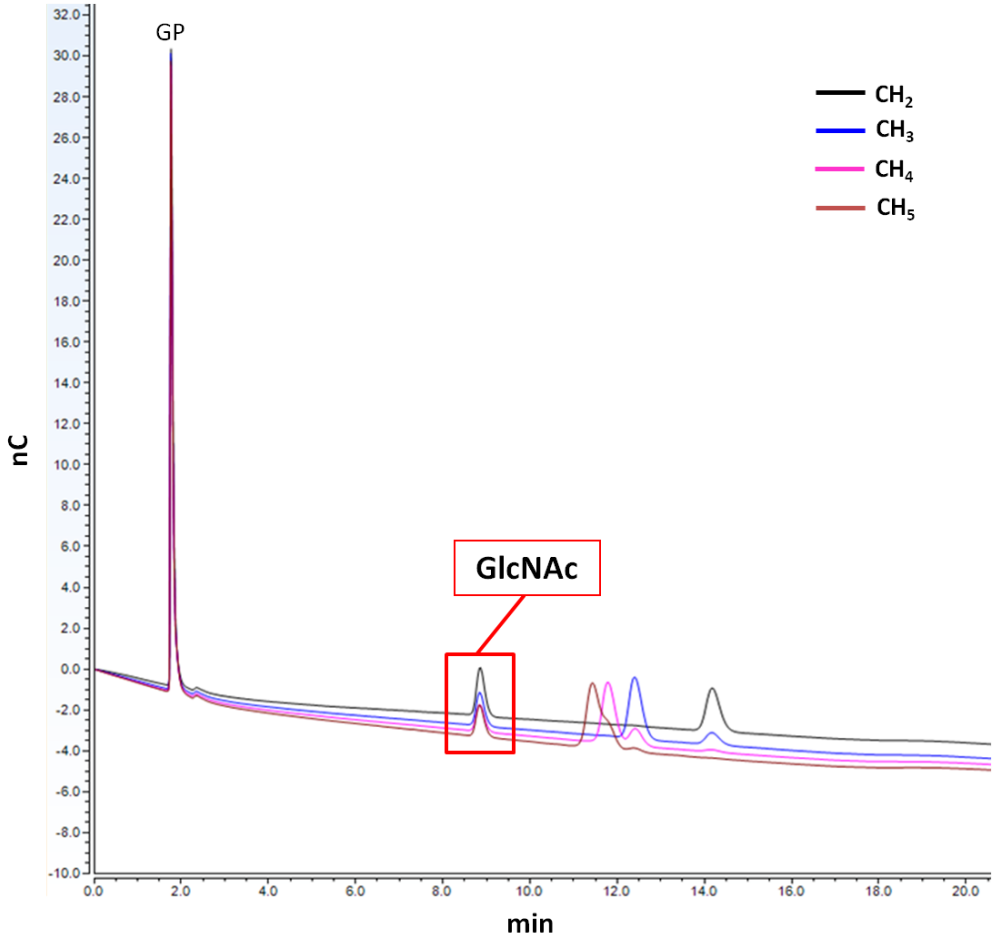

Figure 22: Hydrolysis profile produced by β-N-acetylglucosaminidase separating chitooligosaccharides. A ghost peak was measured in the beginning of the applied chromatography method. β-N-acetylglucosaminidase hydrolyzed four different chitooligos, which eluted in reverse direction (first larger sugars). Chitopentose (CH_5) was formed into N-acetylglucosamine (GlcNAc) and showed little amount of chitotetraose (CH_4) and chitotriose (CH_3), the corresponding peaks of CH_5 are colored in dark red. Whereas chitotetraose produced GlcNAc, plus CH_3 and small amount of CH_2, highlighted in pink. Chitotriose was hydrolyzed to the GlcNAc monomer and CH_2, given in bluish color. Finally, β-N-acetylglucosaminidase hydrolyzed chitobiose to GlcNAc (black line).

3.3.1.2.3 Thermal stability

Applying the continuous assay, β-N-acetylglucosaminidase was examined at a temperature range of 50°C to 90°C, where the initial velocity and the finally resulting specific activity was calculated between a time period of 0 to 140 seconds. In Figure 23, the obtained specific activity using pNPGlcNAc is plotted against temperature, allowing to determine the temperature optimum. A trend has become clear in Figure 23, where the specific activity increased with temperature. It seems that even high activity would be measurable above 90°C.

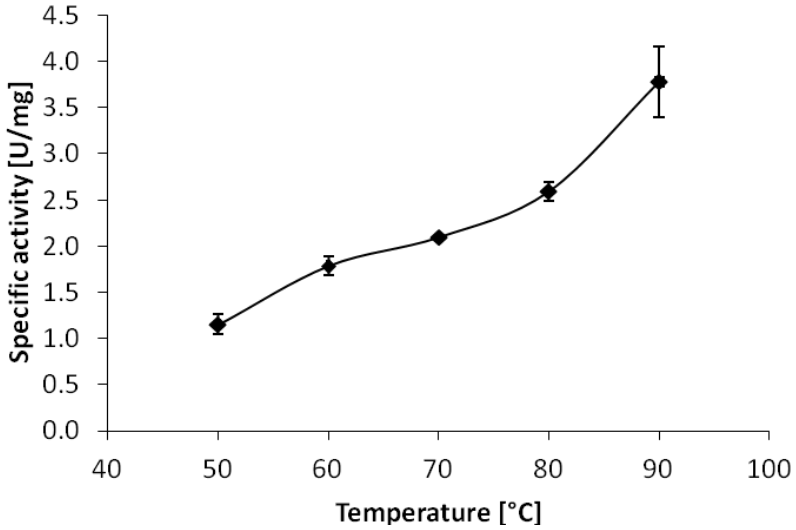

Figure 23: Determination of β-N-acetylglucosaminidase temperature optimum between 50°C to 90°C, where reactions were followed continuously. Measured specific activity values are plotted against the applied temperature range. Highest activity was monitored at 90°C, where also the highest standard deviation was calculated.

The analysis of the temperature profile of β-N-acetylglucosaminidase allowed to generate an Arrhenius plot (Figure 24,), which led to determine the activation energy between the above mentioned temperature range. The outcome of plotting the natural logarithm of the reaction rate against the reciprocal value of temperature was an activation energy of 26.86 ± 0.23 kJ/mol.

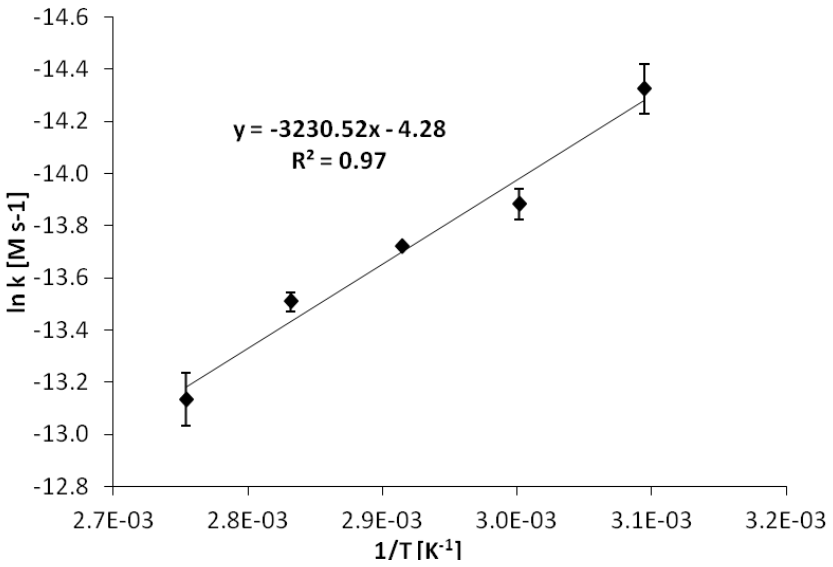

Figure 24: Arrhenius plot for determination of β-N-acetylglucosaminidase activation energy between 50°C to 90°C. The linear regression of the obtained values from the natural logarithm of initial reaction rates reveals a regression coefficient R^2 of 0.97. This plotting of the temperature optima data led to provide an activation energy for β-N-acetylglucosaminidase of 26.86 kJ/mol.

Investigation of thermal stability was performed at 80°C and 90°C, incubating the enzyme for 3 hours. The obtained results of thermal inactivation are represented in Figure 25, where after 30 minutes of incubation at 80°C, the residual activity remained the same. Later on, the activity declined slowly with time and increased from 44% residual activity to 50% at the last measured time point of 180 minutes. Comparing the thermal inactivation data of both applied temperatures, it can be noticed that already from 30 minutes of incubation no activation was measured at 90°C. Therefore the apparent half-life was only calculated for 80°C between 0 to 120 minutes and resulted in 96.27 minutes for β-N-acetylglucosaminidase (Figure 26).

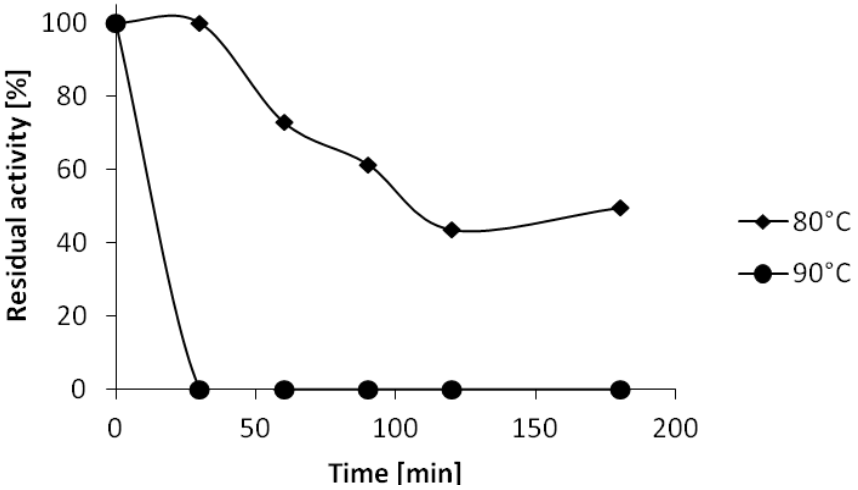

Figure 25: Thermal stability curves of β-N-acetylglucosaminidase, where residual activity of the enzyme is plotted against the time of incubation (0-180 minutes). Stability investigated at 80°C showed 50% of residual activity left, whereas at 90°C β-N-acetylglucosaminidase was inactivated totally already after 30 minutes.

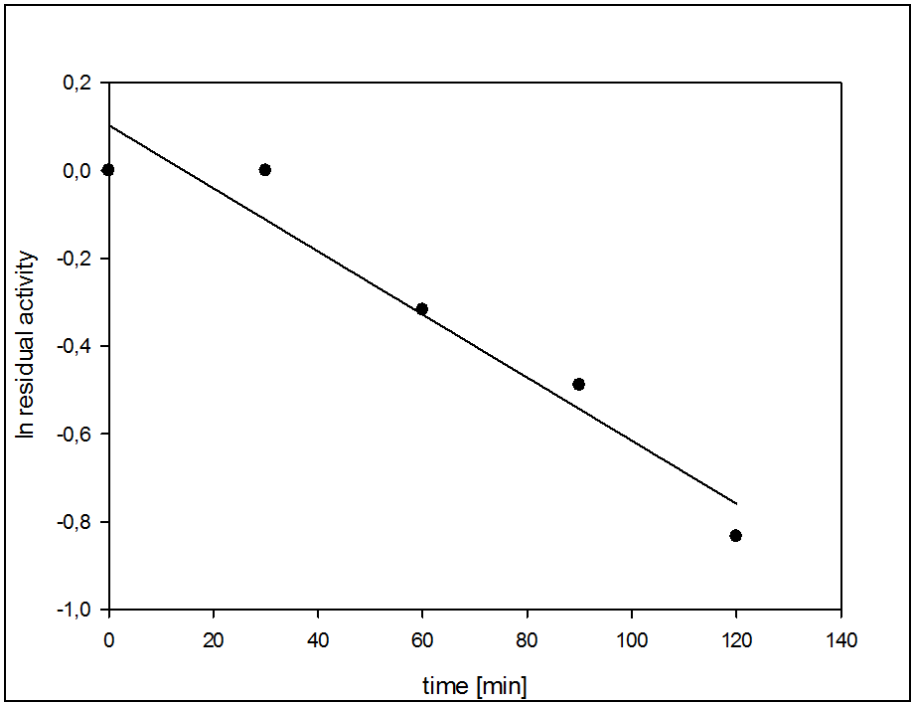

Figure 26: Determination of half-life to the corresponding thermal stability analysis of β-N-acetylglucosaminidase between 0 to 120 minutes of incubation time. An apparent half-life of 96.27 minutes was obtained by plotting the natural logarithm of the measured residual activity against time.

3.3.1.2.4 Determination of kinetic parameters

The application of Lineweaver-Burk relation led to the determination of K_m, V_{max} and k_{cat} values. These results were also compared and fitted by applying Michaelis-Menten model (95% confidence interval, SigmaPlot), which are represented through the saturation curve of β-N-acetylglucosaminidase (Figure 27). Below that figure, the Lineweaver-Burk plot is given (see Figure 28), where a K_m of 0.07 mM and a V_{max} of 6.15 U/mg was determined between a substrate concentration of 0.05 mM to 0.4 mM. From this data a k_{cat} value of 10.76 s^{-1} was calculated. The applied tool provided very similar K_m and V_{max} values of 0.06 mM and 5.96 U/mg, resulting in a k_{cat} of 10.42 s^{-1} and a R^2 of 0.95 for the model fit. All compared values of the kinetic study showed a very good accordance. The only bigger difference was between the k_{cat}/K_m relation, where the calculated data was somewhat lower (162.30 s^{-1}/mM) than the results from the applied tool (174.32 s^{-1}/mM).

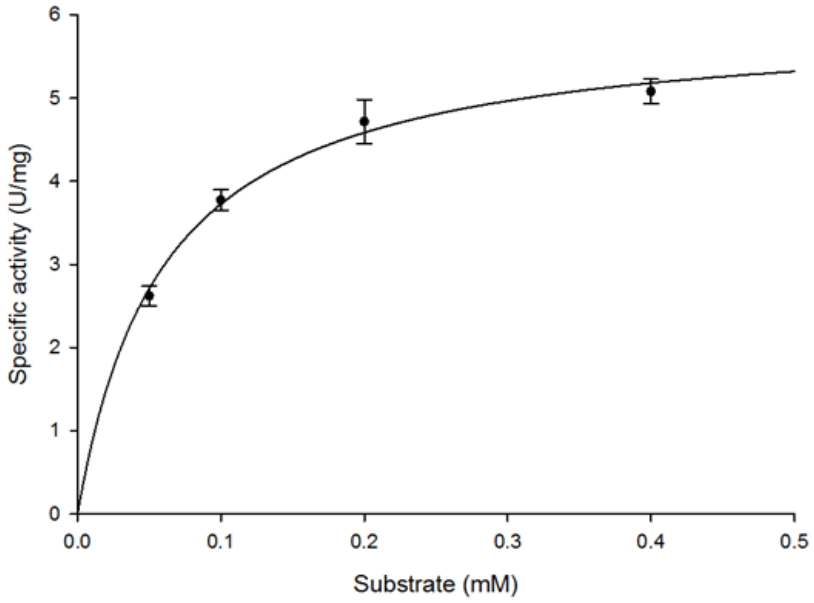

Figure 27: Saturation curve obtained by β-N-acetylglucosaminidase, which follows Michaelis-Menten kinetics. Specific activity is plotted against the investigated substrate concentration range (0.05 to 0.4 mM) and the graph reveals the fitted data with the sigma plot tool, R^2 of 0.95.

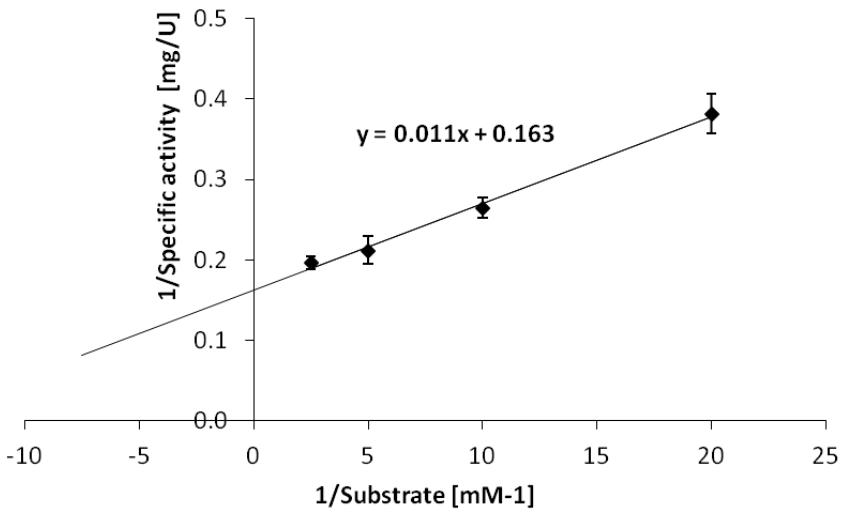

Figure 28: Determination of kinetic parameters for β-N-acetylglucosaminidase by using Lineweaver-Burk plotting at 70°C and pH 5.6. The reciprocal of the specific activity against the reciprocal of substrate concentration between 0.05 to 0.4 mM yielded in the provided kinetic parameters.

