LUND UNIVERSITY

Molecular development of a thermostable -glucosidase for modification of natural products

Khan, Sami

2012

Molecular development of a thermostable β-glucosidase for modification of natural products

Samiullah Khan

LUND UNIVERSITY

Doctoral Thesis
Biotechnology, Department of Chemistry

Academic thesis which, by due permission of the Faculty of Engineering at Lund University, will be publicly defended on Monday 23 January 2012, at 13:15 in lecture hall B, at the Centre for Chemistry and Chemical Engineering, Getingevägen 60, Lund, for the degree of Doctor of Philosophy in Engineering.

Faculty opponent: Dr. Nushin Aghajari, France.

Doctoral Thesis
Biotechnology, Department of Chemistry
Lund University
P.O. Box 124, SE-22100
Sweden

ISBN 978-91-89627-80-2
ISRN LUTKDH/TKBT--12/1134—SE
Printed in Sweden by Media-Tryck, Lund University
Lund 2011

To My Mother

Organization LUND UNIVERSITY	Document name DOCTORAL DISSERTATION
	Date of issue 2011-12-28
Author(s) Samiullah Khan	Sponsoring organization

Title and subtitle

Molecular development of a thermostable β-glucosidase for modification of natural products

Abstract

The gene encoding a β-glucosidase originating from the extreme thermophile *Thermotoga neopolitana* has been cloned and expressed in *Escherichia coli*. The aim was to produce a thermostable enzyme that could be used to remove sugar residues from glucosylated natural products classified as flavonoids by applying a method combining extraction in hot pressurized water with enzymatic hydrolysis.

The β-glucosidase (termed *Tn*Bgl1A in this thesis) is a member of family 1 of glycoside hydrolase (GH1). The enzyme has an apparent unfolding temperature of 101.9 °C and a molecular weight of 52.6 kDa. The activity of *Tn*Bgl1A was first analysed using the model substrate para-nitrophenyl-β-D-glucopyranoside (*p*NPGlc), demonstrating that a single glucosyl residue (typical for β-glucosidases) is released by this enzyme. Hydrolysis of glucosylated forms of the natural product quercetin (the major flavonoid in yellow onion) was found to be dependent on the position of the glucosylation (see below).

Expression in *E. coli* resulted in a relatively large fraction of insoluble target protein. To improve the folding of *Tn*Bgl1A during production, different strategies were then applied: I) the gene was constructed synthetically and codons were optimized to match codon usage in *E.coli*, II) the gene was cloned in frame with a signal peptide to translocate the protein to the periplasmic space, and III) the gene was co-expressed with genes encoding molecular chaperones. Among these strategies the co-expression of the gene with chaperones worked best, and the improved folding resulted in an increased fraction of soluble, active enzyme.

The *Tn*Bgl1A was tested for the hydrolysis of quercetin-glucosides, which are antioxidants classified as flavonoids present in onion, and composed of a polyphenolic backbone glucosylated at two different positions (Q3, Q4´, and the diglucoside Q3,4´). The aglycone form of quercetin is a more potent antioxidant than the glucosylated forms. The activity of *Tn*Bgl1A for Q3 was lower compared to its activity on quercetin-4´-glucoside (Q4´). To improve hydrolysis of Q3, mutations were introduced in the enzyme, based on a structure model of *Tn*Bgl1A. The mutant N221S/P342L showed increased efficiency towards Q3 as well as for Q4´ compared to wild-type. This showed that the position and nature of this residue in the active site is important for substrate specificity and that by careful selection the specificity can be changed for different substrates. Therefore, the mutation studies were extended and the active site region was targeted for further mutagenesis. Among the changes introduced, the mutagenesis of the neighbouring residue, N220S, was also found to influence activity, and this variant had a higher specific activity for quercetin-glucosides.

*Tn*Bgl1A and one of the best performance mutants (N221S/P342L) were immobilised on acrylic support to allow recycling of the enzyme in experiments coupling hot water extraction of quercetin-glucosides with enzymatic hydrolysis. The activity of the immobilised enzyme was analysed in batch experiments using *p*NPGlc as model substrate and showed not only that the enzyme remained active after immobilisation, but that the thermal stability of the enzyme improved slightly. The effect of additives on immobilisation was studied and, as glucose is an activator for the enzyme, the addition of glucose during immobilisation resulted in a slight increase in the specific activity. In the case of bovine serum albumin (BSA), the original activity was recovered after three months of storage by incubation with BSA for 24 hours.

In conclusion, the work in the present thesis shows that *Tn*Bgl1A is an effective biocatalyst for the conversion of quercetin-glucosides to quercetin. It was also found that, in the production step, more soluble enzymes can be obtained if they are co-expressed with chaperones. The specificity of the enzyme can be changed by changing a single amino acid in the active site, and the improved hydrolysis of Q3 found for two single mutants (N220S and N221S) is caused by indirect changes in interactions, which may lead to a better fit of the quercetin backbone in the active site. As the modification of amino acids in the active site requires a deep understanding of the structure, it is hoped that the results reported here can contribute to the creation of, new mutants with better activity guided by predictions based on the current results. In this research, we have shown that both free and immobilised enzyme can be coupled to the hot water extraction process and that the immobilised enzyme can be used in an on-line process for hydrolysis of flavonoid glucosides. Further improvements can also be made in such combined processes, both concerning the conditions for the extraction and hydrolysis.

Key words

Thermotoga neapolitana, *Tn*Bgl1A, Mutagenesis,Flavonoid, Quercetin, Immobilisation, Chaperones

Classification system and/or index terms (if any)

Supplementary bibliographical information	Language English
ISSN and key title	ISBN 978-91-89627-80-2
Recipient´s notes	Number of pages 208 / Price
	Security classification

Distribution by (name and address)

Contents

Acknowledgements

First and above all, I praise God, the almighty for providing me this opportunity and granting me the capability to proceed successfully. Blessing of God be upon His Prophet Muhammad (PBUH), the city of knowledge, who has guided His Ummah to seek knowledge from Cradle to Grave.

This thesis appears in its current form is not solely my personal achievement, but the resultant contribution of many persons to whom I would like to express my gratitude.

I would like thank to:

Prof. Rajni Hatti-Kaul who gave me the opportunity to pursue my PhD studies at the department of Biotechnology.

I find immense pleasure in expressing my sincere regards and profound gratitude to my research supervisor Prof. Eva Nordberg Karlsson, for your guidance and accepting me as a PhD student. I owe the greatest depth of gratitude for your invaluable help, strong motivation, encouragement and support. Apart from your guidance, you have given me the experience, built confidence and lifted my spirits, which cannot be described in words. All of that was very important for somebody who has no previous knowledge about molecular biology/enzyme technology like me. This thesis would not have turned out like that without your help and support. We had a lot of fun at conferences, dinners and I hope that our fruitful collaboration will continue in the future.

My co-supervisor Charlotta Turner from Centre for analysis and synthesis, Lund University, you have been always such a nice supervisor for me, encouraged me, offered me valuable advice and always interested in my work, arranged social events during "SuReTech" meetings. Because of your professional approach I learnt a lot, also thanks for a lot of baby stuff you gave me.

My Co-supervisor Olle Holst for always being so nice, for a lot of recommendation letters whenever I need it, and for urging me to learn Swedish. Tack så mycket.

Gashaw Mamo you have been always supportive and helpful. I took the opportunity of your presence for scientific discussion everywhere I found you, in kitchen, corridor, lab and office.

Chrisitina Wennerberg thanks for your help with cultivation, teaching me how to cast SDS-PAGE, Swedish translation and for the baby stuff.

Fatima Plieva, Harald Kirsebom thank you for introducing me to Cryogel for enzyme immobilisation and Wilfred for your help with Cryogel preparation. Tejas Kulkarni and Tania Pozzo thanks for creating the beautiful structure pictures and for much other fun together. Carla thanks for reading and commenting on my thesis and for the different fun and discussion we had together.

I would like to express my gratitude to all my co-authors and collaboration partners. Sofia Lindhal you were always so nice. I am really pleased to have collaboration with you.

I would like to thank Tahir Mahmood and Shakeel for whom I was involved in supervision. I hope I have not discouraged you from research in the future. I wish you good luck in your future.

My special thanks to Malik Badshah, dear you were always present for my help and never said no, even though you were busy with a lot of your personal stuff. Malik you are simply an obedient young brother.

Present and former colleagues of the fermentation group: Mushtaq Ahmad, Catherine Paul, Peter Falck, Rawana Alkhalili, Ami Patel, Kazi Zubaida, Nihir Shah, Joon, Emanuel and Top for being around me and for discussions about research and other activities.