3.3.1.2.5 Detection of enzyme binding to polymeric substrates - Affinity gel electrophoresis

All native-PAGE gels which were run to analyze specific substrate binding (four different substrates) with β-N-acetylglucosaminidase are presented in the appendix, Figure 35 and Figure 36. This enzyme sample did not even run on the substrate bond gels. Therefore no interaction with substrate was proven.

3.3.1.2.6 Effect of pH and temperature on enzyme activity - optimization study

The optimization of enzyme activity by varying temperature and pH at a time was initially investigated on pNPGlcNAc, at which NAG revealed high substrate specificity. The applied aryl substrate was used to screen NAG´s activity in a broad range. To follow the performance of NAG, the obtained specific activity is plotted against the used pH-values in relation to three different temperatures (see Figure 29). Results of the assay show that, a maximum of activity was detected at around pH 5.4 for all applied temperatures, which rises with increasing temperature. The obtained screening data was sufficient to continue further analysis of NAG´s optimal range on the natural substrate.

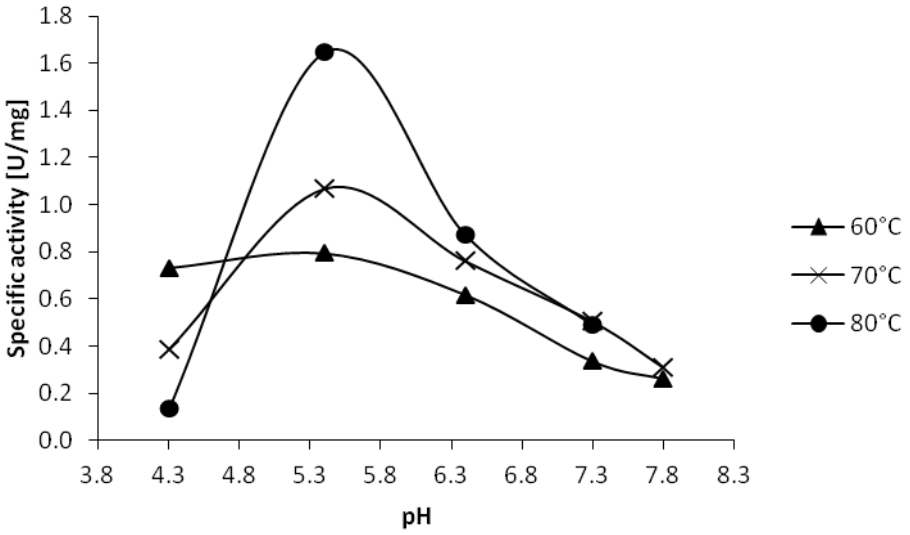

Figure 29: The plot represents the results of the screening for optimal range determination of β-N-acetylglucosaminidase by using pNPGlcNAc as substrate. Obtained specific activity is plotted against the investigated pH-range 4.3 to 7.8 and is in relation to the applied temperatures (60°C, 70°C and 80°C). Optimum detected at pH 5.4, with increasing temperature.

The screening results helped to design the experimental space for more detailed investigation of NAG´s optimal range, by applying another substrate (chitobiose). It was also tried to analyze, if NAG acts similar on the used natural substrate compared to pNPGlcNAc. Therefore, the temperature and pH range was narrowed and a 2x3 factorial design was performed. NAG hydrolyzed chitobiose to N-acetylglucosamine monomer and this hydrolysis product peak was located in the baseline disturbance, which occurred based on the presence of imidazole in the samples. Due to less product formation, the integration of the peak area of the hydrolysis product proved to be difficult. Since, the specific activity should have been determined by the formed amount of product and applied as the response for further analysis with statistical methodology, instead that the relative peak area [%] of chitobiose after 40 minutes incubation time was used as response.

Evaluation of the model revealed that the linear negative and quadratic positive effect of pH as well the linear positive temperature term are significant factors (95% confidence interval). Whereas the quadratic temperature effect and the interaction of the applied parameters, provided not to be a significant model term. These variables are given in the scaled and centered coefficient plot in Figure 30. Furthermore, for validation a basic model statistics is represented in Figure 31, where a regression coefficient R^2 of 0.93, a future prediction precision Q^2 of 0.85, a model validity of 0.66 and reproducibility of 0.91 was obtained. A contour plot (RSM) is shown in Figure 32, where the factors pH and temperature are related to the response (relative area). The plot illustrates that chitobiose was not consumed a lot by β-N-acetylglucosaminidase (monitored 97.5 to 100 % relative peak area) and the least

product formation was measured with high temperature and low pH. In addition with rising pH value, the consumption of chitobiose increases, independent from temperature. Therefore an optimum was determined at pH 6.2 at all applied temperatures.

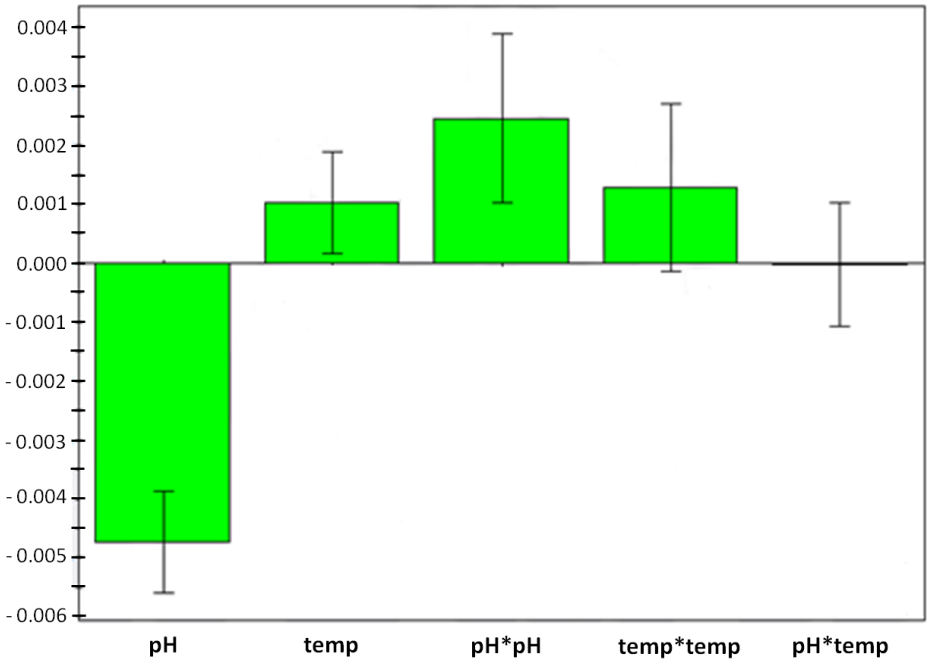

Figure 30: Representing scaled and centered coefficients for the relative peak area of chitobiose (response). Temperature and the interaction term pH have a significant positive effect on the response, pH term exhibits a negative effect. The interaction term of pH*temperature and temp*temp revealed no significance.

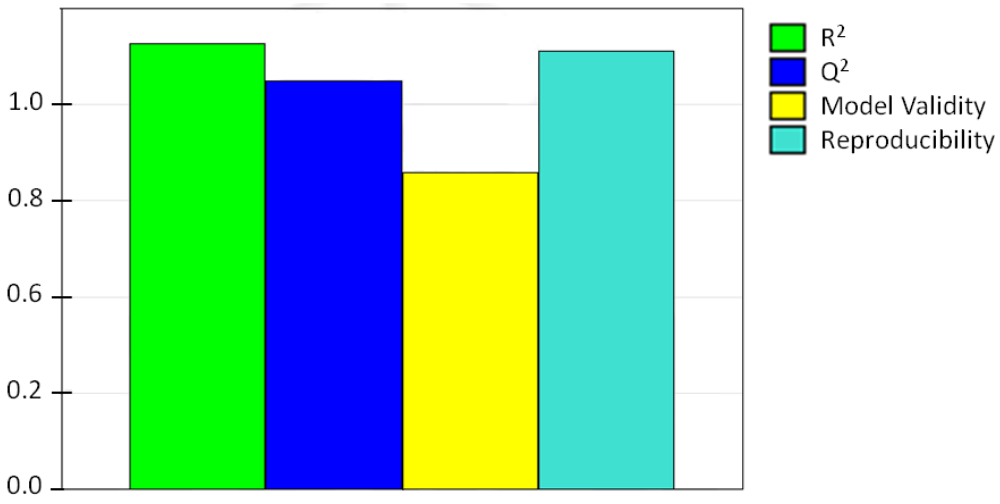

Figure 31: Basis model statistics for the modeled relative peak area, showing the regression coefficient R^2, the prediction precision Q^2, the model validity as well the reproducibility of the regression model.

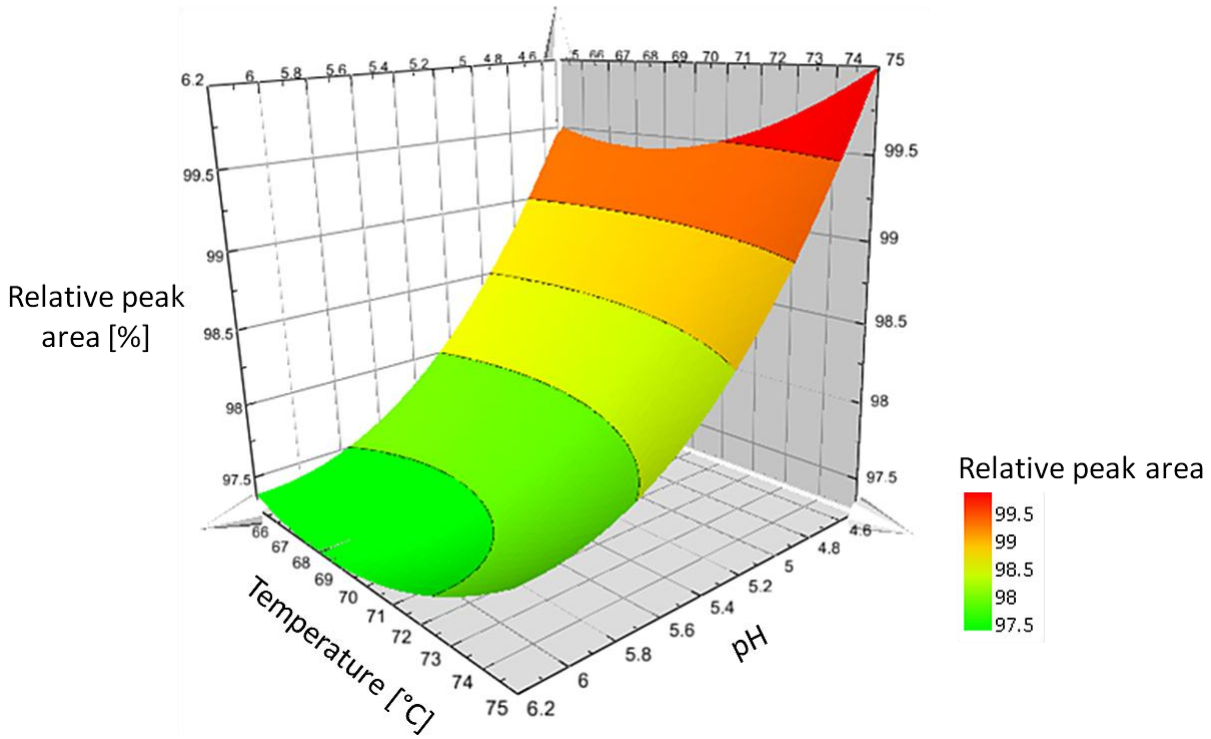

Figure 32: Contour plot obtained for β-N-acetylglucosaminidase. This plot shows the results of the modeled data between a temperature range from 65°C to 75°C and pH 4.5 to 6.2 in relation to the relative peak area of chitobiose.

The results of applying two different substrates showed that NAG acted totally different in regard of optimum values. Enzyme activity rose with increasing temperature at optimal pH, when pNPGlcNAc was used. Whereas on chitobiose, the opposite was detected.

3.3.2 Discussion

β-N-acetylglucosaminidase from the thermohalophilic bacterium *Rhodothermus marinus,* belonging to GH3 family was successfully expressed (*E. coli*) and purified. Few characterized bacterial β-N-acetylglucosaminidases are known, which exhibit large molecular masses in comparison to fungal origin [31]. One of them is the characterized β-N-acetylglucosaminidase of this thesis, with a weight of 104.52 kDa determined by LC-MS, which is in accordance with the obtained 105 kDa from SDS-gel analysis as well as the computed molecular mass of 105.17 kDa. Almost similar molecular weight related to NAG reveals an *Areomonas sp*. with 103 kDa [104], but also the fungal β-N-acetylglucosaminidase from *Paecilomyces persicinus* shows a mass of 100 kDa [105]. A bit smaller in size, but ranked to the enzyme group with large weight is the β-N-acetylglucosaminidase from marine bacteria *Vibrio harveyi* (98 kDa) [106].

The enzyme characterized is very specific towards substrate hydrolysis. Besides being active on one aryl substrate pNPGlcNAc, from the various tested, NAG also hydrolyzed N-acetylglucosamine (GlcNAc) monomers from chitooligosaccharides (CH_n) at the non-reducing end. Detected substrate specificity of NAG is very similar to that of the *S. thermoviolaceus* β-N-acetylglucosaminidase [57], with the only difference that pNP-β-N-acetylgalactosamine was not tested for NAG. High substrate specificity towards pNPGlcNAc and/or chitooligosaccharides was determined for various other β-N-acetylglucosaminidases such as from an *Altermonas sp.* [92]*, Nocardia orientalis* [107], rice seeds [108], *Thermotoga maritima* and *Thermotoga neapolitana* [109]. The substrate specificity results of NAG correspond well with the fact that β-N-acetylglucosaminidases are required to hydrolyze GlcNAc from $GlcNAc_2$, which is often the formed end product by chitinases during chitin degradation, a common raw material in the marine environment [109].

Furthermore the applied HPAEC-PAD method, which helped to detect natural substrate activity led to an interesting phenomena, based on separating chitooligosaccharides. Surprisingly, chitooligosaccharides with higher linked GlcNAc´s eluted earlier than smaller chito-sugars ($CH_5 \rightarrow CH_4 \rightarrow CH_3 \rightarrow CH_2$) (see Figure 22). Although methods were varied, that was the best achievable performance for chitooligosaccharides separation, due to possible options of applicable columns. Chitobiose was only separated properly, without overlapping with other chito-products.

In addition, the influence of imidazole was observable with HPAEC-PAD rather than with the other carried out assays for enzyme characterization. Freshly purified protein did not exhibit

an imidazole peak during separation (see Figure 22), but with time baseline disturbance in presence of imidazole could be monitored, which made integration of product peaks difficult. Therefore the separation of imidazole was necessary and resulted in a very complicated procedure. The process of imidazole separation proved to be unsuccessful, because the protein started to precipitate with time and was only stable for approximately one week. Various reasons may explain the protein precipitation. The centrifugation step may led to concentration of the protein on the membrane, although the volume of protein remained same. Some proteins do not withstand concentrating. Another factor for protein precipitation which plays a role together with the centrifugation step, is the changing in buffer system. Additionally, the purified enzyme containing a histidine tail may be stabilized with the content of imidazole. Therefore, other suggestions for further alternatives to remove imidazole content after purification step are needed, which are discussed in the outlook (section 6).

The determination of temperature optimum of the enzyme-catalyzed reaction revealed that NAG has thermotolerant properties. It showed tendency of exhibiting optimum temperature above 90°C and an activation energy of 26.86 kJ/mol. However, this determination is contrary to the results obtained from thermostability data, where NAG was totally inactivated after 30 minutes at 90°C. Similar characteristics was obtained for the β-N-acetylglucosaminidase of *Nocardia orientalis*, where optimum temperature was determined at 70-75°C, but inactivation of enzyme took already place at 65°C [107]. Hence, it is often difficult to evaluate the temperature optimum, because it varies with the followed reaction time [89]. Various β-N-acetylglucosaminidases, which were analyzed regarding their temperature optima at relatively similar conditions compared to NAG, showed results with high temperature optima. Such as the β-N-acetylglucosaminidase from the soil bacteria *S. thermoviolaceus* at 60°C [57] or the thermophile bacteria *B. stearothermophilus* at 75°C [58]. NAG is comparatively thermostable at 80°C ($t_{1/2}$ = 96.27 min) with almost 50% residual activity after 180 minutes of incubation, since the β-N-acetylglucosaminidases of *T.harziarum* [110] was already inactive after 30 minutes and N-acetylglucosaminidase of the marine zoanthid *Palythoa caribaeorum* showed almost no activity after 1 hour [111].