Thanks Paula Leckius for your big help with printing of the thesis.

Siv for your kindness, being always available for help and to solve problems, Kersti for your support and Frans-Peder for your help with computer and instrument related problems.

My former and present friends at Biotechnology: M. Ibrahim, Yasser, Tarek, Serena, Roya, David, Ravi, Khalid, Mahmoud, Govind, Naresh, M. Takwa Daniel, Georgina, M.Ali, Abolghasem, Reza, Maryam, Javier, Thuy, Ali, Zeeshan, Suhaila, Cecilia, Marlene, Solmaz, Pontus, Victor, Hugo, Deepti, Kosin, Ramin, Nagwa, Amin, Rosa, Ivo, Valentine, Emma and the list is still incomplete……..

I would like to thank Abdoh, Merichel, Jiayin, Arwa and Monica from analytical chemistry.

I would like to express many thanks to all my Pakistani Friends in Lund specially Munir, Nadeem, Murtaza, Asghar, Ashfaq, Taimoor, Naveed Razzaq, Asim, Rashid, Irfan, Kaka, Nadir, Ayaz, Saeed, Mansoor, Mushtaq Farooqi, Ilyas and Liaqat for providing me nice company. I would like to thank Faiz Ullah Shah from Lulea University for providing me every possible support. I still remember when you received me on my arrival.

I would like to express my gratitude to my former supervisor Fariha Hasan and my room-mate Aamer Ali Shah for encouraging me, considering doing PhD from abroad which was not initially my plan. Thanks to all my Quadian and Gomalian friends.

I am grateful to the Higher Education Commission (HEC) of Pakistan for providing my living expenses in Sweden and the Swedish Institute (SI) for managing the scholarship. I am also thankful to the Swedish research council Formas for financially supporting the research work and Kungliga Fysiografiska Sällskapet for travelling grants for the conferences.

Finally I would like to thank my family for all the unconditional love, support and prayers specially my mother whose prayers enabled me to reach to this position. My feelings will not be completed without thanks to my dearest brothers, Sairullah Khan, M. Ali Khan and M. Nisar Khan, for your never-ending support and encouragement and for being always available whenever I need help. To my sisters, your love, continuous encouragement and prayers enabled me to accomplish this humble effort. My sincere thanks to my parents in law who always prayed for smooth finishing of my PhD.

I am indebted to express my thanks to my lovely wife, Sobia and our little angel Eshal. Your love, understanding, encouragement, prayers and sacrifices made all of this possible.

List of Papers

This thesis is a compilation of the following papers, listed with Roman numerals in the order in which they appear in the text.

I. Aglycone specificity of *Thermotoga neapolitana* β-glucosidase 1A modified by mutagenesis, leading to increased catalytic efficiency in quercetin-3-glucoside hydrolysis.
Samiullah Khan, Tania Pozzo, Marton Megyeri, Sofia Lindahl, Anders Sundin, Charlotta Turner and Eva Nordberg Karlsson
BMC Biochemistry (2011), 12**:**11

II. Exploring the possibility of using a thermostable mutant of β-glucosidase for rapid hydrolysis of quercetin glucosides in hot water.
Sofia Lindahl, Anna Ekman, **Samiullah Khan**, Chritina Wennerberg, Pål Börjesson, Per Sjöberg, Eva Nordberg Karlsson and CharlottaTurner
Green Chemistry (2010), 12:1, 159-169
(http://pubs.rsc.org/en/content/articlelanding/2010/gc/b920195p)

III. Immobilisation of thermostable β-glucosidase variants on acrylic supports for biocatalytic processes in hot water.
Samiullah Khan, Sofia Lindahl, Charlotta Turner and Eva Nordberg Karlsson. (Submitted for publication to *Journal of Molecular Catalysis B: Enzymatic)*

IV. An on-line method for pressurised hot water extraction and enzymatic hydrolysis of quercetin glucosides from onions.
Sofia Lindahl, Jiayin Liu, **Samiullah Khan**, Eva Nordberg Karlsson and Charlotta Turner. (Manuscript)

V. Comparison of the active site mutants of *Thermotoga neapolitana* β-glucosidase 1A for hydrolysis of *p*NPGlc and quecetin-3-glucosides.
Samiullah Khan, Tahir Mahmood, Tejas Kulkarni, Sofia Lindahl, Anders Sundin, Charlotta Turner, Derek Logan and Eva Nordberg Karlsson. (Manuscript)

VI. Increased production of the active soluble form of the thermostable beta-glucosidase *Tn*Bgl1A using chaperonin co-expression in *Escherichia coli*.
Samiullah Khan, S M Shariar Shakil, Gashaw Mamo and Eva Nordberg Karlsson. (Manuscript)

My contribution to the papers

Paper I. I participated in planning of the experiments, performed most of the experiments (expression and purification of proteins, collecting thermostability and kinetics data) and contributed in writing and commenting on the paper.

Paper II. I performed minor part of the experiments (expression and purification of proteins) and contributed in writing (minor) and revising of the paper.

Paper III. I participated in planning of the experiments, performed most of the experiments (immobilisation and characterisation of the immobilised enzymes with model substrate *p*NPGlc) and wrote first draft of the manuscript.

Paper IV. I performed minor part of the experiments (expression and purification of proteins, immobilisation of the enzymes and experiments for stability of immobilised enzymes) and contributed to the writing and revising of the paper.

Paper V. I participated in planning of the experiments, performed experimental work, supervised Tahir Mahmood (production and characterisation of the enzymes) and wrote a major part of the paper.

Paper VI. I participated in planning of the experiments, performed some experiments, supervised Shariar Shakil, and contributed in writing the paper.

Abbreviations

BSA	bovine serum albumin
CF	catalytic functions
DSC	differential scanning calorimetry
GHs	glycoside hydrolases
GH1	glycoside hydrolase family 1
hCBG	human β-glucosidase
HPLC	high performance liquid chromatography
NP-LC	normal phase liquid chromatography
NCF	non catalytic functions
*p*NP	para-nitrophenyl
*p*NPGlc	para-nitrophenyl-β-D-glucopyranoside
Q3	quercetin-3-glucoside
Q3,4´	quercetin-3,4´-diglucoside
Q4´	quercetin-4´-glucoside
RP-LC	reversed phase liquid chromatography
TGA	thermogravimetric analysis
UV	ultraviolet
zmGlu1	maize β-glucosidase

Popular science summary

This thesis is a compilation of six research articles and represents research on a thermostable enzyme (called β-glucosidase) that we want to use to remove sugars from natural products combined with extraction of the natural products from onion waste using hot water. The enzyme is originally from a bacterium (*Thermotoga neapolitana*) that grows in hot springs. The enzyme is used to catalyse removal of the sugars in a procedure where water is used as solvent for extraction, to in this way obtain the natural product by an environmentally friendly method. The name of the natural product used in this work is quercetin. As we are using hot water to extract the compound, an enzyme that is stable close to the boiling point of water is desirable. The gene encoding this enzyme was transferred to a host organism, *Eschericia coli* (*E. coli*) that is easy to grow and grows at lower temperature, to make it easier to produce this enzyme in large amounts.

β-glucosidases selectively removes one sugar unit from different target compounds in a reaction called hydrolysis. These target compounds can be oligosaccharides (composed of only a few sugar residues), but also more complex natural compounds where sugars are attached to another type of backbone. In this project the aim was to use the enzyme to hydrolyse one natural product that can act as antioxidant. Our aim was also to use agricultural waste as raw material, as this is desirable from an environmental perspective. As stated above, we have worked with a compound called quercetin which is mainly present in onion wastes as different quercetin-glucosides.

Quercetin is a potent antioxidant that can be used to minimize the risk of cardiovascular and Alzheimer,s diseases in human. The food industries produce a large amount of onion wastes and it is thus a resource to get the antioxidant compound quercetin. The onion wastes consist of its skins, two outer fleshy scales and roots and its production is more than 500,000 tonnes annually in the European Union.

The majority of quercetin molecules in onion wastes have sugar molecules attached to two specific positions (*e.g.* quercetin-3-glucoside, quercetin-3,4´-diglucoside and quercetin-4´-glucosides). The removal of sugar molecules from quercetin is done due to several reasons. For example, quercetin without sugar has higher antioxidising power, it is easier to quantify by analytical methods, and the yield as a uniform product increases.

Removal of sugar molecules can be done by two methods: chemically and enzymatically. In the chemical method the removal of sugar molecules is made by use of high concentration of hydrochloric acid. But this solvent has a negative impact on the environment. The quercetin molecules can also be degraded due to the harsh conditions of hydrochloric acid.