Kinetic parameters were determined at 70°C and near the optimum pH-value on pNPGlcNAc (pH 5.6). Selection of these specific parameters was based on the following considerations. The fact that thermostable enzymes are used for several industrial applications due to their various beneficial aspects, it should be taken in account that these applications are accompanied by some limitations. Working at high temperatures means occurrence of costs

for cooling or materials, which withstand these temperatures. Therefore, it was decided to work at a more moderate temperature.

When comparing K_m of different β-N-acetylglucosaminidases, variations can take place, which can basically be substrate specific [111]. β-N-acetylglucosaminidases were characterized from two different thermophilic bacterial organisms *Thermotoga neapolitana* (nagA) and *Thermotoga maritim*a (cbsA), regarding kinetic parameters [109]. The catalytic efficiency (k_{cat}/K_m) of nagA and cbsA were 5.4-fold, respectively 4.7-fold higher compared to *R. marinus* NAG, which leads to the conclusion that nagA and cbsA share higher substrate specificity than NAG. Hence, it should be taken in account that a low K_m points out that only small concentration of substrate are necessary, to get the enzyme to operate at half of its maximum velocity. This results in, that the enzyme has high affinity towards the applied substrate. NAG with a K_m of 0.07 mM is relative lower than nagA and cbsA with 0.28 mM and 0.36 mM. Furthermore, a low V_{max} reveals that the enzyme is not able to catalyze many reactions in a short period of time, which is the case of the two β-N-acetylglucosaminidases from the *Thermotoga* organisms [109].

The performance of affinity gel electrophoresis, detecting interaction of polymeric substrates with the enzyme NAG was not possible, since NAG did not migrate at all on the native gels. Reasons behind this outcome most probably lie in the difference of pI between the protein and the applied buffer system. A significant difference of 1 pH unit of the protein below the buffer used is needed for efficient migration. Therefore other options have to be considered for successful run of NAG protein on the gels. Such as for example, using another buffer system or changing the position of cathode with anode.

Mostly, whenever β-N-acetylglucosaminidases are investigated based on temperature and pH parameters, the evaluation of these occurs with one parameter at a time [53], [109], [112]. Such investigations can be time consuming and do not allow access of the interaction effects between the applied parameters. In addition, interaction of these parameters may lead to a synergistic effect, which signifies a state where pH and temperature cause higher enzyme activity [113]. Therefore the influence of pH and temperature on the thermostable NAG was analyzed by using two different substrates (applying DoE). The concept behind the screening of NAG on pNPGlcNAc was to determine optimal conditions on the basis of pH and temperature using a broader range, which can be then narrowed when measuring optimum activity conditions on natural substrates. The latter mentioned group of substrates are preferably used in industrial area compared to aryl substrates. Due to this fact, temperature

and pH effect was more precisely analyzed on chitobiose by applying statistical methodology.

NAG revealed a very specific optimum range on pNPGlcNAc. The specific activity was highest at pH 5.4 and increased with temperature (see Figure 29). Although the screening of NAG on pNPGlcNAc investigated both parameters at a time, β-N-acetylglucosaminidases from other organisms can be compared at similar conditions. The β-N-acetylglucosaminidase of soil bacteria *S. thermovialaceus* showed at 60°C an optimum pH of 5, almost similar to NAG [57]. Whereas β-N-acetylglucosaminidases of thermophile bacteria *B. stearothermophilus* revealed high activity at pH 6.5 and 70°C [58]. Higher pH optima compared to NAG were obtained from β-N-acetylglucosaminidases of *T. neapolitana* and *T. maritima* at pH 7-8 at 70°C [109]. The screening on pNPGlcNAc revealed that NAG was more influenced by temperature than pH.

The determined optimal range of pH and temperature for NAG on pNPGlcNAc was narrowed and examined on chitobiose. Less product amount of hydrolyzed GlcNAc and imidazole content made integration of product peak difficult. Hence, the remained amount of applied reactant (relative peak area) was used as response. A relation could be defined, revealing that the less the relative peak area of chitobiose was, the more product was formed and the higher specific activity for NAG was obtained. One consequence of less product formation may be due to high enzyme dilution.

The investigation of the scaled and centered coefficients exhibited valuable results, describing the influence of the analyzed parameters on the response. A significant influence on the consumption of chitobiose had the quadratic pH and the linear pH as well temperature effect. All other analyzed effects, the quadratic temperature term and the interaction term of pH and temperature, did not significantly influence the response. Additionally, from the obtained significant factors, only the linear pH term showed negative effect. The obtained negative effect of this factor means that an increase of it will reduce the amount of reactant. This would lead to form more hydrolysis product, and therefore increase enzyme activity. Although the basic model statistic parameters (R^2, Q^2, model validity and reproducibility) revealed relatively good values (see Figure 31), the results of the model have to be investigated carefully and decided if it is necessary to continue experimenting, which is done by analyzing the contour plot. The least formation of GlcNAc was performed at low pH and high temperature, which increased with higher pH and lower temperature.

Furthermore, the obtained values of the consumed amount of chitobiose differentiate very slightly (see Figure 32). Therefore, it would be preferable to broaden the range through investigating the upper area and positioning new experiments to examine the response, since hydrolysis of chitobiose was less at acid pH range, independent from temperature. Moreover, substrate concentration does also influence the enzyme activity and should be taken in account in this investigation. DoE makes it possible to analyze several factors at a time. These are exercises, which has not been done at this stage of process development.

Analyzing the results of optimal range determination, NAG showed to act different on the applied substrates. Interestingly, on pNPGlcNAc at the optimum pH 5.4, the activity increased with temperature. Whereas on chitobiose, the formation of product was monitored with higher pH (pH-value 6.2) and decreasing temperature. There may exist various explanations for this result, and one of these might be the different time periods with that the enzyme was incubated on the two applied substrates. Even though, by considering the assay time, if still such a contrariness occurs, it would be better to carry out such a study independently on each substrate. This perhaps means the need of more experiments, but such a methodology allows to be more flexible during process development, helping to find optimal conditions as well as allowing to understand the interactions between the applied factors.

3.3.3 Conclusions

In this thesis, a novel member of the glycoside hydrolase family 3 was successfully expressed and purified. It is a β-N-acetylglucosaminidase enzyme from a thermohalophilic bacteria, which was characterized regarding various properties. Additionally, the enzyme was studied concerning optimal working conditions.

The results of the different performed enzyme characterization assays, revealed to be valuable. The relatively heavy β-N-acetylglucosaminidase (approx. 105 kDa) demonstrated to be very specific in regard of its substrate specificity, hydrolyzing GlcNAc monomers (from pNPGlcNAc and CH_2 to CH_5). Furthermore, based on the applied temperature assays, it could be evaluated that determination of the temperature optimum varies with the followed reaction time. Since, β-N-acetylglucosaminidase showed even a higher optimum temperature than 90°C, which was carried out for a shorter time period, but the thermostability data exhibited that β-N-acetylglucosaminidase was totally inactivated after 30 minutes at that specific temperature. These assays helped to understand the enzyme stability. In spite of that, β-N-acetylglucosaminidase revealed thermotolerant properties with almost 50% of residual activity left after 3 hours at 80°C.

Another method led to describe the investigated enzyme in more detail, which was the determination of kinetic parameters, explaining the enzyme´s catalytic mechanism. These results can be summarized, showing that β-N-acetylglucosaminidase had high affinity to the applied substrate pNPGlcNAc and a high V_{max} value was obtained, which stands for the fact that the enzyme is able to catalyze many reactions in a short time.

The presence of imidazole revealed to be problematic, especially when HPAEC-PAD was applied as method to separate di- and oligosaccharides, which made integration of hydrolysis product peaks difficult. Separating the imidazole content from the purified protein, resulted to be difficult, due to protein precipitation with time. Therefore, for further analysis steps, various other procedures have to be examined for reducing imidazole content.

Optimization study led to detailed understanding of temperature and pH influence on the enzyme. The appliance of this study helped to determine an optimal pH and temperature range and the statistical methodology allowed to be more flexible during process development. Moreover, it was possible to gain insight into enzyme stability and to retrieve information. Besides reducing experimental time and material costs, it was also possible to detect optimal working conditions of the enzyme, which can be helpful for specific applications of this enzyme. β-N-acetylglucosaminidase reveled to be very specific towards its optimum values on pNPGlcNAc, where at pH 5.4 the maximum activity was measured,

which increased with temperature. This screening showed that β-N-acetylglucosaminidase is more influenced by temperature than pH, while analyzing its properties on an aryl substrate.

Results of narrowed range on chitobiose showed optima with higher pH and lower temperature, where a significant influence of pH as well temperature was detected by the model. The applied methodology here was effective in estimating the enzyme behavior under different pH and temperature conditions. Nevertheless, further studies are required to improve enzyme activity, which is possible with the established basis of this optimization investigation.

4 Overall discussion

Novel enzymes of the thermophilic organism *Rhodothermus marinus* were characterized regarding biochemical properties and the obtained results will be compared in this section.

One enzyme of interest (Rmar_2069) did not contain the His_6-tagged tail at the C-terminus, which therefore required a subcloning procedure. In the consequence of not correct cut of applied restriction enzymes, further experiments to achieve subcloning of enzyme were omitted. In addition the measured plasmid concentration before induction was already very low, which leads to believe that it was dealt with a low copy number plasmid. Therefore, the gene was successfully amplified using genomic DNA. Since the last mentioned process step of amplification was not one of the goals of the thesis, characterization with the GH3 enzyme Rmar_1080 was continued. The next step after amplification will be to produce and check the expression level of protein.

Analyzing the characterized enzymes showed that differentiation started already in the length of their amino acid sequence. Therefore, the approximately 17 kDa heavier β-N-acetylglucosaminidase compared to the molecular mass of β-glucosidase enzyme can be explained. Classification of both enzymes as members of GH3 family was mainly based on the amino acid sequence similarity. The aligned sequences displayed low sequence conservation (Figure 39, appendix), especially BGL and NAG shared 16% of similarity. This result is not uncommon among GH3 enzymes. It has been also reported for the β-glucosidase of *T. neapolitana* or β-N-acetylglucosaminidases of *B. subtilis,* when compared these enzymes to barley exohydrolase ExoI [61], [65]. Hence, the low sequence homology as well the large amount of insertions and deletions made homology modeling complicated.

The applied amino acid sequences for the MSA of several GH3 enzymes, where some of these have been structurally characterized, were used to predict the catalytic residues of NAG and BGL as well as to understand their domain architecture. Since the catalytic nucleophile (aspartic acid) is highly conserved among the GH3 family enzymes, always being present in first domain [61], [63], [66] and could be therefore identified for NAG and BGL. Prediction of catalytic acid/base residue revealed to be difficult, because it does not show high conservation among GH3 enzymes [66]. Various β-glucosidases from GH3 family show to exhibit glutamic acid as acid/base residue [60], [61], [63], but determining this catalytic residue for BGL proved to be complicated, since several acid/base catalysts came to consideration, but without certainty. In the case of β-N-acetylglucosaminidase of *B. subtilis*, the residue H234 of the Asp-His dyad was detected to act as the acid/base catalyst, which is supported by Asp232 [65]. The D(S/T)H motif comprises the Asp-His dyad and is conserved among β-N-acetylglucosaminidases [65]. This motif, which is located near the highly

conserved motif KHFPGHG of β-N-acetylglucosaminidases [65], is lacking in family 3 β-glucosidases (Figure 10D). These patterns were also detected for NAG and the β-N-acetylglucosaminidase from an *Alteromonas sp*. (Cht60), which are applied to identify members of this subfamily. In addition, the provided evidence for the catalytic acid/base of *B. subtilis* β-N-acetylglucosaminidase, helped to predict the Asp-His dyad as the acid/base catalyst for NAG and also Cht60. Further structural analysis as well as kinetic data is needed for BGL to determine the catalytic acid/base residue, but also for NAG the predicted residues have to be proved.

Analyzing the domain architecture of both enzymes, it was pointed out that these GH3 enzymes are composed of three domains. The first two domains, glycoside hydrolase family 3 N-terminal and C-terminal domain are found in both enzymes, but differentiation between these occurs in the third domain. For BGL a fibronectin III-like domain (Fn-III) was suggested compared to a β-lactamase domain for NAG. As already mentioned very few crystal structures of GH3 enzymes are stored in the Protein Data Bank [114]. Although several three domain family 3 β-glucosidases are known [61] and [63], however to date none have been investigated for β-N-acetylglucosaminidase.

The third domain of NAG was predicted to represent a β-lactamase structural motif. β-lactamases are widely expressed in bacteria and are responsible for the resistance mechanism against β-lactam antibiotics [115], [116]. This is done by catalyzing the opening as well as hydrolysis of the β-lactam ring of β-lactam antibiotics (e.g. penicillins) [117]. β-lactamases are classified into four groups [117], where the most common class A (penicillinase-type), can be found on transmissible plasmids in Gram-negative bacteria [118]. How exactly this domain plays a role in the overall structure of NAG and if domain-domain interactions take place, has to be investigated further.

Besides knowing the function of NAG´s third domain, only assumptions can be made for Fn-III domain of BGL. In modular proteins, the Fn-III like domains are the most common folds [67], but have not been studied in detail [119], which therefore leads to suggest hypothesis of this domain. For proteins of the GH3 family, where the third domain has a Fn-III fold, some show to be of unknown function such as from *T. neapolitana* [61] or *A. aculeatus* [63]. Whereas for a GH3 family β-glucanase from marine bacterium *Pseudoalteromonas sp.*, the Fn-III domain seems not to be a carbohydrate-binding domain and reveals no relation in sequence and structure to the other glycoside hydrolase 3 domains [62]. This domain finds its role in protein stability and supporting substrates to orient [62]. Other investigations on Fn-III modules conclude that they act as cellulose disruptors, improving hydrolytic ability of cellulases [120]. Another hypothesis was proposed, where the Fn-III modules revealed to

function as spacers between other domains of the protein or work as linker peptides, which are able to link with other enzyme modules and optionally can be extended [67], [121]. May one of these various mentioned functions can be assigned to the third domain of BGL, but for exact knowledge BGL needs to be further studied based on crystal as well as three dimensional structure.

Investigating substrate specificity of characterized enzymes, BGL revealed broad substrate specificity than NAG, which is more selective for N-acetyl-β-D-glucosamine. Similar results were obtained when comparing both with other analyzed β-N-acetylglucosaminidases [92], [109], [112] (but expectations exist [52]) and β-glucosidases [60], [36], [100], [102]. Differences in substrate specificity may be explained through the variation of enzymic nucleophile and general acid/base residue, which make the entrance to the active site of GH3 enzymes. Substrate specificity variation was also detected among GH3 β-glucosidases. Such as for *T. neapolitana* and barley, exhibiting a difference in loop structure in the first two domains, which is located at the entrance to the active site [61]. In addition, for the β-N-acetylglucosaminidase of *B. subtilis* a hypothesis was made, that the enzyme´s high substrate specificity might be traced back to the proper orientation of the determined Asp-His dyad during substrate binding [65]. Furthermore, it is important to mention that those GH3 enzymes which have been purified are tested in a narrow range of possible substrates [66] and it would be of interest to broaden the selection and characterize the enzymes in more detail. This is also true for the assay applied investigating enzyme´s binding towards specific substrate, while carrying out affinity gel electrophoresis. Neither for BGL nor NAG, substrate binding could be detected. Although NAG showed no migration on the gels.

Besides exhibiting difference in substrate specificity, it was even more interesting to compare both enzymes in regard of applied temperature assays. Already during optimum temperature determination, the two members of GH3 family showed their ability to work at high temperature ranges. For BGL lower temperature optimum was obtained (60°C) [93] in contrast to NAG, which showed maximum activity at 90°C (appendix, Figure 41). Primarily characterization of BGL was carried out in previous work as mentioned already [93], where activation energy during temperature optimum assay was not determined and therefore could not be compared to evaluated E_a of NAG. These results of optimum temperature describe the relative temperature stabilities, because temperature optima are dependent on substrate concentration, pH as well incubation time [21]. More information about how stable an enzyme is, gives the investigation of thermal stability.

Comparing thermostability of these enzymes at 80°C revealed that, NAG performed better in regard of residual activity than BGL. Almost 50% of activity was left for NAG, whereas for BGL approximately half of NAG´s residual activity was measured after 180 minutes of incubation (appendix, Figure 42). Furthermore, investigating the obtained apparent half-life of enzyme catalyzed reaction, revealed for BGL at 70°C a $t_{1/2}$ of 79 minutes and for NAG at higher temperature (80°C) a $t_{1/2}$ of 96 minutes.

The β-N-acetylglucosaminidase was characterized based on kinetic parameters determination, which helps to understand enzyme´s metabolism. Due to time limitation enzyme kinetics of β-glucosidase could not be analyzed, but needs to be studied. Since investigation of enzyme kinetics is important, in terms of not only figuring out enzyme´s metabolism and controlling activity, but also revealing its catalytic mechanism and detecting if the enzyme is inhibited.

When investigating enzymes hydrolysis products through application of HPAEC-PAD as method, the chromatograms indicated always the existence of a ghost peak. It was tried to investigate the peak occurrence, due to column installation, contamination of injection port or sample contamination from syringe, rinse vials etc., but could not be traced back to these mentioned explanations. The peak did not only occur during carrying out the assays for the current work, but also in other experimental data from colleagues. As this peak did not influence the analyzed samples, the method was continued. Furthermore, the content of imidazole was visible, when applying HPAEC-PAD. Already mentioned above trials of imidazole separation were carried out for NAG, and the work of M. Jenkins [93] discussed that BGL shows tendency to precipitate, which led to not conduct a dialysis step. In both cases imidazole removal respectively its reduction remained unsuccessful. Therefore, other solutions are needed to solve the issue of imidazole presence in protein samples.

Performance of substrate specificity and temperature assays helped to design further experiments based on determining maximum activity at specific optimum temperature as well pH conditions. The designed experiment of varying both parameters at time, allowed to investigate the GH3 enzymes on different levels, regarding pH and temperature interactions as well as giving the possibility to be more flexible during process development. Statistical methods were applied to model obtained data and investigate effects of parameters. This could only be performed for NAG, because BGL revealed high variation with each measurement of enzyme activity and a shift of optimum with temperature. It is supposed that BGL is very sensitive towards pH and/or temperature changes. Hence, comparison on the basis of model results cannot be carried out. Nevertheless it is possible to compare their

results based on their performance regarding optimal ranges, even though the assays were performed on different substrates.

The determination of maximum activity of BGL and NAG, applying aryl substrates, led to observe a similar trend. In both cases of the investigated glycoside hydrolase family 3 enzymes, the activity rose, with the increase of temperature, independent from pH. Interestingly, the highest enzyme activity for BGL as well NAG was determined at pH 5.4, at each measured temperature point. Whereas the evaluated results for these enzymes on natural substrates were relatively different. BGL´s enzyme activity varied a lot with applied temperature range, where no real trend could be followed. In contrast to BGL, NAG revealed that with decrease of temperature and increase of pH, more chitobiose was hydrolyzed to GlcNAc. These enzymes showed with this investigation that they act differently on aryl and natural substrates.

The characterized enzymes of thermophilic bacteria *R. marinus* proved to be very valuable and interesting candidates for industrial applications. That is due to the fact that both enzymes exhibited the ability to work at high temperatures, with specified half-lives as well as provided results for maximum activity under specific conditions with various substrates.

As β-N-acetylglucosaminidase is known to play a key role during chitin degradation, it might be applied to act during the recycling of the each year highly formed amount of chitin in marine environment [56]. NAG would be an optimal candidate for this kind of appliance, since it shows thermotolerant properties, it would be possible to work at evaluated temperatures, which would lead to an increase of reaction rate and reduction of conversion time but also enzyme amount [21].

In the last years, the aim to produce environmental-friendly products from renewable biomass (e.g. lignocellulosic biomass) increased, which therefore leads researchers to search for novel enzymes that reveal to be stable towards the harsh conditions used during biomass degradation [49]. The characterized BGL would be a preferred enzyme of interest to apply for cellulosic biomass degradation, because it exhibits thermostability and broad substrate specificity, with the assumption of being bifunctional (showed activity on pNPX and pNPG, as well xylobiose and cellooligosaccharides). The appliance of multifunctional enzymes would lead to the advantage that one enzyme has the ability to hydrolyze several oligosaccharides at a time. Such a β-glucosidase, with bifunctional character would reduce enzyme costs, because it could be directly used after steam is applied during saccharification to make lignocellulosic biomass more accessible for hydrolysis and therefore the step of pre-cooling is not needed [44]. Additionally, in that way the time of process could be reduced and

energy saved [122]. Moreover, the use of thermostable BGL would also reduce microbial infections, which would lead to improvement of fermentation yields and qualities [122].

5 Overall conclusions

The novel enzymes β-glucosidase and β-N-acetylglucosaminidase were effectively characterized based on biochemical level, through bioinformatics analysis and by applying a new approach of carrying out an optimization study. These enzymes from marine bacteria exhibited to belong to the glycoside hydrolase family 3 enzymes, which was proven through the investigation based on bioinformatics.

Recombinant protein production and expression as well as purification was successfully carried out for β-N-acetylglucosaminidase. Whereas for further characterization of β-glucosidase an already prepared batch was used, due to purification complications. The MSA was helpful in predication of catalytic acid/base and nucleophile residues of both enzymes. Although the catalytic nucleophile could be determined with high certainty, the acid/base residue proved to be more difficult. In addition, it was possible to detect a difference in the third domain of the enzyme´s domain architecture, where β-glucosidase exhibited a Fn-III-type domain and β-N-acetylglucosaminidase a β-lactamase domain. To understand the domain architecture of these enzymes and to exactly determine the residues, crystal structure and three dimensional investigation is needed.

The characterized enzymes are very interesting in their properties, which show to be promising for various application areas. Due to their thermostability and other determined characteristics (which are summarized below), these valuable enzymes could contribute in the industrial enzyme technology area. Since, both enzymes proved to exhibit thermotolerant properties and β-glucosidase also showed the indication of bifunctionality.

Furthermore, during HPAEC-PAD analysis, the integration of hydrolysis product peaks revealed to be complicated, since these were located in the baseline disruption, which was caused by the content of imidazole in the protein sample. The removal of imidazole proved to be difficult, due to protein precipitation. Therefore, it is very important for further enzyme investigation to solve the problem with the reduction of imidazole in the samples.

Besides, all other applied characterization methods, the analysis of optimal working conditions by determining high enzyme activity through varying temperature and pH, were very effective to understand enzyme stability. Although, this investigation occurred some challenges, such as the imidazole content or the variability of the dataset from β-glucosidase, it still formed the basis for further analysis procedures for optimal activity determination as a new approach. Analyzing the enzyme's activity influencing parameters with such an methodology will not only lead to save experimental time and costs but also to develop a process, which can be applied for any other enzyme, which needs to be characterized. The optimization study gained insight of how easy and quick it is to

understand enzyme´s stability by taking in account the parameters, which influence the enzyme as well as their interactions.

All results of this work are summarized in the following.

Summary of results

- The MSA revealed 16% similarity between NAG and BGL. Catalytic nucleophile was found for NAG at D283 and for BGL at D306, whereas the acid/base residue could only be predicted for NAG at D207 together with H209.
- BGL was lower in molecular weight with 87.18 kDa compared to NAG 104.52 kDa, which was determined with LC-MS.
- Optimum temperature of BGL was 60°C [93] and for NAG a maximum activity was obtained at 90°C. Additionally for NAG, the activation energy was determined, which was 26.86 kJ/mol.
- Thermostability data revealed for BGL higher residual activity at 70°C compared to 80°C, where 80% were left after 3 hours and an apparent half-life of 79 minutes determined between 0 to 60 minutes. NAG showed total inactivation at 90°C after 30 minutes and at 80°C a residual activity of 50% was measured after 180 minutes, with an apparent half-life of 96 minutes, between 0 to 120 minutes.
- Broad substrate specificity was detected for BGL, with high activity on pNPX and xylobiose. Whereas NAG was very specific towards tested substrates and showed higher activity towards pNPGlcNAc and chitooligos (CH_2 to CH_5).
- Determination of kinetic parameters was only carried out for NAG, where a catalytic efficiency of k_{cat}/K_m = 162.30 s^{-1}/mM was obtained.
- The carried out affinity gel electrophoresis with particular substrates, exhibited no binding of both investigated enzymes.
- Studying effect of temperature and pH on both enzymes by applying aryl substrates revealed at the different investigated temperature ranges, always pH 5.4, as a value where BGL and NAG revealed highest activity.

- BGL optimal range on xylobiose shifted with temperature.
- NAG revealed optimal values with increasing pH and decreasing temperature (pH = 6.2 and 65°C-70°C). The model data showed that chitobiose consumption was influenced by both investigated parameters (linear pH and temperature as well as quadratic pH effect).

6 Outlook

In the current work, two novel enzymes were investigated and revealed valuable results. Although the applied assays provided much information about the enzymes of interest, this work does also exhibit some experimental limitations, which can be improved in future work. Therefore, recommendations for these are going to be discussed in the next three sections.

Protein production and purification procedures

The recombinant protein production was performed in 500 ml shake flasks. In the future it would be time saving to produce the proteins in a bioreactor and to scale up. Therefore, it is possible to work with the same batch of prepared protein. Furthermore, it would be beneficial to change the lysing procedures. Since, only applying an ultrasonication step did not yield in a high amount of ruptured cells, afterwards lysing buffer was added. However, the removal of lysing buffer sometimes proves to be difficult, which makes the purification step complicated. Although a purification step is in between, a certain amount of lysing buffer is often still present. Other possible lysing steps, which may be more effective could be to use a homogenizer.

The presence of imidazole made various analysis difficult. As the removal of imidazole through buffer exchange together with a centrifugation step by using a cut-off membrane led to protein precipitation, other possibilities need to be suggested. Either applying a desalting column or following the purification step with a size-exclusion chromatography. Another option would be to add carboxypeptidase during purification, which removes the His-tag and reduces the need of imidazole for solubility.

Characterization

As mentioned above in the overall discussion, to date only few GH3 family enzymes have been structurally characterized. Therefore, if the imidazole content of the protein samples can be eliminated or reduced through the above outlined suggestions, then the protein structure can be examined by applying circular dichroism. Additionally, another possibility to understand the structure in more detail would be the performance of crystal structure trials.

Molecular mass of the proteins showed differences between the computed weight and the weight determined through LC-MS. For exact determination, it would be beneficial to investigate the amino acid sequence, by carrying out N-terminal sequencing.

Although, both enzymes were tested against a wide variety of potential substrates, there still exist some more which can be examined. The work of M. Hong and colleagues, who characterized a β-glucosidase from a GH1 family show some, which were not tested for BGL

(such as pNP-β-D-fucopyranoside, pNP-β-D-lactopyranoside or pNP-β-D-maltopyranoside) [48]. Furthermore, various characterized β-N-acetylglucosaminidases showed to be active on pNP-β-N-acetylgalactosamine [57], [108] which was not screened for NAG. Therefore, it would be of interest to investigate NAG´s substrate specificity on the mentioned substrate. In addition, the assumption that BGL is a bifunctional enzyme should be strengthened by analyzing its substrate specificity between pNPG, pNPX and cellooligos, xylobiose, on which it was active.

Analyzing the binding of enzyme to any of the applied polymeric substrates during affinity gel electrophoresis was negative. Since a small range of different substrates were tested and many more exist as possible candidates, such as laminarin or lichenin, this analysis should be continued. Furthermore, NAG did not migrate at all on the native gels while applying affinity electrophoresis. Two solutions exist to solve this problem: either to change the electrodes positions (anode ↔ cathode) or to use a different buffer system, where the pH is significantly 1 pH value higher than the investigated protein. For this purpose, it is necessary to determine the pI of the proteins through isoelectric focusing, which was not done at this stage of enzyme characterization.

Due to time limitation of this research, the kinetic parameters could only be determined for NAG on pNPGlcNAc. Also BGL needs to be analyzed based on its kinetics parameters on pNPX and it would be interesting to follow these parameters as well on natural substrates.

The thermostability assay obtained very informative results for the characterized enzymes. Besides the applied range of temperature for these enzymes, BGL could be analyzed at 60°C, monitoring BGL´s half-life to investigate how stable it is at its temperature optimum. On the other hand NAG could be screened and examined at 70°C, since various other assays were conducted at that specific temperature. For further investigations, the proteins can be analyzed by their unfolding property, carried out through applying **d**ifferential **s**canning **c**alorimetry (DSC).

Optimizing enzyme activity

While the activity of BGL was investigated by varying pH and temperature, data at 55°C revealed variability with each run. Therefore, BGL´s behavior can be further examined by extending the range to analyze enzyme activity at 70°C. It may be possible that less variability is detected and no shift of optimum with temperature, which would allow to statistically validate and model the data. Since, the optimum range determination by using xylobiose as substrate revealed more than one maxima, the applied range should also be screened on pNPX for certainty.

The hydrolysis of chitobiose at the narrowed range obtained from the screening on pNPGlcNAc, resulted in less product formation. Hence, the determined values differentiated very slightly, which made some conclusions difficult to detect. Reasons behind this, already discussed above, could be the high diluted enzyme. As the applied CCF design (RSM) offers the possibility to investigate many factors at a time, the substrate concentration, which does also influence the enzyme activity, should be taken in account as a dependent variable. In addition, besides the mentioned possibilities, the range of the analyzed enzyme should be extended towards basic pH-range as represented in Figure 33, where the red circles illustrate the additional points to be investigated for further enzyme activity optimization.

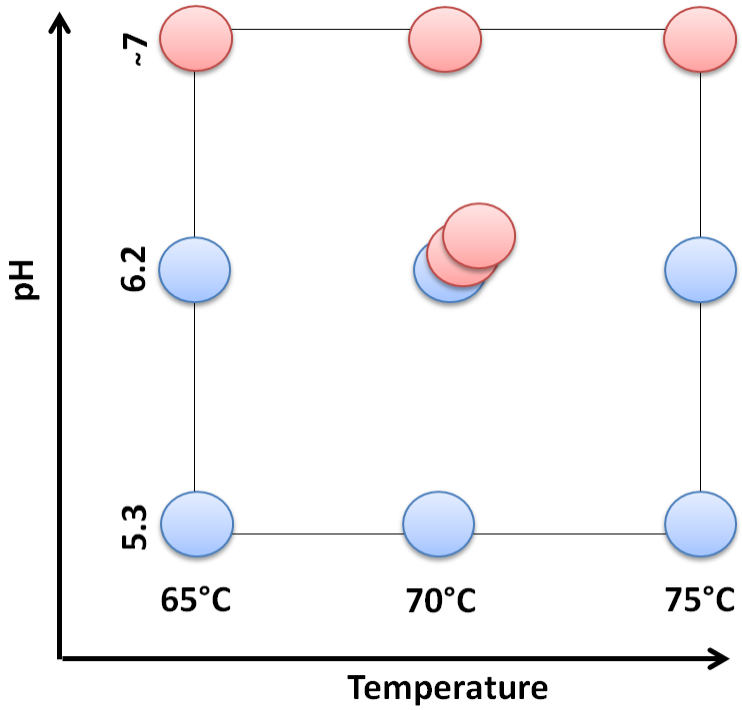

Figure 33: Experimental design, showing an extension of pH-range. The currently conducted enzyme reaction runs are represented in blue circles, whereas the additional points are marked red. On the x-axis the temperature is plotted and on the y-axis the pH-range.