In our research project, hot water is used for the extraction of quercetin species and the beta-glucosidase enzyme is used for the removal of sugar molecules from quercetin. This method is environmentally friendly, as both water and enzymes had little bad impact on the environment as well as on the persons working with it. The advantage of using hot water is that the solubility of quercetin species also increases and then you do not need organic solvents (such as methanol that have been used traditionally). Most enzymes are however denatured and destroyed when used at high temperature as they in general are sensitive to high temperature. Therefore in our research we used a special enzyme isolated from an organism growing at high temperature (*T. neapolitana)* for bioconversion of quercetin-glucosides to quercetin. It is known that enzymes isolated from microorganisms living at high temperature keep their high temperature stability when produced in microorganisms growing at low temperature (e.g. *E. coli).* Thus the hot water extract from onions wastes can be mixed with the enzymes to get the quercetin molecules without much cooling of the extract.

As mentioned before the quercetin in onion can be found in different forms and sugars are attached at different positions on the quercetin backbone. The β-glucosidase used in this thesis can efficiently hydrolyse sugars at some positions, but not all. Quercetin-3-glucosides were difficult to hydrolyse. Therefore the wild-type enzyme was engineered (changing the amino acid at a specific place in the enzyme) and different variants of the enzyme were obtained. The wild-type and the different mutants were tested for the hydrolysis of quercetin-3-glucosides and few of the engineered enzymes obtained higher activity to this substrate compared to wild-type.

Our next goal was to be able to reuse the enzyme, by attaching it to a solid support and combine the enzymatic modification with extraction of quercetin-glucosides from onions in an online process. Therefore wild-type and one of the engineered enzymes was immobilised on acrylic support materials and some successful trials have been done by passing the yellow onion hot extract on immobilised enzymes.

In conclusion, this work shows that enzymes can be used as aids in new processes developed to be more environmentally friendly. Moreover, it is a step towards using by-products to get more value from biomass. Further work in this field is important to better use the resources on our planet and is a small step in the direction of sustainable development.

1. Introduction

Carbohydrates are essential components of life and are present in biomass in various sizes and shapes. The majority of carbohydrate materials in nature are in the form of polysaccharides, but a number of different compounds are glycosylated, e.g. glycoprotein and glycolipids, and glycosylation plays an important role in a many biological processes. The shortest soluble saccharides are used by organisms to obtain energy and sustain life. Oligosaccharides and polysaccharides are often involved in molecular recognition, protein stability, molecular signalling, and cell adhesion. They are the building blocks of plants, animals and bacteria [1].

Due to the continuous growth in human population and economic development, fossils resources such as oil are becoming limited, while demand for them is increasing. In addition, the increase in human activity often results in the release of greenhouse gases which cause severe problems with global warming. Therefore there is increasing interest in finding renewable resources, both for chemical production or energy generation in order to as a means of decrease our dependency on fossil materials.

Renewable biomass materials are a promising and abundant source of energy generation. The estimated production of biomass is 60 Gt/year and 53 Gt/year from terrestrial and marine ecosystems, respectively [2]. In addition, biomass materials include a wide variety of forestry and agricultural products and wastes. Such wastes are produced in tons every day and, if unmanaged or improperly managed, their decomposition results in significant accumulation of waste on land, and in water and air. Most agro-wastes today are not utilised, or are only used for the fertility of land. However, this waste, if properly managed, can provide a cheap and abundant source of energy. In addition, value can be added to waste biomass by processing with new technologies to extract high value compounds such as flavonoids. These high value compounds are present in low amounts but could increase the market value of waste if isolated. In addition, their extraction with careful selection of technology will not greatly change the energy value/weight of the waste material. As a result, it can still be used for biogas production, animal feed and incineration, etc.

Flavonoids are a high value compound that has attracted interest recently because they are antioxidants, thus have the ability to scavenge free radicals,

which are associated with various diseases. In addition, antioxidants are used as additives in food, cosmetics and pharmaceuticals to prolong the shelf life. Flavonoids are not always used as antioxidants but can also be used as colorants, for example anthocyanins [3]. Flavonoids can be found in onions, carrots, and apples and in the wastes from vegetable and fruit industries. In wastes such as quercetin molecules in onion waste, they are in the form of quercetin-3-glucoside, quercetin-3,4´-diglucoside and quercetin-4´-glucosides as well as aglycone and minor amounts of other glucosidic forms. Many studies have reported that the aglycone form of quercetin has higher antioxidative ability than their glycosylated counterpart. Because of their positive effects on health, they are supplied in their aglycone form by a large number of food industries [4].

The objective of this research project was to contribute to the development of alternative sustainable/green methods by combining enzymatic hydrolysis of glycosylated flavonoids with extraction methods that have only a minor effect on the composition of the remaining waste material. For this purpose the idea was to utilise a thermostable enzyme, as the extraction was using pressurised hot water as solvent instead of organic solvent [5-7]. This step should be directly followed by enzymatic hydrolysis instead of chemical hydrolysis, with HCl used as catalyst for the conversion of glucosylated compounds to aglycone [8]. The organic solvents and chemical catalysts used in the traditional method are not only harmful for the environment but lead to harsh conditions that can also result in the breakdown of the target compounds. Our newly developed method is using thermostable enzymes. Enzymes are generally denatured at the temperatures required for hot water extraction. However, if thermostable enzymes are used instead, the extraction and hydrolysis can be combined. To improve the sustainability of the method we thus replaced the organic extraction solvents with hot water, and the chemical catalyst with a biocatalyst.

Immobilisation of enzymes often increase stability, and thermostable enzymes can be further stabilised by immobilisation [9]. Thus by using immobilised enzymes, an online flow system can be built and the hot extract can be passed through the immobilised enzymes. Biocatalysts are attractive alternatives to chemical catalysts as they are renewable and biodegradable, and normally function in environmentally friendly solvents such as water [10, 11]. These properties of enzymes are hampered sometimes as they are often too labile to perform their functions. In order to make them competitive in

processing of renewable resources, their potential to operate at high temperatures is very important. During the last decade, interest in obtaining novel, robust biocatalysts from nature has increased and many new extremophilic microorganisms have been isolated. Enzymes of thermophilic origin are typically thermostable and retain their thermal stability after heterologous production in mesophiles, thus showing that these properties are intrinsic [12, 13]. Thermophilic enzymes often function optimally at or close to the growing conditions of the organisms from which they originate.

Since the process for hydrolysis of high value compounds investigated in this study requires enzymes that are operationally stable at high temperatures, *Thermotoga neapolitana* (*T. neapolitana*), β-glucosidase was investigated. It is highly thermostable and a potentially favourable candidate for this as well as other similar applications. As no wild-type enzyme is perfect or ideal and each application demands an enzyme with specific properties, protein engineering is a useful tool for modifying the properties of enzymes. In combination with 3-D structure studies and molecular biology techniques, enzymes can be rationally tailored to produce the desired properties. In this project the objectives were the following:

- to form a better understanding of *Tn*Bgl1A in the hydrolysis of glycosylated flavonoids with respect to its catalytic properties in solution and in immobilised form
- to investigate the enzyme's changes in substrate specificity and thermostability by site directed mutagenesis.

1.1. Scope of the Thesis

This thesis is a part of the research project "Sustainable Resources Technologies," where the long-term goal is to develop sustainable and environmentally friendly methods for extraction and bioconversion of glucosylated high value compounds present in agricultural by-products and industrial wastes. The specific goal in this thesis has been to develop a thermostable β-glucosidase to improve bioconversion of flavonoids present in extracts from yellow onion to their aglycone form. Quercetin glucosides from onions are glucosylated at two positions. One part of this work deals with the construction and characterisation of a set of mutants of a β-glucosidase from *T. neapolitana* and the evaluation of their effect on substrate specificity in general. Another more specific goal was to improve the catalytic efficiency in hydrolysis of quercetin-3-glucoside **(Papers I** and **V)**. The enzyme and a

mutated variant were studied in solution using a hot water extraction method developed for this purpose **(Papers I, II** and **V).** In addition, covalent immobilisation of the enzyme and of one of the developed variants **(Papers III** and **IV)** has been used to evaluate the possibility of reusing the β-glucosidase hydrolysis reactions, together with the hot water extractions. The effects of mutations on thermostability of mutant enzymes were studied by differential scanning calorimetry. Finally the expression of *Tn*Bgl1A was optimised using different strategies such as replacing rare codons with commonly used codons in *E. coli*, exporting the target protein to a periplasmic space using pel B signal petide, and using co-expression with chaperonins for proper folding of enzymes **(Paper VI)**.

2. Development of a sustainable method for extraction and bioconversion of glucosylated antioxidants

The research project "Sustainable Resources Technologies" is an interdisciplinary collaboration aiming to develop sustainable methods for extraction and bioconversion of glucosylated antioxidants (which are high value compounds) from agricultural and industrial by-products (Figure 1). Our part, the task was to develop a thermostable enzyme for the selective removal of sugars attached to antioxidants present in selected materials. The processing of these materials using sustainable and cleaner technologies not only results in the production of novel compounds but also in the reduction of greenhouse gases **(Paper II)** and the possibility of still using the material as a source of energy.

Flavonoids are an example of a high value compound found in materials such as onions, carrots and apples, as well as in wastes from vegetable and fruit industries. Quercetin is a potent flavonoid antioxidant present in onion skin, and due to its antioxidant properties, it can scavenge free radicals such as reactive oxygen species and reactive nitrogen species. It can thus minimize the risk of cardiovascular and Alzheimer's disease. These high value compounds can be extracted from the wastes without changing their overall composition significantly, and the remaining waste can still be used in anaerobic digestion for biogas production, animal feed or for composting and incineration.

The traditional and commonly used method for extraction of quercetin is solid-liquid extraction with aqueous-methanol and a high concentration of hydrochloric acid for the hydrolysis/deglycosylation. A problem with using hydrochloric acid is the harsh conditions it produces, which may lead to the breakdown of extracted compounds [8, 14, 15]. In addition, it is harmful for the environment.

With the introduction of Green Chemistry in 1990, scientists started to develop processes in line with the definition of Green Chemistry i.e. "the design of chemical products and processes to reduce or eliminate the use and generation of hazardous substances" [16]. In 1998 Anastas and Warner published the 12 principles of Green Chemistry. By following these principles, chemist and others can analyse and improve the sustainability of

an already existing process or develop a new environmentally friendly process [16].

The extraction/isolation was carried out using pressurised hot water at 120°C and 50 bar [17]. This high temperature process requires enzymes which are stable under the reaction conditions and can efficiently hydrolyse the glucosylated molecule. The enzyme presented in this thesis is from a hyperthermophile and is extremely stable, with an unfolding temperature of 101.9°C. It was thus a promising candidate for this process and other similar applications. Our goal was also to develop the thermostable enzyme from *T. neapolitana* so that it would be able to hydrolyse glycosylations at different positions on the quercetin backbone. Use of this method and the best performance mutant N221S/P342L was proven to be a more environmentally friendly alternative by Life Cycle Assessment compared to the conventional method **(Paper II)**.

The enzyme and variants also functioned when immobilised on two epoxy-activated support materials (Cryogel and Eupergit® C) at temperatures in the range 80-95°C **(Papers III** and **IV)**. Analysis using the model substrate *para*-nitrophenyl-β-D-glucopyranoside (*pNPGlc*) showed that the immobilised enzyme had even higher stability than the free enzyme. To combine the immobilised enzymes with pressurised hot water extraction, the enzyme was reused in an online mini-flow system. Application trials on onion extracts show successful recycling of the enzyme **(Paper IV)**. This method, which is in accordance with most of the principles of Green Chemistry, was successfully used for the pressurised hot water extraction and enzymatic hydrolysis of high value antioxidant compounds from low value agricultural and industrial waste and can be used for other similar applications.

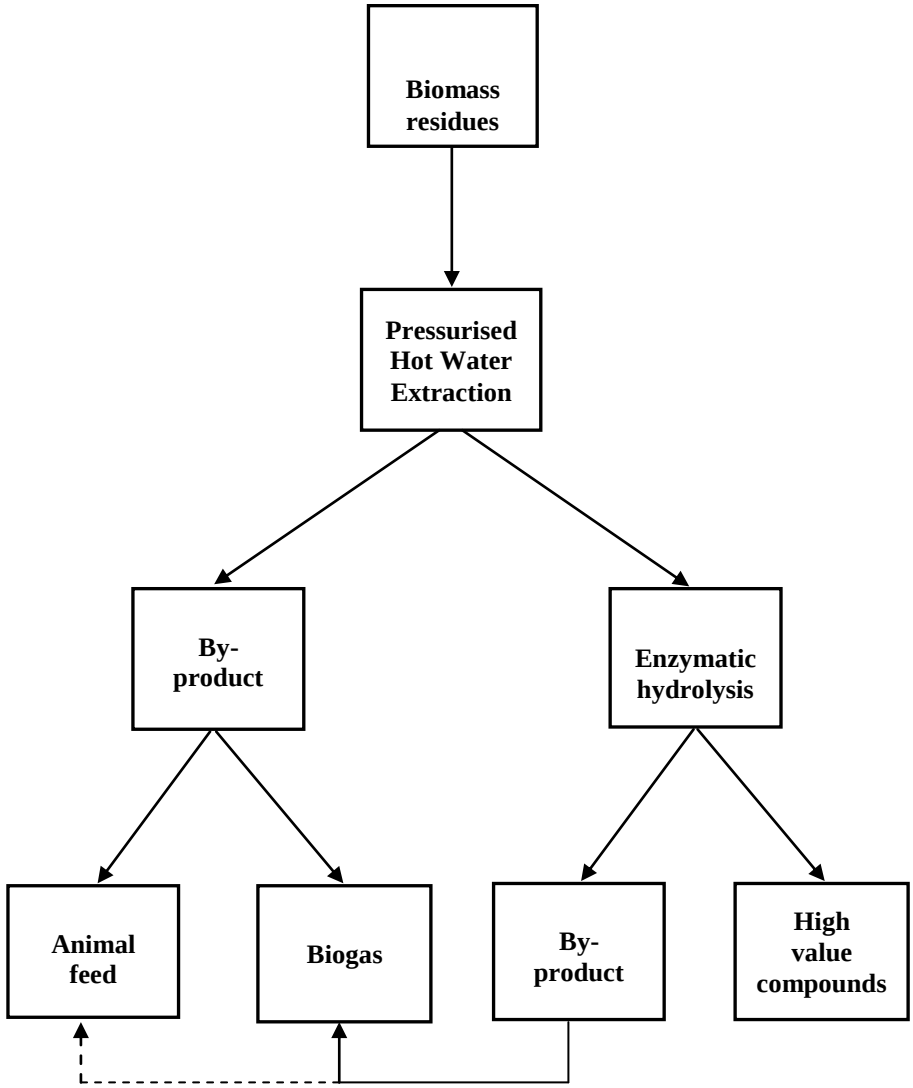


Figure 1. Schematic view of production of high value compounds from biomass waste using pressurised hot water extraction followed by enzymatic hydrolysis.

3. Sources and production of thermostable enzymes

Many environments once considered hostile and "extravagant" for any life form are now known to be oasis essential for vast flora of microorganisms known collectively as extremophiles [18-23]. An ecological system is defined as extreme if it exists in an environment where human life is not possible. Extreme conditions include temperatures, salinity, pH, pressure, dryness and desiccation inhospitable to humans, and extremophiles are special organisms with the ability to survive optimally where one or more of these conditions are extreme [19, 23]. In 1974, RD MacElroy [24] introduced the term extremophile. Extremophiles belong to all three domains of life. Two of them, namely archaea and bacteria, are prokaryotes but extremophiles can be found also in the eukaryotic domain [18, 20, 25].

From the mid 1970's and onward many novel extremophiles have been isolated, and the field is growing to the extent that now there are regular international extremophile symposia and research on the subject is published in well-respected international journals such as *Extremophiles, Microbial Life under Extreme conditions* and *Archaea*. The term extremophiles has been used to refer to organisms adapted to extreme environments, but enzymes produced by these organisms often function optimally at or close to the growing conditions of their originating organisms. A brief overview of different extremophiles is given in Table 1 below.

Table1. Extremophiles and their growth conditions adapted from [18, 26, 27].

Classification	Optimal growth condition
Hyperthermophiles	80-121°C
Thermophiles	55-80°C
Mesophiles	20-55°C
Psychrophiles	2-20°C
Halophiles	2-5 M NaCl
Acidophiles	pH < 4
Alkaliphiles	pH > 9
Piezophiles	130 Mpa

3.1. Thermophiles

Living organisms can be grouped into the following four main classes defined by temperature range: psycrophiles, mesophiles, thermophiles and hyperthermophiles [28, 29]. Thermophiles are extremophiles that grow optimally at a broad range of temperatures above 55°C [26]. This group is further subdivided into organisms optimally growing at 55-80°C, referred to as thermophiles [26]. Organisms growing above 80°C are called hyperthermophiles [30], all of which belong to the domains archaea and bacteria [31]. *Pyrolobus* and *Pyrococcus* are examples of archaea growing optimally between 103 and 113°C [30].