7 Appendix

7.1 Gels of subcloning, SDS-PAGE and affinity electrophoresis analysis

Subcloning experiment:

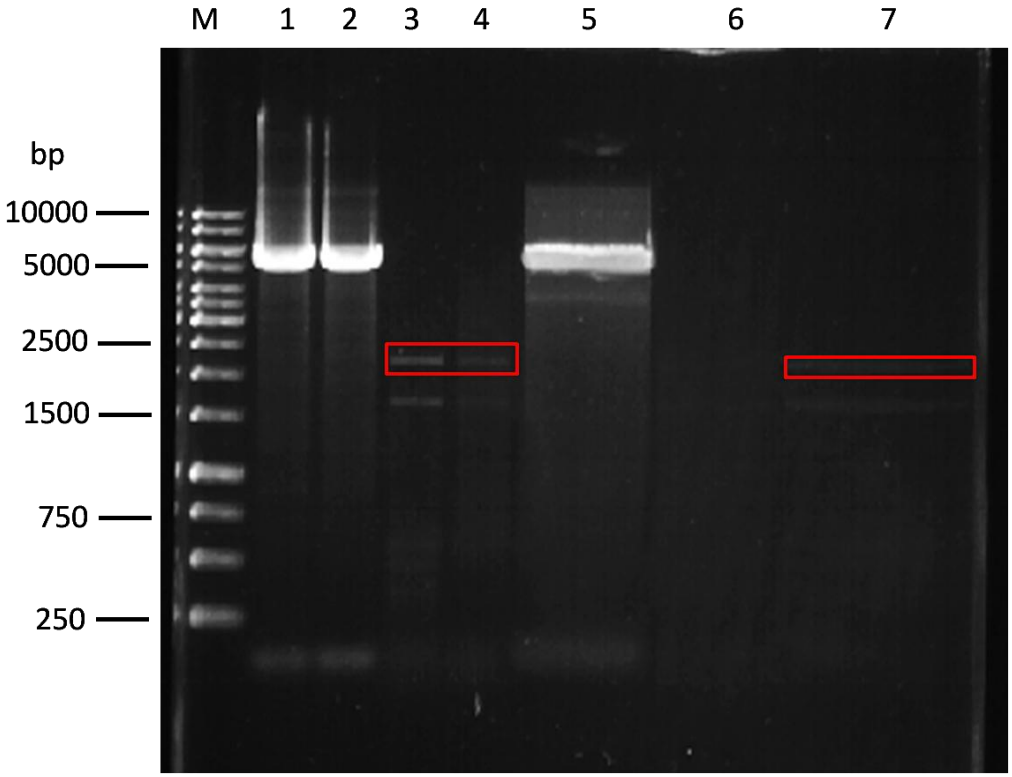

Figure 34: 1% agarose gel showing the digestion of various plasmids, which were digested with NdeI and HindIII. Lane M is the 1 kb generuler DNA ladder. Lane 1,2 shows pET21b plasmid and lane 5 the pET28B plasmid at around 5000 - 6000 bp . The lanes 3, 4, 6 and 7 reveal the pJOE3075 plasmid at slightly below 2500 bp, which is less in its concentration compared to the other plasmids.

Affinity gel electrophoresis:

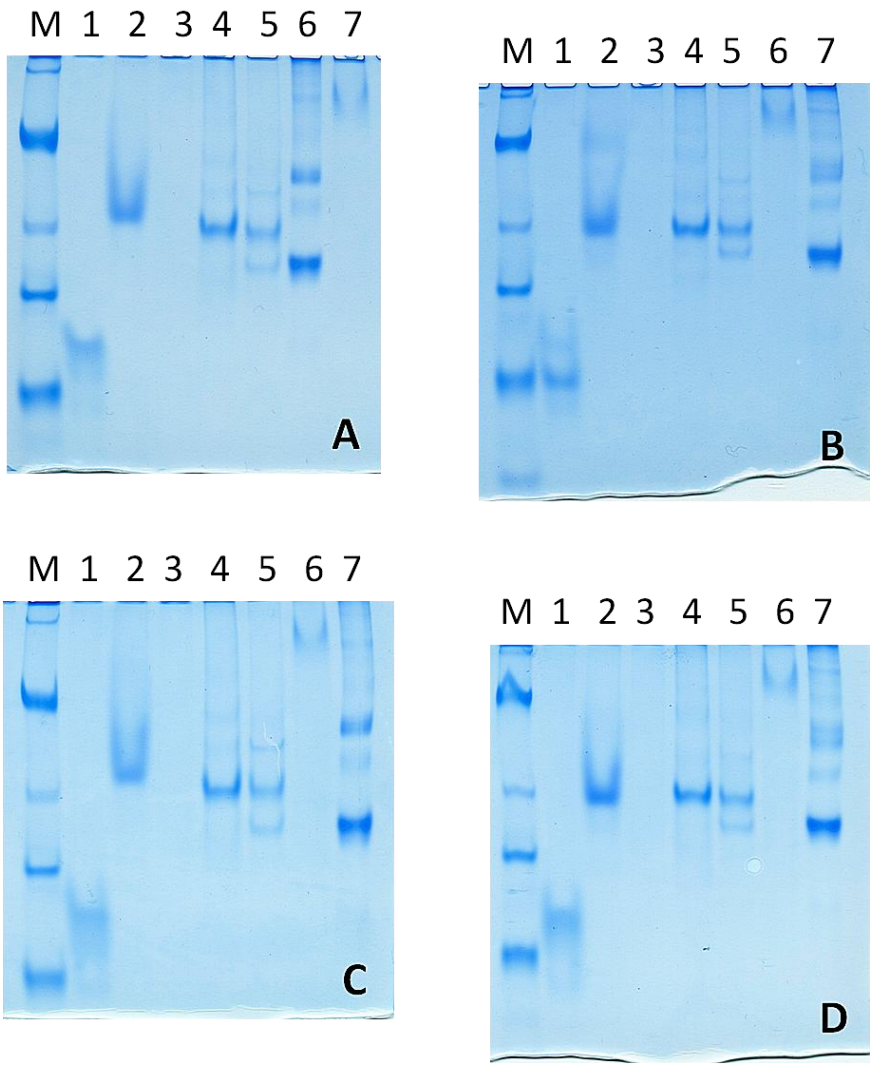

Figure 35: 6.5% native gels, prepared for affinity gel electrophoresis. Figures A and B are gels, where β-glucan was applied as substrate. Whereas for the gels in figure C and D, the substrate hydroxyethyl cellulose was used. On the left side of the figure are the control gels (A, C), where no enzyme was added and on the right side of the figure the gels loaded with sample positioned (B, D). Sample gels are shifted, due to gel alignment. The following is valid for all shown gels, if not stated different. Lane M indicates the used protein ladder, NativeMark (20-1,200 kDa) and in lane 1 a sample of unknown origin was run. Lane 3 belongs to investigated NAG and lane 4 to BGL. In lane 2 the third domain of the TnBgl3 enzyme of *Thermotoga neapolitana* was run. Lane 5 shows the sample of a GH3 enzyme from *Rhodothermus marinus*. In lane 7 (figure A), lane 6 (figure B), lane 6 (figure C and D), another GH3 family member from *R. marinus* is given. The enzyme TnBgl3 was run in the lane 6 (figure A), lane 7 (figure B) and lane 7 (figure C and D). All samples show not binding with the applied substrates.

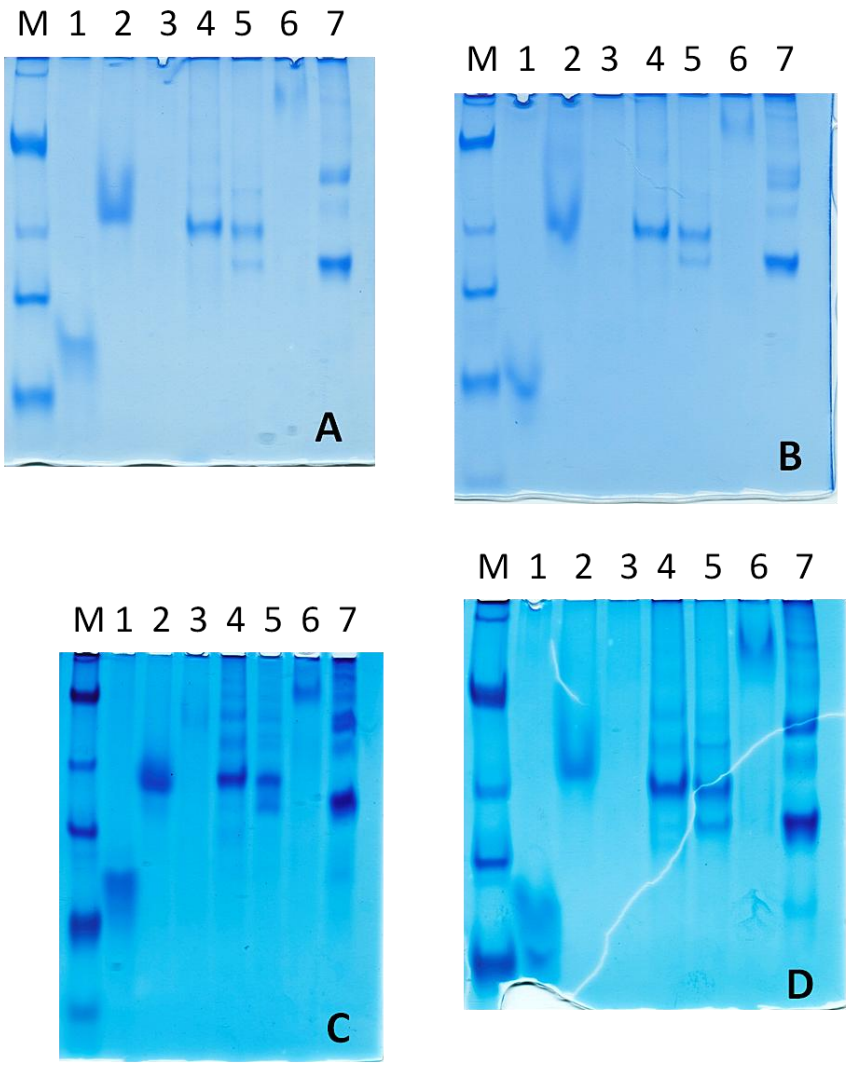

Figure 36: Representing affinity gel electrophoresis, 6.5% native gels. In figure A and B gels are shown, which were run with bound oat spelt xylan. For the gels in figure C and D, xyloglucan was applied as substrate. The figures shown on the left side are the control gels (A, C), where no enzyme was added and on the right side the sample (B, D). Sample gels are shifted, due to gel alignment. The following description of the examined samples is valid for all shown gels. The lane M indicates the used protein ladder, NativeMark (20-1,200 kDa). Lane 1 shows a sample run with unknown origin, lane 2 the third domain of TnBgL3 from *Thermotoga neapolitana*, lane 3 investigated enzyme NAG and in lane 4 BGL. In lane 5 and 6 members of the GH3 family were run. The last lane 7 reveals the migration of TnBGl3. These gels do also reveal no binding of the GH3 enzymes with the substrates.

SDS-PAGE:

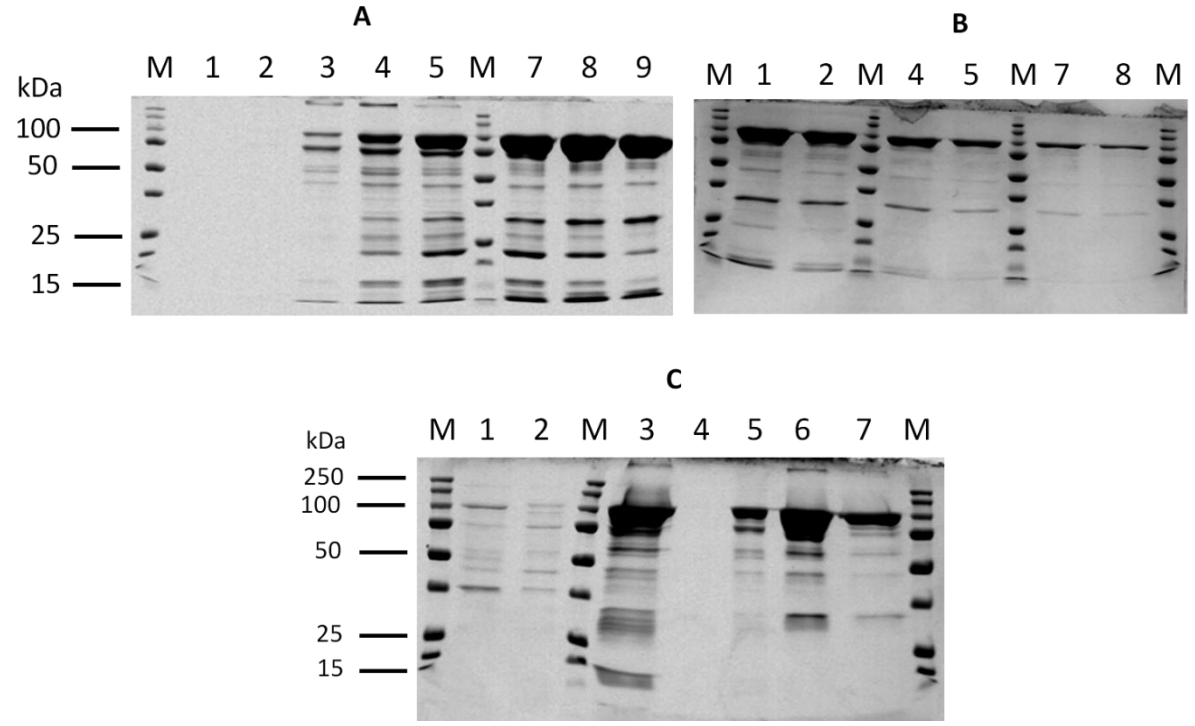

Figure 37: SDS-PAGE gels of protein samples (NAG) from various sources. In all gels, the lane M describes the applied 1 kDa protein ladder. In gels A and B collected fractions during carrying out IMAC of the second batch of produced protein, are loaded in the various pockets of the gel, where some of them show samples with lower protein amount and less additional bands (such as lane 4-8, gel B) compared to others (lanes 4-9, gel A). The gel C, contains protein samples from the purification step (lane 3, 6, 7), expression level check (lane 1 is third hour sample taken during protein production in insoluble form and lane 2, zero hour sample in soluble form) and imidazole removal performance (lane 5 shows the protein sample after buffer exchange and centrifugation step and lane 4 the flow through of this procedure). The protein of interest, NAG is found at approximately 105 kDa, in all shown gels.

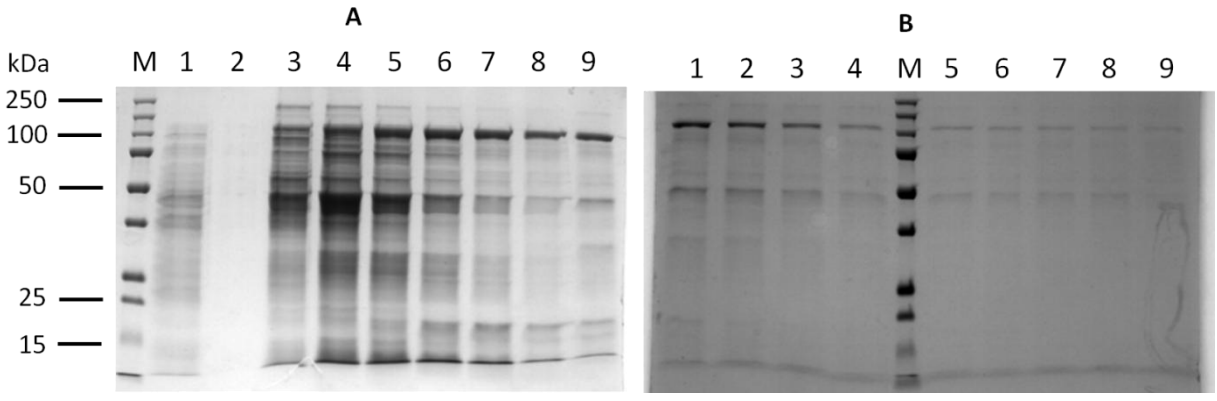

Figure 38: SDS-PAGE gels for evaluation of purification level of the third produced batch of protein. Lane M indicates the applied 1kDa protein ladder. Figure B shows the protein of interest, where NAG is lower in concentration, but contains less additional bands compared to the samples in gel A (lanes 3-6). NAG can be found at approximately 105 kDa.

7.2 Bioinformatics analysis - MSA

In Figure 39, the complete MSA of the investigated GH3 family members is shown. Due to the large amino acid sequences, the MSA is divided into three parts, which contain different domains.

Figure 39: Multiple sequence alignment of GH3 family β-glucosidases and β-N-acetylglucosaminidases obtained from ClustalW2 program [79], [80]. The amino acid sequence of the investigated β-glucosidase (BGL) and β-N-acetylglucosaminidase (NAG) from *Rhodothermus marinus* were compared to the β-glucosidase from *Thermotoga neapolitana* TnBgl3B (UniProt: Q0GC07), *Hordeum vulgare* ExoI (UniProt: Q9XEI3) and *Aspergillus aculeatus* AaBGl1 (UniProt: P48825) as well as to the β-N-acetylglucosaminidase from *Bacillus subtilis* BsNAGZ (UniProt: P40406) and *Alteromonas sp. strain O-7* Cht60 (UniProt: P48823). Identical or similar amino acid residues of the aligned GH3 family members are denoted by asterisks and dots, respectively. In this figure, the first domain of investigated GH3 family members is shown.

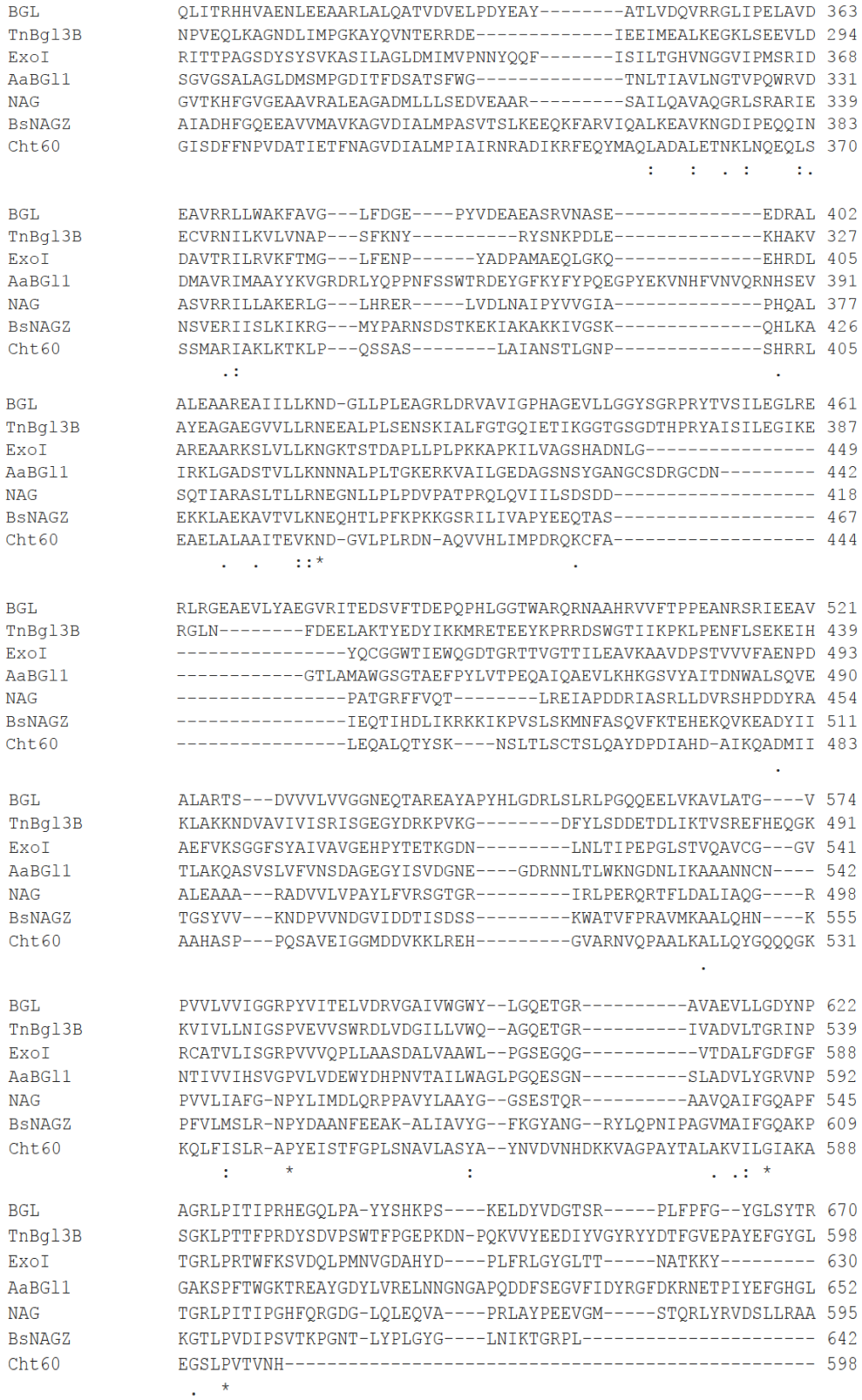

Figure 31: Continued multiple sequence alignment of GH3 family β-glucosidases and β-N-acetylglucosaminidases obtained from ClustalW2 program [79], [80]. In this figure, the first domain of all aligned glycosidases is continued, without that of ExoI and the second domain of investigated GH3 family members is shown.

```
BGL             GFRRVHLG-----------------PGERKVVEIELG-PEAFAFYGLEMERVVEAGWFD 767
TnBgl3B         AFHKTRLLN----------------PGESEEVVLEIP-VRDLASFNGEEWVVEAGEYEV 696
ExoI            ------------------------------------------------------------
AaBGl1          STDLKASSGDPYYGVDTAEHVPEGATDGSPQPVLPAGGGSGGNPRLYDELIRVSVTVKNT 772
NAG             AMQLYEAGKLDLDAPVVRYLPEFGQNGKERVTIRQLLS-HTAGLAPFHPFHRMGITTAEA 710
BsNAGZ          ------------------------------------------------------------
Cht60           ------------------------------------------------------------

BGL             VLVGGNSEELISVPLEITEGCNLGR---------------------------------- 792
TnBgl3B         RVGASSRNIKLKGTFSVGEERRFKP---------------------------------- 721
ExoI            ------------------------------------------------------------
AaBGl1          GRVAGDAVPQLYVSLGGPNEPKVVLRKFDRLTLKPSEETVWTTTLTRRDLSNWDVAAQDW 832
NAG             VRQAILSDSLIYEPGTQSRYSDLGMIVLGWVIERITGQPLDRYAAEHIFRPLGMRHTGFR 770
BsNAGZ          ------------------------------------------------------------
Cht60           ------------------------------------------------------------
BGL             ------------------------------------------------------------
TnBgl3B         ------------------------------------------------------------
ExoI            ------------------------------------------------------------
AaBGl1          VITSYPKKVHVGSSSRQLPLHAALPKVQ-------------------------------- 860
NAG             PVGRPDTTVVPTEIDTIFRHRLIQGEVHDETAWILGGVAGHAGLFSTAEDLARFAYMLVN 830
BsNAGZ          ------------------------------------------------------------
Cht60           ------------------------------------------------------------

BGL             ------------------------------------------------------------
TnBgl3B         ------------------------------------------------------------
ExoI            ------------------------------------------------------------
AaBGl1          ------------------------------------------------------------
NAG             EGRIGGRPFLKPETIRLFTTPVDPERAGTRALGWDTRSREGYSSAGRLFGSRSFGHTGFT 890
BsNAGZ          ------------------------------------------------------------
Cht60           ------------------------------------------------------------

BGL             ------------------------------------------------------------
TnBgl3B         ------------------------------------------------------------
ExoI            ------------------------------------------------------------
AaBGl1          ------------------------------------------------------------
NAG             GTSIWIDPDQQLFVILLTNRVYPTRENRKHLAVRARLADLAYEALIGPPTLNLNALLP 948
BsNAGZ          ------------------------------------------------------------
Cht60           ------------------------------------------------------------
```

Figure 31: Continued multiple sequence alignment of GH3 family β-glucosidases and β-N-acetylglucosaminidases obtained from ClustalW2 program [79], [80].. In this figure, the second domain of AaBGl1 is continued and the third domain of investigated GH3 family members is shown.

7.3 Purification

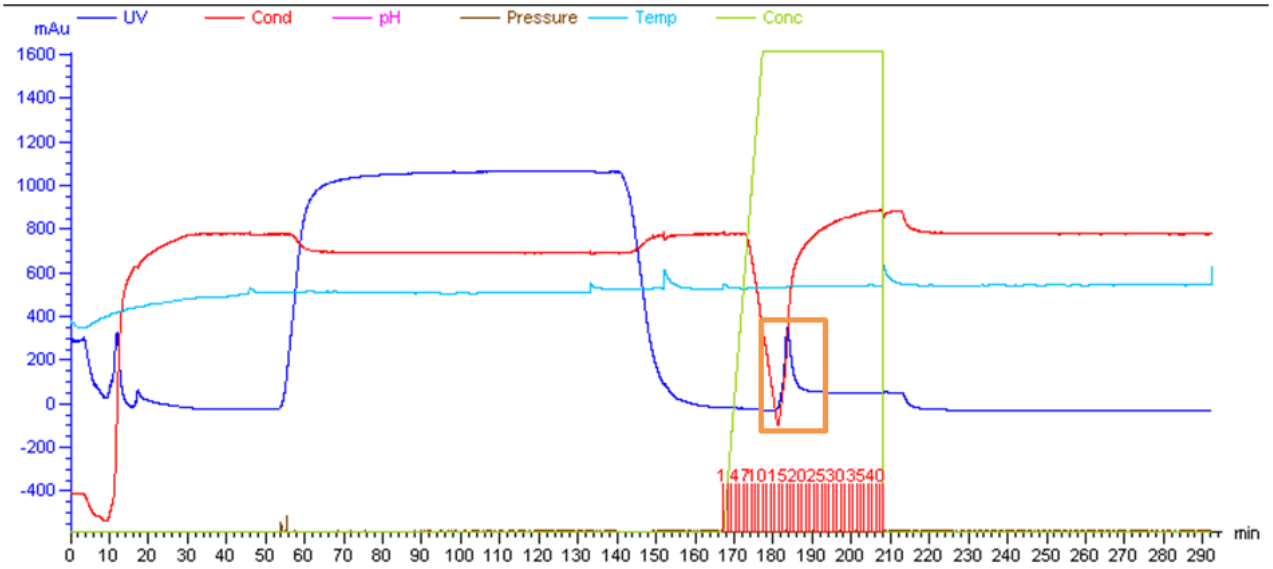

Figure 40: Purification profile of β-N-acetylglucosaminidase applying IMAC. The purification step was very quick and clean, which can be followed by the orange colored box, indicating the eluted peak containing the protein of interest. The protein was detected by measuring the UV-absorbance at 280 nm (blue line). In addition, conductivity (red line) was also followed, which decreased as soon as the protein eluted. Temperature (light blue line) and pressure (brown line) were constant during purification. The concentration is indicated as green colored line.

7.4 Thermal stability

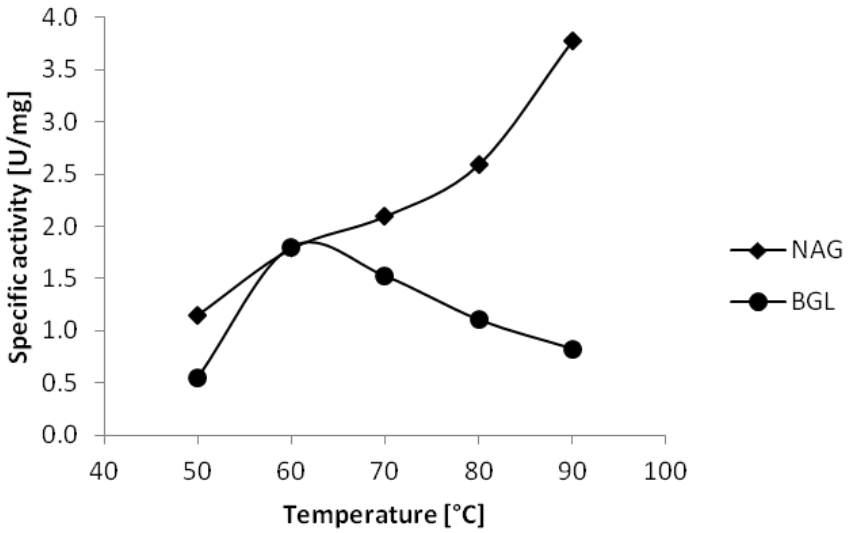

Figure 41: Determination of β-N-acetylglucosaminidase (NAG) and β-glucosidase (BGL) (Data of BGL obtained from previous work of M. Jenkins [93]) temperature optimum between 50°C to 90°C, where reactions were followed continuously. Measured specific activity values are plotted against the applied temperature range. Highest activity was monitored at 90°C for NAG and 60°C for BGL.

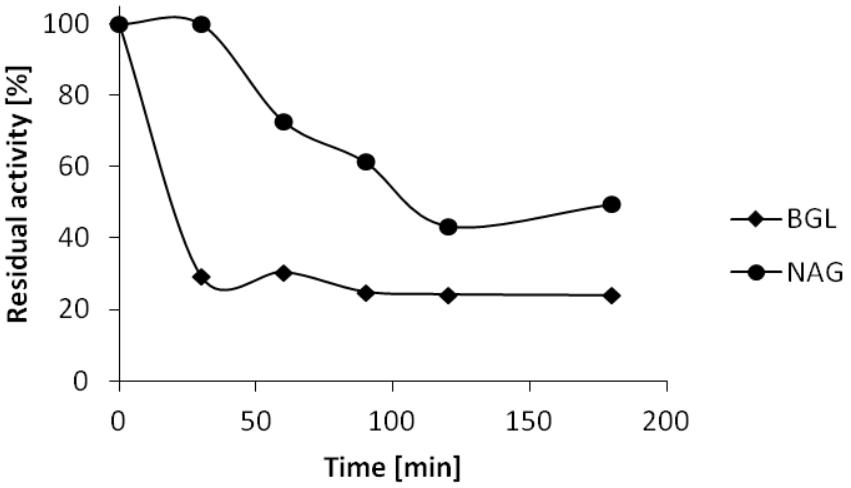

Figure 42: Thermal stability curves of β-N-acetylglucosaminidase (NAG) and β-glucosidase (BGL) at 80°C. The residual activity was determined between 0 to 180 minutes. NAG reveals higher residual activity compared to BGL. The activity of NAG decreases slowly, whereas BGL shows rapid decrease in activity, remaining almost same during the three hours of incubation.

7.5 Effect of temperature and pH - optimization study

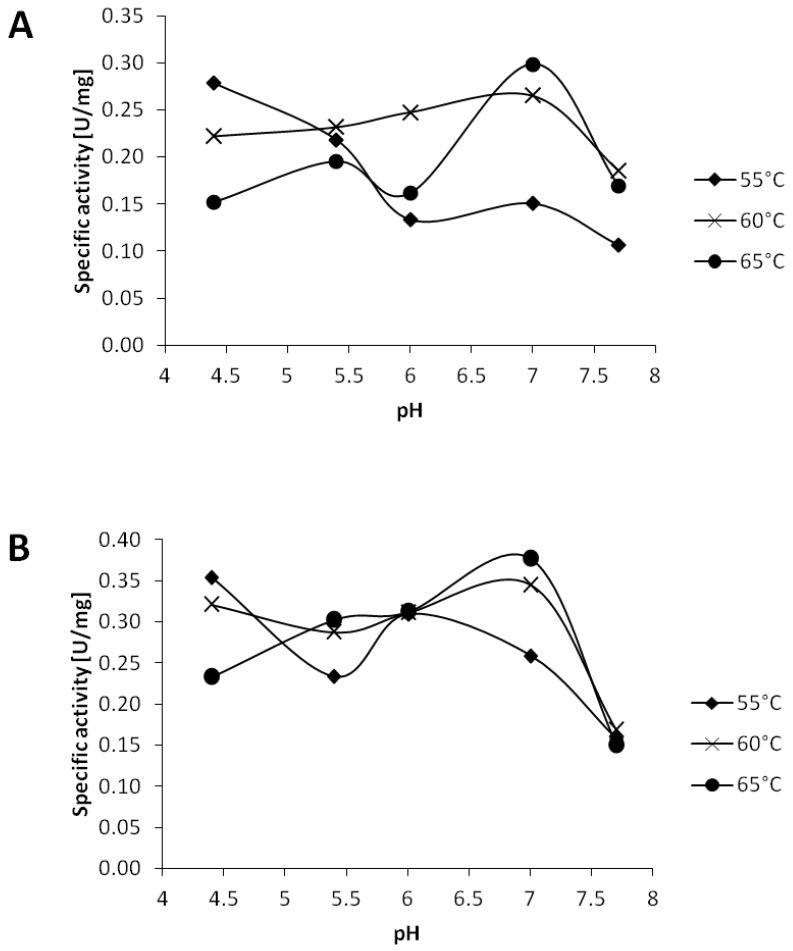

Figure 43: Optimization of specific activity of β-glucosidase with xylobiose as substrate. Determined activity is plotted against a broad range of pH-values in regard to three different applied enzyme assay incubation temperatures, at which in figure A the first run and in figure B the second run is represented. Investigating these runs shows the variability between them, especially at 55°C.

7.6 References

[1] J. R. Cherry and A. L. Fidantsef, “Directed evolution of industrial enzymes: an update,” *Current Opinion in Biotechnology*, vol. 14, no. 4, pp. 438–443, Aug. 2003.

[2] Business Communications Company Inc., “Global Markets for Enzymes in Industrial Applications,” 2012. [Online]. Available: http://www.bccresearch.com/market-research/biotechnology/enzymes-industrial-applications-markets-bio030g.html.

[3] D. Demirjian, F. Morís-Varas, and C. Cassidy, “Enzymes from extremophiles,” *Current opinion in chemical biology*, pp. 144–151, 2001.

[4] M. de Champdoré, M. Staiano, M. Rossi, and S. D’Auria, “Proteins from extremophiles as stable tools for advanced biotechnological applications of high social interest.,” *Journal of the Royal Society, Interface / the Royal Society*, vol. 4, no. 13, pp. 183–91, Apr. 2007.

[5] P. Turner, G. Mamo, and E. N. Karlsson, “Potential and utilization of thermophiles and thermostable enzymes in biorefining.,” *Microbial cell factories*, vol. 6, p. 9, Jan. 2007.