The upper limit of life was once thought to be 113°C (*Pyrolobus fumarii*) [32] but the archaeal strain121, the most heat tolerant thermophile, was recently isolated with the ability to survive at 121°C [33]. Hyperthermophiles, which grow at temperature 90 to 95°C, are dominated by *archaea* from the phyla Crenarchaeota, Euryarchaeota, Korarcheota and Nanoarchaeota, but a few species of bacteria from the orders *Thermotogales* and *Aquificales* are also assigned to this group (Figure 2) [30, 34]. Hyperthermophiles are believed to be the earliest evolved organisms [18] and have been isolated from a number of marine and terrestrial geothermally-heated habitats, including hot springs [35], geysers, hydrothermal vents, volcanic areas, and deep sea hydrothermal vents from all over the world, including areas in Italy, Japan, USA, New Zealand, Russia and Iceland [35-37].

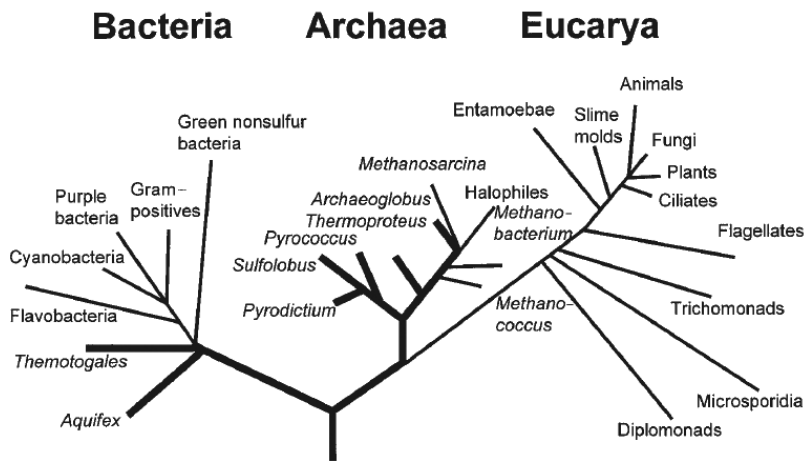


Figure 2. Universal phylogenetic tree. Hyperthermophilic genera are depicted as thick lines. Reprinted with kind permission from [38].

To survive at high temperatures thermophiles produce thermostable enzymes which are resistant to irreversible denaturation and are optimally active at temperatures up to 120°C. Enzymes from hyperthermophiles can be successfully expressed in mesophilic organisms, which make them attractive for application. Thermophile enzymes have many properties in common with mesophilic enzymes including sequence similarities up to 40-85%, superimposition of 3-D structures and the same catalytic mechanisms. However, thermophiles retain their thermostability and catalytic properties when expressed in mesophiles, showing that these properties are genetically encoded. The intrinsic and extrinsic bases of thermostabilisation are still unknown. However, many studies have been carried out to understand the molecular basis of their thermostabilty, and this subject is the focus of several reviews [13, 29, 34, 38-41]. These studies indicate that no single factor is responsible for protein stabilisation.

Suggestions regarding the structural features responsible for the intrinsic stability of thermophiles can be made if there is one unique enzyme isolated from all sorts of organisms and their 3-D structures are compared [42]. However, several intrinsic factors have been identified that have obvious roles in protein stabilisation and are worth mentioning here: the amino acid composition of proteins, ion-pair interactions on the surface of proteins, increased packing density with the hydrophobic core, disulphide bridges,

extensive hydrogen bonding, shortening of loop size, decrease in cavity size and surface hydrophobic residues. Some extrinsic factors have also been seen to be responsible for protein stabilisation such as compatible solutes, ions, stabilizing peptides, substrates and metal ions [39].

The thermal stability of a protein can be determined either from its transition state from folding to unfolding or by measuring its stability in terms of half-life ($t_{1/2}$). A protein with high melting temperature (T_m) is called a thermostable protein. The melting temperature of β-glucosidase from *T. neapolitana* (Figure 3) measured with differential scanning calorimetry (DSC) **(Paper I)** and the enzyme activity was stable at 80°C.

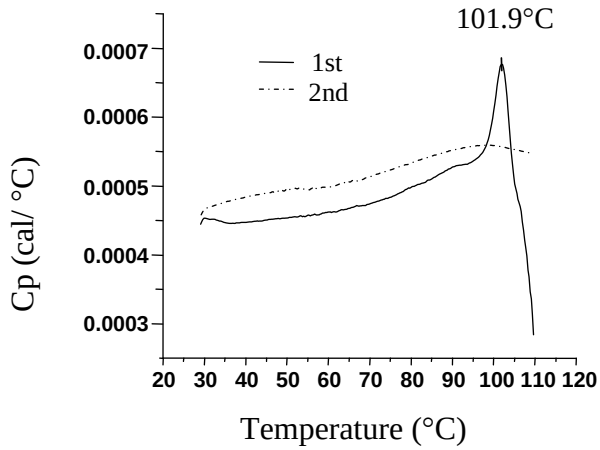


Figure 3. Thermogram of *Tn*Bgl1A (1 mg/mL in 20 mM citrate phosphate buffer, pH 5.6)

3.2. The Genus *Thermotoga*

T. neapolitana is a hyperthermophilic Gram-negative bacterium that belongs to the genus *Thermotoga.* One of its enzymes (*Tn*Bgl1A) is the main focus of this investigation [43]. It was isolated from a shallow submarine hot spring in the Bay of Naples, near Lucrino, Italy [44]. It can grow in the temperature range of 50-95°C with optimum temperature 77°C and pH 7.5 with a doubling time of 45 minutes [43, 44]. The cells from the genus *Thermotoga* are rod-shaped, non-motile and non-sporulating. The genus name "toga" is derived from the sheath-like outer structure that usually balloons over the ends [45]. It was thought to be an obligate anaerobe, but recent research has shown that it is a microaereophile using reduced level of oxygen to produce

hydrogen [46, 47]. *T. neapolitana* is highly robust, easily adapted to varying conditions and can catabolise and grow on different sugars, both mono and polysaccharides [48]. The genus *Thermotoga* is one of the most examined sources of hyperthermophilic glycosyl hydrolases [49] and encodes the largest number of glycosyl hydrolases of any bacterial or archeal genome [50]. Carbohydrate degrading enzymes of *T. neapolitana* can be expressed in *E. coli.* and due to their extreme stability at high temperature, have great potential for biotechnological applications at high temperature [51]. This enzyme reported on in this thesis was used for bioconversion of quercetin glucosides to quercetin **(papers I-V).** The enzyme studied is from *T. neapolitana* and on the basis of substrate specificity the enzyme is classified in the glycosyl hydrolase family GH1. It is thus termed *Tn*Bgl1A in this thesis. Since it acts on the β-D-1,4 linkage and is releasing a terminal glucose, it is therefore referred to as β-glucosidase [52].

3.3. Recombinant production of thermostable enzymes in *E. coli*

E. coli is an extensively used Gram-negative bacterium and, for many reasons, is a popular choice for recombinant protein production for commercial purposes. *E. coli* can grow at high cell densities. Its genetic information is well established and has been heavily researched. The bacterium provides ease of simple process scale-up and requires inexpensive and simple media for growth [53]. Moreover, several cloning and expression vectors have been developed that are commercially available, and that can achieve high levels of protein production. These vectors generally contain a selection marker, cloning cassette with selective restriction sites and an origin of replication. Purification tags (e.g. His tag **Paper I**) can be included to facilitate the production of the target product. Recombinant production in *E. coli* occurs inside the cells in the cytoplasm or can be directed to the periplasmic space by cloning the gene in frame with a sequence encoding a leader peptide [54].

Successful protein production for therapeutic purposes and biochemical or structural studies depends on the expression, purification and solubility of proteins [55]. Expression is well established and techniques for purification have greatly improved during the last years. Purification of protein of thermophilic origin expressed in *E. coli* has been further simplified recently and partial purification can be achieved by heat treatment, as most of the host

protein can be precipitated and removed [56, 57]. However, production of soluble protein in *E. coli* for purification is still an increasing problem in the field.

One problem with heterologous protein expression in *E. coli* is the accumulation of over-expressed protein as inclusion bodies in the cells. This problem can be overcome sometimes by exchange of the rarely used codon to more commonly used ones in *E. coli*. The degeneracy of the genetic code allows rewriting of the open reading frame, and the same protein can be produced by changing up to one-third of the nucleotide. To improve our target protein production, the same strategy of degeneracy of the genetic code was used and a bgla gene was constructed synthetically (with codons adapted to *E. coli* K12) and cloned **(Paper VI)**, in the same way as described by Turner et al [17]. The expression of this synthetic gene did not increase the solubility of the protein significantly, and the synthetic gene was expressed in *E. coli* with almost the same expression level as for the wild-type gene **(Paper VI)**.

The codon optimisation of the synthetic gene did not solve the problem of protein solubility and therefore another strategy was applied. The synthetic gene was cloned so as to translocate the protein into periplasmic space by including a leader peptide in the cloning design at the N-terminal end (Figure 4). Periplasmic expression may promote proper folding of protein into tertiary structure, providing protection against intracellular degradation by proteases as well as easy purification from other cellular components. Moreover, most of the proteins are soluble, as it is not a favourable place for the formation of insoluble inclusion bodies [58]. As was expected the target protein was exported into the periplasm and we obtained a slightly higher level of soluble protein. However, as the improvement in solubility was negligible **(Paper VI)**, co-expression with cytoplasmic folding modulators was the next step.

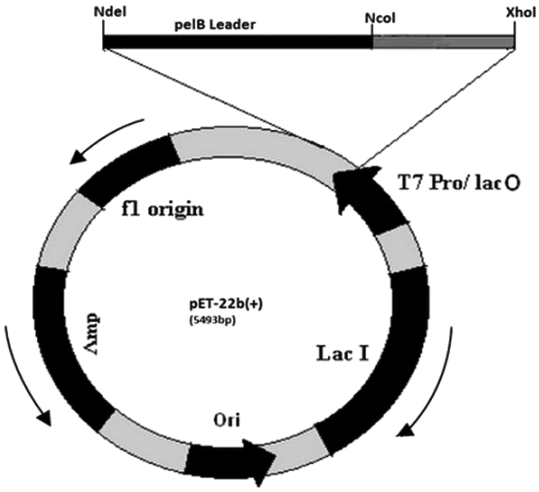


Figure 4. Schematic diagram of the construct used for heterologous expression of a synthetic gene with pleB leader sequence and His tag, cloned under the control of a T7 lac promoter in pET-22b (+) vector

Due to the high concentration of macromolecules in cytoplasmic environments, overexpressed recombinant protein folding is often assisted by molecular chaperones [53]. Chaperones do not change the protein folding pathway but increase recovery yield by preventing aggregation of newly synthesised non-native polypeptide chains and assembles subunits from aggregation [53]. The most commonly used chaperones in *E. coli* are DnaK, DnaJ, GrpE, GroEL, GroES. They have a better cooperative role if the chaperones are used in the form of teams such as DnaK-DnaJ.GrpE and GroEL-GroES, and the refolding efficiency in vivo can be increased [59, 60]. Nishihara, et al. have constructed five different plasmids containing genes for these chaperone teams [61, 62], (see Table 2).

It has been shown that co-expression of a target protein with one of these chaperone teams increases recovery of expressed proteins in the soluble fraction [61, 62]. Our results show that co-expression of protein with these cytoplasmic modulators increased the soluble fraction in all cases and was maximum using plasmid pG-KJE8, encoding both chaperonin teams where we obtained a 6.4-fold higher activity (U/ mL) compared to the wild-type **(paper VI)**.

Table 2. Molecular Chaperones used for co-expression with *Tn*Bgl1A for solubility enhancement.

Plasmid	Chaperones	Promoter	Inducer	Selection Marker	Origin
pG-KJE8	DnaK-Dnaj-GrpE GroES-GroEL	*araB* *Pzt-1*	L-Arabinose Tetracycline	Cm	pACYC
pGro7	GroES-GroEL	*araB*	L-Arabinose	Cm	pACYC
pKJE7	DnaK-Dnaj-GrpE	*araB*	L-Arabinose	Cm	pACYC
pG-Tf2	GroES-GroEL, tig	*Pzt-1*	Tetracycline	Cm	pACYC
pTf16	tig	*araB*	L-Arabinose	Cm	pACYC

3.4. Application of thermostable enzymes

People started to consider the use of biocatalytic processes to replace chemical catalysis after the first use of industrial enzymes in detergent applications in 1915 [63]. Today enzymes are used not only in detergents but also in textile, paper and pulp, food and feed and leather industries. In many industries enzymes are used for sustainable development due to their environmental advantages, for example for improving the use of raw materials and decreasing the amount of wastes products. Enzymes from extremophiles are in increasing demand because of their unique properties and are expected to fill a part of the gap between biological and chemical processes [22].

Thermophiles and hyperthermophiles are valuable sources of many industrial enzymes. Thermostable enzymes offer many advantages in biotechnological processes and their stability make them feasible for harsh industrial processes as they are normally resistant to chemical denaturants. Enzymatic reactions at elevated temperatures can result in higher reaction rates, which in turn can increase process yield because of increased solubility and bioavailability of substrates and products, decreased viscosity, increased diffusion rate and reduced risk of contamination [39]. The above-mentioned properties make thermostable enzymes a hot topic for study and their use has important implications for many biotechnological applications [13, 64]. Some examples

of applications are enzymatic digestion of starch, food colouring and brewing and detergent industries. Many of these industrial enzymes are hydrolases and those acting on carbohydrates are called glycoside hydrolases (GH). However, few of them are used in the hydrolysis of flavonoids and isoflavonoids glycosides [17, 65].

4. Glycoside hydrolases and β-glucosidases

Carbohydrates are abundant in nature and are linked to each other by glycosidic bonds, one of the most stable bonds in naturally occurring biopolymers [66]. Glycoside bond formation occurs by elimination of water molecules as a result of complex metabolic pathways within organisms [67], while the cleavage of glycoside bonds is as a result of addition of one water molecule with the help of glycoside hydrolysing enzymes [1]. Many organisms utilise carbohydrates for energy uptake and as a carbon source, and the selective hydrolysis of glycosidic bonds is therefore crucial [1]. Because of the diversity and complexity of carbohydrates, there is a large assortment of glycoside hydrolases present in nature. These enzymes are highly specific in nature and can catalyse either alpha or beta linkages. They can hydrolyse polysaccharides, oligosaccharides at the end (exo-acting) or in the middle (endo-acting) by breaking the glycosidic bond.

Glycoside hydrolases (GHs) are characterised by many protein folds with known 3-D structures. However, the overall active site topologies are classified into three major classes: pocket, cleft and tunnel topologies (Figure 5). Pocket topologies, also termed crater topologies, typically recognize the non-reducing end of monosaccharides such as β-glucosidases (which are studied in **Papers I-VI**), β-gluconases, and exopolysaccharides such as glucoamylase and β-amylase [1]. The enzymes with cleft topology are endo-acting polysaccharides and are able to accommodate several sugar units. Examples are endocellulases, lysozymes and chitinases [1]. The third type of topology is derived from the cleft one, which firmly holds the polysaccharides while the product is released and a substrate such as in glucosidases is processively cleaved [68]. In this thesis the focus is on β-specific broad specificity glucosidase with pocket topology from *T. neapolitana* **(Papers I-VI)**.

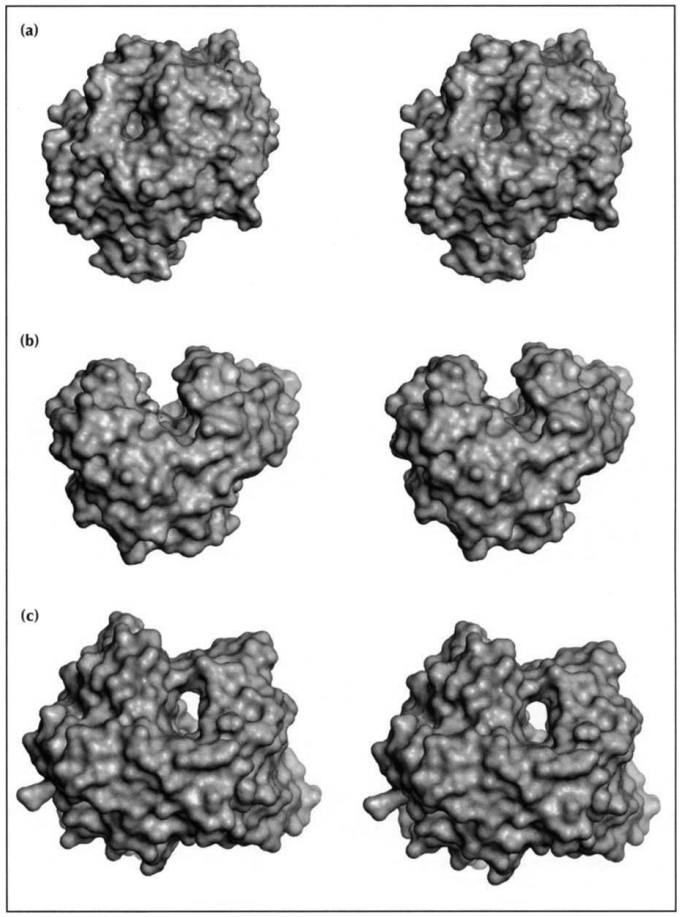


Figure 5. The three types of active site topologies found in glycosyl hydrolases. (a) The pocket (glucoamylase from *A. awamori*). (b) The cleft (endoglucanase E2 from *T. fusca*). (c) The tunnel (cellobiohydrolase II from *T. reesei*). Reprinted with kind permission from [1].