[6] S. Kumar, C. J. Tsai, and R. Nussinov, “Factors enhancing protein thermostability.,” *Protein engineering*, vol. 13, no. 3, pp. 179–91, Mar. 2000.

[7] G. Haki, “Developments in industrially important thermostable enzymes: a review,” *Bioresource Technology*, vol. 89, no. 1, pp. 17–34, Aug. 2003.

[8] A. Giordano, G. Andreotti, A. Tramice, and A. Trincone, “Marine glycosyl hydrolases in the hydrolysis and synthesis of oligosaccharides.,” *Biotechnology journal*, vol. 1, pp. 511–530, 2006.

[9] C. Vieille and G. Zeikus, “Hyperthermophilic enzymes: sources, uses, and molecular mechanisms for thermostability,” *Microbiology and Molecular Biology Reviews*, vol. 65, no. 1, pp. 1–43, 2001.

[10] F. Niehaus, C. Bertoldo, M. Kähler, and G. Antranikian, “Extremophiles as a source of novel enzymes for industrial application.,” *Applied microbiology and biotechnology*, vol. 51, no. 6, pp. 711–29, Jun. 1999.

[11] M. Podar and A.-L. Reysenbach, “New opportunities revealed by biotechnological explorations of extremophiles.,” *Current opinion in biotechnology*, vol. 17, no. 3, pp. 250–5, Jun. 2006.

[12] K. Mavromatis, N. Ivanova, I. Anderson, A. Lykidis, S. D. Hooper, H. Sun, V. Kunin, A. Lapidus, P. Hugenholtz, B. Patel, and N. C. Kyrpides, “Genome Analysis of the Anaerobic Thermohalophilic Bacterium Halothermothrix orenii,” *PLoS ONE*, vol. 4, no. 1, p. e4192, Jan. 2009.

[13] X. Huang, Y. Zhao, Y. Dai, G. Wu, Z. Shao, Q. Zeng, and Z. Liu, “Cloning and biochemical characterization of a glucosidase from a marine bacterium Aeromonas sp. HC11e-3.,” *World journal of microbiology & biotechnology*, vol. 28, no. 12, pp. 3337–44, Dec. 2012.

[14] J. Kennedy, J. R. Marchesi, and A. D. Dobson, “Marine metagenomics: strategies for the discovery of novel enzymes with biotechnological applications from marine environments.,” *Microbial cell factories*, vol. 7, p. 27, Jan. 2008.

[15] G. A. Alfredsson, “Rhodothermus marinus, gen. nov., sp. nov., a thermophilic, halophilic bacterium from submarine hot springs in Iceland,” *Journal of general microbiology*, no. 1988, pp. 299–306, 1988.

[16] M. Nolan and B. Tindall, “Complete genome sequence of Rhodothermus marinus type strain (R-10T),” *Standards in genomic sciences*, pp. 283–291, 2009.

[17] S. Bjornsdottir and T. Blondal, “Rhodothermus marinus: physiology and molecular biology,” *Extremophiles*, pp. 1–16, 2006.

[18] G. O. Hreggvidsson, E. Kaiste, O. Holst, G. Eggertsson, A. Palsdottir, G. O. Hreggvidsson, E. V. A. Kaiste, O. Holst, and G. Eggertsson, “An Extremely Thermostable Cellulase from the Thermophilic Eubacterium Rhodothermus marinus These include : An Extremely Thermostable Cellulase from the Thermophilic Eubacterium Rhodothermus marinus,” 1996.

[19] F. Baneyx, “Recombinant protein expression in Escherichia coli,” *Current opinion in biotechnology*, 1999.

[20] M. Del Pozo, “Microbial β-glucosidases from cow rumen metagenome enhance the saccharification of lignocellulose in combination with commercial cellulase cocktail,” *Biotechnology for …*, vol. 5, no. 1, p. 73, Jan. 2012.

[21] O. P. Ward and M. Moo-Young, “Thermostable enzymes,” *Biotechnology advances*, vol. 6, no. 1, pp. 39–69, 1988.

[22] C. J. Yeoman, Y. Han, D. Dodd, C. M. Schroeder, R. I. Mackie, and I. K. O. Cann, “Thermostable enzymes as biocatalysts in the biofuel industry.,” *Advances in applied microbiology*, vol. 70, no. 10, pp. 1–55, Jan. 2010.

[23] “CAZypedia-Carbohydrate-active enzymes.” [Online]. Available: http://www.cazypedia.org/. [Accessed: 01-Nov-2013].

[24] D. G. Naumoff, “Hierarchical classification of glycoside hydrolases.,” *Biochemistry. Biokhimiia*, vol. 76, pp. 622–635, 2011.

[25] B. L. Cantarel, P. M. Coutinho, C. Rancurel, T. Bernard, V. Lombard, and B. Henrissat, “The Carbohydrate-Active EnZymes database (CAZy): an expert resource for Glycogenomics.,” *Nucleic acids research*, vol. 37, no. Database issue, pp. D233–8, Jan. 2009.

[26] “Carbohydrate Active Enzymes database.” [Online]. Available: http://www.cazy.org/.

[27] B. Henrissat and G. Daviest, “Structural and sequence-based classification of glycoside hydrolases,” *Current Opinion in structural biology*, pp. 637–644, 1997.

[28] D. E. Koshland, “Stereochemistry and the mechanism of enzymic reactions,” *Biol. Revs. Cambridge Phil. Soc.*, vol. 28, pp. 416–36, 1953.

[29] T. V Vuong and D. B. Wilson, “Glycoside hydrolases: catalytic base/nucleophile diversity.,” *Biotechnology and bioengineering*, vol. 107, no. 2, pp. 195–205, Oct. 2010.

[30] D. Wilson, “Cellulases and biofuels,” *Current Opinion in Biotechnology*, vol. 20, no. 3, pp. 295–299, Jun. 2009.

[31] M. Scigelova and D. H. G. Crout, “Microbial β-N-acetylhexosaminidases and their biotechnological applications,” *Enzyme and Microbial Technology*, vol. 25, no. 1–2, pp. 3–14, Jul. 1999.

[32] D. Faure, “The family-3 glycoside hydrolases: from housekeeping functions to host-microbe interactions,” *Applied and environmental microbiology*, vol. 68, no. 4, pp. 1485–1490, 2002.

[33] D. Dodd, S. Kiyonari, R. I. Mackie, and I. K. O. Cann, “Functional diversity of four glycoside hydrolase family 3 enzymes from the rumen bacterium Prevotella bryantii B14.,” *Journal of bacteriology*, vol. 192, no. 9, pp. 2335–45, May 2010.

[34] G. Dotsenko, O. Sinitsyna, S. Hinz, J. Wery, and A. Sinitsyn, “Characterization of a GH family 3 β-glycoside hydrolase from Chrysosporium lucknowense and its application to the hydrolysis of β-glucan and xylan.,” *Bioresource technology*, vol. 112, pp. 345–9, May 2012.

[35] N. Aït, N. Creuzet, and J. Cattaneo, “Properties of β-glucosidase purified from Clostridium thermocellum,” *Journal of general microbiology*, pp. 569–577, 1982.

[36] L. Ruttersmith and R. Daniel, “Thermostable β-glucosidase and β-xylosidase from Thermotoga sp. strain FjSS3-B.1,” *Biochimica et Biophysica Acta*, vol. 1156, pp. 167–172, 1993.

[37] D. K. Watt, H. Ono, and K. Hayashi, “Agrobacterium tumefaciens beta-glucosidase is also an effective beta-xylosidase, and has a high transglycosylation activity in the presence of alcohols.,” *Biochimica et biophysica acta*, vol. 1385, no. 1, pp. 78–88, Jun. 1998.

[38] Y. Bhatia, S. Mishra, and V. S. Bisaria, “Microbial beta-glucosidases: cloning, properties, and applications.,” *Critical reviews in biotechnology*, vol. 22, no. 4, pp. 375–407, Jan. 2002.

[39] J. R. Ketudat Cairns and A. Esen, “β-Glucosidases.,” *Cellular and molecular life sciences : CMLS*, vol. 67, no. 20, pp. 3389–405, Oct. 2010.

[40] T.-H. Park, K.-W. Choi, C.-S. Park, S.-B. Lee, H.-Y. Kang, K.-J. Shon, J.-S. Park, and J. Cha, “Substrate specificity and transglycosylation catalyzed by a thermostable beta-glucosidase from marine hyperthermophile Thermotoga neapolitana.,” *Applied microbiology and biotechnology*, vol. 69, no. 4, pp. 411–22, Dec. 2005.

[41] W.-Y. Jeng, N.-C. Wang, M.-H. Lin, C.-T. Lin, Y.-C. Liaw, W.-J. Chang, C.-I. Liu, P.-H. Liang, and A. H.-J. Wang, “Structural and functional analysis of three β-glucosidases from bacterium Clostridium cellulovorans, fungus Trichoderma reesei and termite Neotermes koshunensis.,” *Journal of structural biology*, vol. 173, no. 1, pp. 46–56, Jan. 2011.

[42] J. Langston, N. Sheehy, and F. Xu, “Substrate specificity of Aspergillus oryzae family 3 beta-glucosidase.,” *Biochimica et biophysica acta*, vol. 1764, no. 5, pp. 972–8, May 2006.

[43] S. Fang, J. Chang, Y. S. Lee, W. Guo, Y. L. Choi, and Y. Zhou, “Cloning and characterization of a new broadspecific β-glucosidase from Lactococcus sp. FSJ4.,” *World journal of microbiology & biotechnology*, Jul. 2013.

[44] D. Liu, R. Zhang, X. Yang, Z. Zhang, S. Song, Y. Miao, and Q. Shen, “Characterization of a thermostable β-glucosidase from Aspergillus fumigatus Z5, and its functional expression in Pichia pastoris X33.,” *Microbial cell factories*, vol. 11, no. 1, p. 25, Jan. 2012.

[45] R. Breves and K. Bronnenmeier, “Genes encoding two different beta-glucosidases of Thermoanaerobacter brockii are clustered in a common operon.,” *Applied and environmental microbiology*, vol. 63, no. 10, pp. 3902–3910, 1997.

[46] D. Faure, J. Desair, and V. Keijers, “Growth of Azospirillum irakense KBC1 on the Aryl β-Glucoside Salicin Requires either salA or salB,” *Journal of bacteriology*, vol. 181, no. 10, pp. 3003–3009, 1999.

[47] R. J. Gruninger, X. Gong, R. J. Forster, and T. a McAllister, “Biochemical and kinetic characterization of the multifunctional β-glucosidase/β-xylosidase/α-arabinosidase, Bgxa1.,” *Applied microbiology and biotechnology*, Aug. 2013.

[48] M.-R. Hong, Y.-S. Kim, C.-S. Park, J.-K. Lee, Y.-S. Kim, and D.-K. Oh, “Characterization of a recombinant beta-glucosidase from the thermophilic bacterium Caldicellulosiruptor saccharolyticus.,” *Journal of bioscience and bioengineering*, vol. 108, no. 1, pp. 36–40, Jul. 2009.

[49] L. Bao, Q. Huang, L. Chang, Q. Sun, J. Zhou, and H. Lu, “Cloning and characterization of two β-glucosidase/xylosidase enzymes from yak rumen metagenome.,” *Applied biochemistry and biotechnology*, vol. 166, no. 1, pp. 72–86, Jan. 2012.

[50] M. D. Balcewich, K. a Stubbs, Y. He, T. W. James, G. J. Davies, D. J. Vocadlo, and B. L. Mark, “Insight into a strategy for attenuating AmpC-mediated beta-lactam resistance: structural basis for selective inhibition of the glycoside hydrolase NagZ.,” *Protein science : a publication of the Protein Society*, vol. 18, no. 7, pp. 1541–51, Jul. 2009.

[51] F. Liu, T. Liu, M. Qu, and Q. Yang, “Molecular and biochemical characterization of a novel β-N-acetyl-D-hexosaminidase with broad substrate-spectrum from the Aisan corn borer, Ostrinia furnacalis.,” *International journal of biological sciences*, vol. 8, no. 8, pp. 1085–96, Jan. 2012.

[52] C. Mayer, D. J. Vocadlo, M. Mah, K. Rupitz, D. Stoll, R. a J. Warren, and S. G. Withers, “Characterization of a beta-N-acetylhexosaminidase and a beta-N-acetylglucosaminidase/beta-glucosidase from Cellulomonas fimi.,” *The FEBS journal*, vol. 273, no. 13, pp. 2929–41, Jul. 2006.

[53] Q. Yang, T. Liu, F. Liu, M. Qu, and X. Qian, “A novel beta-N-acetyl-D-hexosaminidase from the insect Ostrinia furnacalis (Guenée).,” *The FEBS journal*, vol. 275, no. 22, pp. 5690–702, Dec. 2008.

[54] R. Caspi, T. Altman, J. M. Dale, K. Dreher, C. a Fulcher, F. Gilham, P. Kaipa, A. S. Karthikeyan, A. Kothari, M. Krummenacker, M. Latendresse, L. a Mueller, S. Paley, L. Popescu, A. Pujar, A. G. Shearer, P. Zhang, and P. D. Karp, “The MetaCyc database of metabolic pathways and enzymes and the BioCyc collection of pathway/genome databases.,” *Nucleic acids research*, vol. 38, no. Database issue, pp. D473–9, Jan. 2010.

[55] H. Orikoshi, N. Baba, and S. Nakayama, “Molecular analysis of the gene encoding a novel cold-adapted chitinase (ChiB) from a marine bacterium, Alteromonas sp. strain O-7,” *Journal of bacteriology*, vol. 185, no. 4, pp. 1153–1160, 2003.

[56] T. Tanaka, S. Fujiwara, and S. Nishikori, “A unique chitinase with dual active sites and triple substrate binding sites from the hyperthermophilic archaeon Pyrococcus kodakaraensis KOD1,” *Applied and environmental microbiology*, 1999.

[57] H. Tsujibo, N. Hatano, and T. Mikami, “A Novel beta-N-Acetylglucosaminidase from Streptomyces thermoviolaceus OPC-520: Gene Cloning, Expression, and Assignment to Family 3 of the Glycosyl Hydrolases,” *Applied and environmental microbiology*, vol. 64, no. 8, pp. 2920–2924, 1998.

[58] K. Sakai, M. Narihara, Y. Kasama, M. Wakayama, and M. Moriguchi, “Purification and characterization of thermostable beta-N-acetylhexosaminidase of Bacillus stearothermophilus CH-4 isolated from chitin-containing compost.,” *Applied and environmental microbiology*, vol. 60, no. 8, pp. 2911–5, Aug. 1994.

[59] N. Dahiya, R. Tewari, and G. S. Hoondal, “Biotechnological aspects of chitinolytic enzymes: a review.,” *Applied microbiology and biotechnology*, vol. 71, no. 6, pp. 773–82, Aug. 2006.

[60] J. N. Varghese, M. Hrmova, and G. B. Fincher, “Three-dimensional structure of a barley beta-D-glucan exohydrolase, a family 3 glycosyl hydrolase.,” *Structure (London, England : 1993)*, vol. 7, no. 2, pp. 179–90, Feb. 1999.

[61] T. Pozzo, J. L. Pasten, E. N. Karlsson, and D. T. Logan, “Structural and functional analyses of beta-glucosidase 3B from Thermotoga neapolitana: a thermostable three-domain representative of glycoside hydrolase 3.,” *Journal of molecular biology*, vol. 397, no. 3, pp. 724–39, Apr. 2010.

[62] Y. Nakatani, S. M. Cutfield, N. P. Cowieson, and J. F. Cutfield, “Structure and activity of exo-1,3/1,4-β-glucanase from marine bacterium Pseudoalteromonas sp. BB1 showing a novel C-terminal domain.,” *The FEBS journal*, vol. 279, no. 3, pp. 464–78, Feb. 2012.

[63] S. Kentaro, S. Jun-Ichi, and N. Young-Woo, “Crystal structures of glycoside hydrolase family 3 beta-glucosidase 1 from Aspergillus aculeatus,” *Biochemical society*, vol. 221, pp. 211–221, 2013.

[64] K. a Stubbs, M. Balcewich, B. L. Mark, and D. J. Vocadlo, “Small molecule inhibitors of a glycoside hydrolase attenuate inducible AmpC-mediated beta-lactam resistance.,” *The Journal of biological chemistry*, vol. 282, no. 29, pp. 21382–91, Jul. 2007.

[65] S. Litzinger, S. Fischer, P. Polzer, K. Diederichs, W. Welte, and C. Mayer, “Structural and kinetic analysis of Bacillus subtilis N-acetylglucosaminidase reveals a unique Asp-His dyad mechanism.,” *The Journal of biological chemistry*, vol. 285, no. 46, pp. 35675–84, Nov. 2010.