4.1. Beta-Glucosidases and their substrates

Beta-glucosidases (β-D-glucopyranoside glucohydrolases) are enzymes that hydrolyze glycosidic bonds between two glucose residues or between carbohydrate and non-carbohydrate moieties (aglycones) that can be alkyl or aryl residues and that release non-reducing terminal glucosyl residues. These enzymes are widely distributed in nature and found in all domains of living organisms, starting from bacteria to highly evolved mammals. Due to their key role in many biological processes and biotechnological applications, β-glucosidases have been the subject of much research. β-glucosidases belong

to the families GH1, GH3, GH5, GH9, GH30 and GH116 of glycosyl hydrolases (GH). Enzymes in GH1, GH5 and GH30 have a $(\beta / \alpha)_8$ barrel tertiary structure and are classified as one GH clan A. In insects and plants, β-glucosidases release cyanides from cyano-glucoside precursors and are involved in defence mechanisms [69]. Human deficiency of lysosomal acid β-glucosidase results in the accumulation of glycosylceramides in the lysosome of reticulo-endothelial cells, which is normally hydrolysed by this enzyme. As a result organs such as the spleen, liver and lymph nodes enlarge and therapeutic doses of purified β-glucosidases are recommended [70].

Extensive research has been conducted on microbial β-glucosidases. *Agrobacterium fecalis* β-glucosidases was the earliest enzyme characterised. The gene encoding this enzyme was cloned and expressed in *E. coli* [71]. Some examples of other β-glucosidases cloned and characterised from bacteria are *Bacillus polymyxa* bglA [72], *Clostridium thermocellum* bglA [73], *Thermotoga maritima* (*T. maritima*) bglA [74] and *T. neapolitana* bglA, (the enzyme studied in **Papers I-VI**). Today most of the research on microbial β-glucosidases focuses on their biotechnological application rather than their endogenous role in microrganisms [75]. In bacteria and fungi they work in agreement with cellulases and exo-β-glucanases, thus catalysing the short chain oligosaccharides and cellobiose. Their efficiency drops as the glucose chain length increases [76, 77]. Cellobiose is an inhibitor of exo- and endo-glucanases and thus needs to be removed for significant hydrolysis of cellulose, which is facilitated by supplementation of β-glucosidases, in particular by those with high tolerance of glucose [69, 78].

β-glucosidases releases aromatic compounds from the glucosidic precursors present in fruits and fermentation products, and thus play a very important role in flavour industries [79]. These enzymes are also added to food and beverages to improve their quality and nutritional value by release of antioxidants, vitamins and other useful compounds from their glucoside form [80]. Quercetin is a potent antioxidant present in onion as quercetin-3-glucoside, quercetin-3,4´-diglucoside and quercetin-4´-glucosides. It can be released from its glucoside by β-glucosidase treatment. Many studies support the fact that the aglycone form of many vitamins and antioxidants present in plants are absorbed quickly in the small intestines, which thereby increases their bioavailability. Therefore, it appears that animal feed treatment with crude β-glucosidases may provide important advantages [69]. However this

is controversial and there are studies supporting the relatively fast absorption of the glycoside form of flavonoid antioxidants [81].

4.2. Classification of β-glucosidases

There is no single method present to classify β-glucosidases. In principle they are either classified according to specificity or sequence similarity. They are assigned to three classes on the basis of substrate specificity.

(a) Aryl β-glucosidases that act on aryl glucosides
(b) Cellobioases that break down cellobiose into glucose units
(c) Broad substrate specificity β-glucosidases that act on large numbers of substrates

Most of the characterised enzymes belong to class (c).

One of the most simple and frequently used classification methods has been the use of the EC number system for a given enzyme, upon the approval of the International Union of Biochemistry and Molecular Biology (IUBMB). In this system glycoside hydrolase was given the code E.C 3.2.1.x, where x represents the substrate specificity or cleavage type, e.g. β-glucosidase was given the code E.C 3.2.1.21 [82]. The traditional system fails because the corresponding EC number and name give us very little information about the mode of action or structure. In addition, it does not deal with the problems of enzymes having multiple specificity and varying relationships with other glycoside hydrolases.

β-glucosidases can also be classified on the basis of the stereochemistry of hydrolysis and can be assigned to the classes of retaining or inverting enzymes, depending on the anomeric configuration, five or six membered sugar ring and axial or equatorial leaving group. This method is not globally accepted for β-glucosidases classification, although stereochemistry is still of high value in hydrolysis [83].

In 1991 a new system of classification was proposed by Henrissat, where glycosyl hydrolases were assigned to GH families based on their amino acid sequence composition; the system has been updated over time [52, 84]. This new system has been widely accepted in the scientific community and has proven to be immensely versatile. Based on the principle of structure and sequence associations, much useful structural and stereochemical information can be derived from the amino acid sequence. This system has many advantages as each family shares the same 3-D structure, active site residues and stereochemical mechanisms. This facilitates prediction about new

structures of enzymes and is a powerful analytical tool for using homology modelling [85]. For example, the β-glucosidase from *T. neapolitana* presented in this thesis has been assigned to GH1 and its homology model was built by us **(Paper I)** on the available crystal structure from *T. maritima*. Enzymes are grouped into glycoside hydrolase families according to this classification system and to date 129 families have been identified [75]. Related families of common structure and evolutionary origin are grouped together into clans with similar folds, molecular mechanisms and catalytic machinery [85].

4.3. Flavonoids as substrates for β-glucosidases

As mentioned above in the classification section, β-glucosidases have broad substrate specificity and are active on different types of substrates including cellulose, β-glucan and flavonoids. In this thesis the activity and selectivity of *Tn*Bgl1A towards flavonoid substrates has been explored.

4.3.1. Flavonoids

Flavonoids are a large group of natural compounds present in fruits and vegetables. They are composed of two benzene rings linked by a heterocyclic pyrene ring (Figure 6A). Quercetin, isorhamnetin, kaempferol, hesperitin and naringenin are examples of flavonoids and, because of their positive effects on health, they are supplied as food supplements in their aglycone form by a large number of food industries [4]. Flavonoids have a common carbon skeleton and are subdivided into flavonols, flavones, isoflavones, flavanones, anthocyanidins and flavanols on the basis of group attached to the carbon skeleton. Flavonoids have the ability to scavenge free radicals, which are associated with various diseases and have therefore attracted interest recently for their function as antioxidants [86].

The health problem of free radical generation is not only a result of chemical and radiation injury but is normally generated during respiration, other cellular activities and microbial defence. Free radicals are highly reactive chemical entities because they are unstable due to their lack of a stable number of outer shell electrons. By reacting with proteins, lipids and nucleic acids, they can cause carcinogenesis and chronic degenerative processes in the human body [87].

Flavonoids contain numerous double bond and hydroxyl groups that can donate electrons to stabilise free radicals and can thus act as anti-oxidants.

The electron scavenging properties of flavonoids are thus associated with their structure, thus giving these compounds properties that defend against oxidative stresses. In doing so, they reduce the risk of heart diseases, slow down or prevent the development of cancer and slow the aging process in cells responsible for degenerative diseases. To avoid oxidative stress during the physiological imbalance of reactive oxygen species, it is important to have flavonoids in our daily diet [88, 89].

Over 6500 flavonoids have been identified [90]. They are found in plants, fruits, vegetables, ferns, barks, stems, roots, tea and wine [91, 92]. They are secondary metabolites and their role in plants is to protect against UV-radiation diseases, infections and insect invasion [90, 92]. In plants, flavonoids are mainly present as β-glycosides and are bound to at least one sugar such as glucose, rhamnose, xylose or galactose. Glucose is the most commonly bound sugar. The most favoured glucosylation site is the 3-postion, and the least preferred one is the 7-position [93]. Glycosylation can occur with different types of sugars, for example, with quercetin-3,4´-diglucoside or kaempferol-3-robinoside-7-rhamnoside. Glycosides have higher water solubility than its aglycone form; however, glycosylation can decrease its antioxidative potency [90].