[66] a J. Harvey, M. Hrmova, R. De Gori, J. N. Varghese, and G. B. Fincher, “Comparative modeling of the three-dimensional structures of family 3 glycoside hydrolases.,” *Proteins*, vol. 41, no. 2, pp. 257–69, Nov. 2000.

[67] J.-G. Jee, T. Ikegami, M. Hashimoto, T. Kawabata, M. Ikeguchi, T. Watanabe, and M. Shirakawa, “Solution structure of the fibronectin type III domain from Bacillus circulans WL-12 chitinase A1.,” *The Journal of biological chemistry*, vol. 277, no. 2, pp. 1388–97, Jan. 2002.

[68] L. Eriksson, E. Johansson, N. Kettaneh-Wold, C. Wikström, and S. Wold, *Design of Experiments: Principles and Applications*, 3rd ed. Umeå, Sweden: Umetrics, 2008.

[69] C. Mandenius and A. Brundin, “Bioprocess optimization using design-of-experiments methodology,” *Biotechnology progress*, 2008.

[70] “Citric Acid – Na2HPO4 Buffer Solutions, pH approx. 2.6–7.6.” [Online]. Available: http://www.sigmaaldrich.com/life-science/core-bioreagents/biological-buffers/learning-center/buffer-reference-center.html#citric.

[71] “Potassium Phosphate Buffer.” [Online]. Available: http://www.unl.edu/cahoonlab/phosphate buffer.pdf.

[72] “Glycine– NaOH Buffer, pH 8.6–10.6, PKA = 9.78.” [Online]. Available: http://microscopy.berkeley.edu/Resources/instruction/buffers.html.

[73] M. Punta, P. C. Coggill, R. Y. Eberhardt, J. Mistry, J. Tate, C. Boursnell, N. Pang, K. Forslund, G. Ceric, J. Clements, A. Heger, L. Holm, E. L. L. Sonnhammer, S. R. Eddy, A. Bateman, and R. D. Finn, “The Pfam protein families database,” *Oxford Journals*, vol. 40, no. November 2011, pp. 290–301, 2012.

[74] E. Quevillon, V. Silventoinen, S. Pillai, N. Harte, N. Mulder, R. Apweiler, and R. Lopez, “InterProScan: protein domains identifier.,” *Nucleic acids research*, vol. 33, no. Web Server issue, pp. W116–20, Jul. 2005.

[75] A. Marchler-Bauer, S. Lu, J. B. Anderson, F. Chitsaz, M. K. Derbyshire, C. DeWeese-Scott, J. H. Fong, L. Y. Geer, R. C. Geer, N. R. Gonzales, M. Gwadz, D. I. Hurwitz, J. D. Jackson, Z. Ke, C. J. Lanczycki, F. Lu, G. H. Marchler, M. Mullokandov, M. V Omelchenko, C. L. Robertson, J. S. Song, N. Thanki, R. a Yamashita, D. Zhang, N. Zhang, C. Zheng, and S. H. Bryant, “CDD: a Conserved Domain Database for the functional annotation of proteins.,” *Nucleic acids research*, vol. 39, no. Database issue, pp. D225–9, Jan. 2011.

[76] A. Marchler-Bauer, J. B. Anderson, F. Chitsaz, M. K. Derbyshire, C. DeWeese-Scott, J. H. Fong, L. Y. Geer, R. C. Geer, N. R. Gonzales, M. Gwadz, S. He, D. I. Hurwitz, J. D. Jackson, Z. Ke, C. J. Lanczycki, C. a Liebert, C. Liu, F. Lu, S. Lu, G. H. Marchler, M. Mullokandov, J. S. Song, A. Tasneem, N. Thanki, R. a Yamashita, D. Zhang, N. Zhang, and S. H. Bryant, “CDD: specific functional annotation with the Conserved Domain Database.,” *Nucleic acids research*, vol. 37, no. Database issue, pp. D205–10, Jan. 2009.

[77] A. Marchler-Bauer and S. H. Bryant, “CD-Search: protein domain annotations on the fly.,” *Nucleic acids research*, vol. 32, no. Web Server issue, pp. W327–31, Jul. 2004.

[78] T. U. Consortium, “Update on activities at the Universal Protein Resource (UniProt) in 2013,” *Oxford Journals*, vol. 41, no. November 2012, pp. 43–47, 2013.

[79] M. A. Larkin, G. Blackshields, N. P. Brown, R. Chenna, P. A. Mcgettigan, H. Mcwilliam, F. Valentin, I. M. Wallace, A. Wilm, R. Lopez, J. D. Thompson, T. J. Gibson, and D. G. Higgins, “Clustal W and Clustal X version 2 . 0,” *Oxford Journals*, vol. 23, no. 21, pp. 2947–2948, 2007.

[80] M. Goujon, H. Mcwilliam, W. Li, F. Valentin, S. Squizzato, J. Paern, and R. Lopez, “A new bioinformatics analysis tools framework at EMBL – EBI,” *Oxford Journals*, vol. 38, no. May 2010, pp. 695–699, 2010.

[81] H. Block, B. Maertens, A. Spriestersbach, N. Brinker, J. Kubicek, R. Fabis, J. Labahn, and F. Schäfer, “Reprint of: Immobilized-Metal Affinity Chromatography (IMAC): A Review.,” *Methods in Enzymology*, vol. 463, pp. 439–473, Sep. 2009.

[82] U. K. Laemmli, “Cleavage of structural proteins during the assembly of the head of bacteriophage T4.,” *Nature*, vol. 227, pp. 680–685, 1970.

[83] P. Smith, R. Krohn, and G. Hermanson, “Measurement of protein using bicinchoninic acid,” *Analytical biochemistry*, vol. 85, pp. 76–85, 1985.

[84] “ExPASy - SIB Bioinformatics Resource Portal.”

[85] Kroll K., “Characterization of thermostable ß-glucosidase Tn Bgl3B and evaluation of possibilities in transglycosylation reactions,” 2007.

[86] L. Michaelis, M. L. Menten, R. S. Goody, and K. A. Johnson, “Die Kinetik der Invertinwirkung The Kinetics of Invertase Action translated by,” no. February, 1913.

[87] H. Lineweaver and D. Burk, “The determination of enzyme dissociation constants,” *Journal of the American Chemical Society*, 1934.

[88] J. Burdon and S. Haldane, “A note on the kinetics of enzyme action,” *Biochem J.*, 1925.

[89] A. Cornish-Bowden, *Fundamentals of Enzyme Kinetics*, Fourth. 2012, pp. 33–39, 251–280.

[90] V. V Kadnikov, A. V Mardanov, O. A. Podosokorskaya, S. N. Gavrilov, I. V Kublanov, A. V Beletsky, E. A. Bonch-osmolovskaya, and N. V Ravin, “Genomic Analysis of Melioribacter roseus , Facultatively Anaerobic Organotrophic Bacterium Representing a Novel Deep Lineage within Bacteriodetes / Chlorobi Group,” vol. 8, no. 1, 2013.

[91] G. Xie and D. Bruce, “Genome sequence of the cellulolytic gliding bacterium Cytophaga hutchinsonii,” *Applied and environmental microbiology*, vol. 73, no. 11, pp. 3536–3546, 2007.

[92] T. Hiroshi, F. Kazuhiro, T. Hiromi, and M. Katsushiro, “Gene sequence, purification and characterization of N-acetyl-β-glucosaminidase from a marine bacterium, Alteromonas sp. strain O-7,” *Gene*, vol. 146, pp. 111–115, 1994.

[93] M. Jenkins, “Functional characterization of four glycoside hydrolase 3 enzymes isolated from Rhodothermus marinus,” (Report), Lund University, Department of Biotechnology, 2012.

[94] N.-Y. Gu, J.-L. Kim, H.-J. Kim, D.-J. You, H.-W. Kim, and S.-J. Jeon, “Gene cloning and enzymatic properties of hyperthermostable beta-glycosidase from Thermus thermophilus HJ6.,” *Journal of bioscience and bioengineering*, vol. 107, no. 1, pp. 21–6, Jan. 2009.

[95] J. Pei, Q. Pang, L. Zhao, S. Fan, and H. Shi, “Thermoanaerobacterium thermosaccharolyticum β -glucosidase : a glucose-tolerant enzyme with high specific activity for cellobiose,” *Biotechnology for biofuels*, vol. 5:31, 2012.

[96] Y. Kim, S. Yeom, and D. Oh, “Characterization of a GH3 family β-glucosidase from Dictyoglomus turgidum and its application to the hydrolysis of isoflavone glycosides in spent coffee grounds,” *Journal of agricultural and food chemistry*, pp. 11812–11818, 2011.

[97] K. Goyal, P. Selvakumar, and K. Hayashi, “Characterization of a thermostable β-glucosidase (BglB) from Thermotoga maritima showing transglycosylation activity,” *Journal of Molecular Catalysis B: Enzymatic*, vol. 15, pp. 45–53, 2001.

[98] A. Karnaouri, E. Topakas, T. Paschos, I. Taouki, and P. Christakopoulos, “Cloning, expression and characterization of an ethanol tolerant GH3 β-glucosidase from Myceliophthora thermophila.,” *PeerJ*, vol. 1, p. e46, Jan. 2013.

[99] S. M. Pitson, R. J. Seviour, and B. M. McDougall, “Purification and characterization of an extracellular beta-glucosidase from the filamentous fungus Acremonium persicinum and its probable role in beta-glucan degradation.,” *Enzyme and microbial technology*, vol. 21, no. 3, pp. 182–90, Aug. 1997.

[100] C. Riou, J. Salmon, and M. Vallier, “Purification, characterization, and substrate specificity of a novel highly glucose-tolerant β-glucosidase from Aspergillus oryzae,” *Applied and environmental microbiology*, vol. 64, pp. 3607–3614, 1998.

[101] P. Christakopoulos, P. W. Goodenough, D. Kekos, B. J. Macris, M. Claeyssens, and M. K. Bhat, “Purification and characterisation of an extracellular beta-glucosidase with transglycosylation and exo-glucosidase activities from Fusarium oxysporum.,” *European journal of biochemistry / FEBS*, vol. 224, no. 2, pp. 379–85, Sep. 1994.

[102] P. Murray, N. Aro, C. Collins, A. Grassick, M. Penttilä, M. Saloheimo, and M. Tuohy, “Expression in Trichoderma reesei and characterisation of a thermostable family 3 beta-glucosidase from the moderately thermophilic fungus Talaromyces emersonii.,” *Protein expression and purification*, vol. 38, no. 2, pp. 248–57, Dec. 2004.

[103] S. Vroemen, J. Heldens, C. Boyd, B. Henrissat, and N. T. Keen, “Cloning and characterization of the bgxA gene from Erwinia chrysanthemi D1 which encodes a beta-glucosidase/xylosidase enzyme.,” *Molecular general genetics MGG*, vol. 246, pp. 465–477, 1995.

[104] M. Ueda and A. Motoo, “Purification and some properties of beta-N-Acetylglucosaminidase from Aeromonas sp. 10S-24,” *Bioscience, Biotechnology, and Biochemistry*, vol. 56, pp. 1204–1207, 1992.

[105] L. Eriquez and M. Pisano, “Purification and characterization of an extracellular beta-n-acetylhexosaminidase from Paecilomyces persicinus.,” *Journal of bacteriology*, vol. 137, no. 1, pp. 620–626, 1979.

[106] R. Soto-Gil and J. Zyskind, “N, N’-diacetylchitobiase of Vibrio harveyi. Primary structure, processing, and evolutionary relationships.,” *Journal of Biological Chemistry*, vol. 264, no. 25, pp. 14778–14783, 1989.

[107] F. Nanjo, M. Ishikawa, R. Katsumi, and K. Sakai, “Purification, properties, and transglycosylation reaction of beta-N-acetylhexosaminidase from Nocardia orientalis.,” *Agricultural and Biological Chemistry*, vol. 54, no. 4, pp. 899–906, 1990.

[108] Y.-L. Jin, Y.-Y. Jo, K.-Y. Kim, J.-H. Shim, Y.-W. Kim, and R.-D. Park, “Purification and characterization of beta-N-acetylhexosaminidase from rice seeds.,” *Journal of biochemistry and molecular biology*, vol. 35, no. 3, pp. 313–9, May 2002.

[109] K.-H. Choi, J. Y. Seo, K.-M. Park, C.-S. Park, and J. Cha, “Characterization of glycosyl hydrolase family 3 beta-N-acetylglucosaminidases from Thermotoga maritima and Thermotoga neapolitana.,” *Journal of bioscience and bioengineering*, vol. 108, no. 6, pp. 455–9, Dec. 2009.

[110] J. Lisboa De Marco, M. C. Valadares-Inglis, and C. R. Felix, “Purification and characterization of an N-acetylglucosaminidase produced by a Trichoderma harzianum strain which controls Crinipellis perniciosa.,” *Applied microbiology and biotechnology*, vol. 64, no. 1, pp. 70–5, Mar. 2004.

[111] D. S. L. Souza, M. F. Grossi-de-Sa, L. P. Silva, O. L. Franco, J. E. Gomes-Junior, G. R. Oliveira, T. L. Rocha, C. P. Magalhães, B. M. Marra, M. Grossi-de-Sa, E. Romano, C. M. de Sá, E. Kombrink, A. V Jiménez, and L. R. D. Abreu, “Identification of a novel beta-N-acetylhexosaminidase (Pcb-NAHA1) from marine Zoanthid Palythoa caribaeorum (Cnidaria, Anthozoa, Zoanthidea).,” *Protein expression and purification*, vol. 58, no. 1, pp. 61–9, Mar. 2008.

[112] S. Roseman, “Molecular Cloning and Characterization of a Novel beta -N-Acetyl-D-glucosaminidase from Vibrio furnissii,” *Journal of Biological Chemistry*, vol. 271, no. 52, pp. 33433–33439, Dec. 1996.

[113] C. S. Farinas, M. M. Loyo, A. Baraldo, P. W. Tardioli, V. B. Neto, and S. Couri, “Finding stable cellulase and xylanase: evaluation of the synergistic effect of pH and temperature.,” *New biotechnology*, vol. 27, no. 6, pp. 810–5, Dec. 2010.

[114] H. M. Berman, J. Westbrook, Z. Feng, G. Gilliland, T. N. Bhat, H. Weissig, I. N. Shindyalov, and P. E. Bourne, “The Protein Data Bank.,” *Nucleic acids research*, vol. 28, no. 1, pp. 235–42, Jan. 2000.

[115] K. C. Usher, L. C. Blaszczak, G. S. Weston, B. K. Shoichet, and S. J. Remington, “Three-dimensional structure of AmpC beta-lactamase from Escherichia coli bound to a transition-state analogue: possible implications for the oxyanion hypothesis and for inhibitor design.,” *Biochemistry*, vol. 37, no. 46, pp. 16082–92, Nov. 1998.

[116] M. Oliva, O. Dideberg, and M. J. Field, “Understanding the acylation mechanisms of active-site serine penicillin-recognizing proteins: a molecular dynamics simulation study.,” *Proteins*, vol. 53, pp. 88–100, 2003.

[117] J. R. Knox and P. C. Moews, “Beta-lactamase of Bacillus licheniformis 749/C. Refinement at 2 A resolution and analysis of hydration.,” *Journal of molecular biology*, vol. 220, no. 2, pp. 435–55, Jul. 1991.

[118] J. I. Campbell, S. Scahill, T. Gibson, and R. P. Ambler, “The phototrophic bacterium Rhodopseudomonas capsulata sp108 encodes an indigenous class A beta-lactamase.,” *The Biochemical journal*, vol. 260, no. 3, pp. 803–12, Jul. 1989.

[119] E. Little, P. Bork, and R. F. Doolittle, “Tracing the spread of fibronectin type III domains in bacterial glycohydrolases.,” *Journal of molecular evolution*, vol. 39, pp. 631–643, 1994.

[120] I. Kataeva, R. Seidel, and A. Shah, “The fibronectin type 3-like repeat from the Clostridium thermocellum cellobiohydrolase CbhA promotes hydrolysis of cellulose by modifying its surface,” *Applied and environmental microbiology*, 2002.

[121] M. Alahuhta, Q. Xu, R. Brunecky, W. S. Adney, S.-Y. Ding, M. E. Himmel, and V. V Lunin, “Structure of a fibronectin type III-like module from Clostridium thermocellum.,” *Acta crystallographica. Section F, Structural biology and crystallization communications*, vol. 66, no. Pt 8, pp. 878–80, Aug. 2010.

[122] D. Liu, R. Zhang, X. Yang, Y. Xu, Z. Tang, W. Tian, and Q. Shen, “Expression, purification and characterization of two thermostable endoglucanases cloned from a lignocellulosic decomposing fungi Aspergillus fumigatus Z5 isolated from compost.,” *Protein expression and purification*, vol. 79, no. 2, pp. 176–86, Oct. 2011.