Quercetin is one of the most common flavonoids and is found in many different plants. Onion is one important source, where it occurs at highest concentrations [94]. In dried red onion it accounts for up to 2.1% w/w [95]. According to one study (Hollman and Arts, 2000), the amount of quercetin in onion is around 300 mg/kg of onion, 100 mg/kg of broccoli, 50 mg/kg of apples, 40 mg/kg of blackcurrants and 30 mg/kg of tea. The onion by-products only in Sweden produced 1000-5000 ton each year [Unpublished data]. The average daily intake is from 15 to 40 mg/day while its increased intake decreases the risk of cardiovascular diseases [92]. However, the content of quercetin is not fixed and varies a great deal from sample to sample and is different in different parts and varieties of the plants, e.g. red onion has higher amounts than yellow onion, and onion skin has higher amounts than the internal bulb layers. In this thesis the study is based mainly on quercetin from yellow onion **(Papers I-V)**.

Quercetin molecules in onion are in the form of quercetin-3,4´-diglucoside, quercetin-3-glucoside and quercetin-4´-glucoside as well as aglycone (Figure 6B-E) and other glucosides. As mentioned above, position-3 is a favoured glycosylation site for many glycosylated flavonoids but on the other hand this

position is the most difficult one for enzymatic hydrolysis. Many studies have reported that the aglycone form of quercetin has higher antioxidative ability than its glycosylated counterpart. Hence it would be interesting to obtain pure quercetin, but to do so it must be hydrolysed. *Tn*Bgl1A is a promising enzyme for the hydrolysis of quercetin glucosides and was thus studied in the present investigation **(Papers I-V)**. This enzyme can be used in many applications and was successfully used for deglycosylation of other interesting molecules such as kaempferol-3-glucoside, Isorhamntin-3-glucoside, daidzein and genestein which are present in different sources (see **Paper V**).

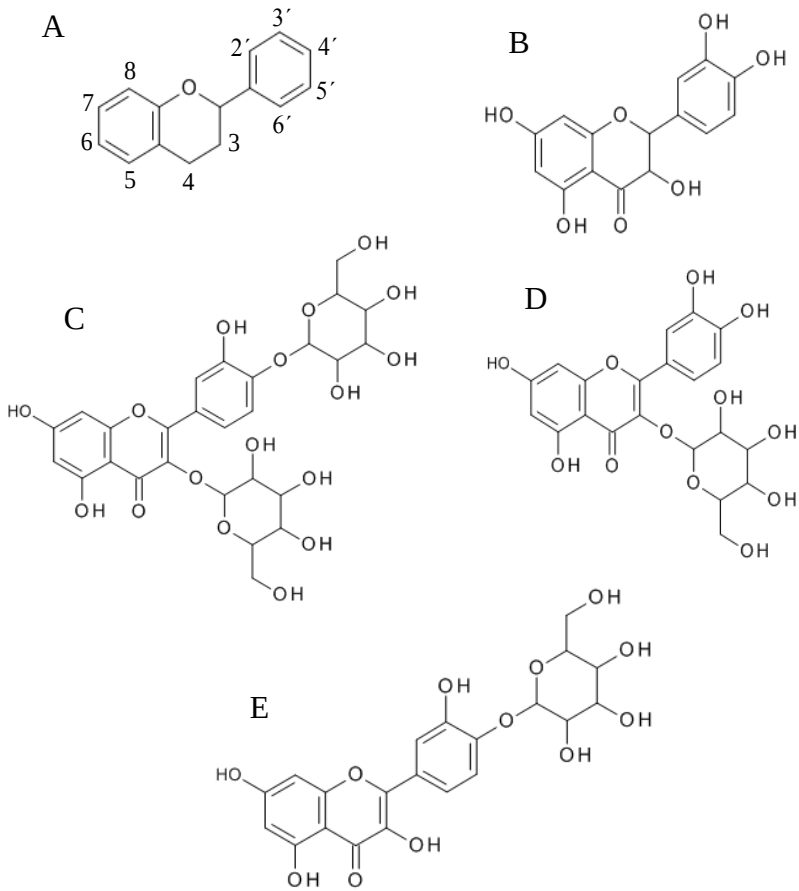


Figure 6. Structure of (A) Flavonoid backbone, (B) quercetin (C) quercetin 3,4´-diglucoside (D) quercetin-3-glucoside and (E) quercetin-4´-glucoside

4.4. Methods for estimating β-glucosidase activity

Several methods both sensitive and easy to use are available for the activity measurement of β-glucosidase depending on the substrate used.

4.4.1. Spectrophotometric Assay

Spectrophotometric methods have been in use for many years have many advantages over chemical methods and chromatographic separation and has been used in many studies [96-100]. The instrument used for spectrophotometric measurement, the spectrophotometer, consists of a light source, a monochromator and a detector. This method is rapid and allows direct measurement of the sample. It has the advantage that an intact sample can be measured and is thus required in many on-line measurements. This method is environmentally friendly and no extra chemicals are used to measure the sample. The method allows measurement on several quality parameters. It is well suited for the activity measurement of hydrolytic enzymes when a chromogenic substrate such as *p*NPGlc is used. *p*NPGlc releases para-nitrophenyl (*p*Np) upon hydrolysis which is measured at 405 nm. Light with a wavelength close to 400 nm is violet and appears yellow to the human eye; thus yellow can be observed in the cuvette after hydrolysis. The drawback with this method is that it fails to distinguish between substrate and product when they both absorb light in the same wavelength, for example, quercetin and quercetin-3-glucoside. For the accurate quantification of product when a natural substrate such as quercetin glucoside is used, the alternative is HPLC.

4.4.2. HPLC

High Performance Liquid Chromatography is a powerful separation technique introduced in the late 1960s. It has the ability to resolve mixtures with similar analytes. The sample mixture is distributed between a stationary and mobile phase. The stationary phase is either solid or porous surface active materials in small-particle form, and the mobile phase is always liquid. The liquid transports the eluted compounds to the detector, where it is recorded as bell-shaped curves. The recorded signals are known as peaks and provide qualitative and quantitative information about the sample [101]. HPLC is one of the most popular and widely used separation techniques today due to its reproducibility, high accuracy in quantitative determinations,

and suitability for separating thermally labile and non-volatile species [102]. In HPLC the commonly used detector is the UV-Vis detector. It is reliable, robust and suitable if the analyte absorbs in the UV-Vis region 190-800 nm. Separation in HPLC is based on polarity and the distribution of the analyte between the stationary phase and the mobile phase.

There two types of HPLC: normal phase HPLC (NP-HPLC) and reversed phase HPLC (RP-HPLC). RP-HPLC is more frequently used and it is a more environmentally friendly alternative since less toxic organic solvents are used in the mobile phase. RP-HPLC is used for separation of polyphenols and the most commonly used stationary phase is C18. The C18 chains are usually bound into silanol groups on silica particles. Throughout the **papers I-V** a RP-HPLC was used with a Zorbax SB-C18 column (2.1x100 mm, 3.5 micron) from Agilent Technologies. With both types of HPLC methods, the response of either a single or several components is measured and quantified. The area and height of the peaks are proportional to the amount of compounds injected, and the calibration curve can be derived from peak areas or peak heights for various solutions of standard concentration. The concentration of unknown samples can be determined from the calibration curve.

4.5. Glycoside Hydrolase Family 1

The β-glucosidase from *T. neapolitana*, used in this thesis is classified by sequence similarity under the glycoside hydrolase family 1 (GH1). Family 1 glycosidases are a very relevant group of enzymes because of the diverse biological roles they play in nature and their occurrence in all domains of living organisms [103]. GH1 is one of the most characterised enzyme families and holds enzymes with the greatest number of different substrate specificities such as β-glucosidase, β-galactosidase, β-mannosidase, β-glucuronidase, β-D-fucosidase, phlorizin hydrolase, exo-β-1,4-glucanase, 6-phospho-β-galactosidase, 6-phospho-β-glucosidase, strictosidine β-glucosidase, lactase, amygdalin β-glucosidas, prunasin β-glucosidase, raucaffricine β-glucosidase, thioglucosidase, β-primeverosidase, isoflavonoid 7-*O*-β-apiosyl β-glucosidase, hydroxyisobutyrate hydrolase and β-glycosidase [75]. This family currently contains 3329 enzymes, among which 225 have been characterised [75]. In GH1 two glutamic acids are involved in the catalysis in the active site. One acts as nucleophile and the second as catalytic acid/base [75, 104]. In plant myrosinases the catalytic acid/base is

absent but it is replaced by a cofactor ascorbate [105]. All GH1 β-glycosidases share the common $(\beta / \alpha)_8$ barrel tertiary structure [75]. The active site of β-glycosidases is in the β-barrel towards the C-terminal portion and is surrounded by loops which link the α-helices to the β-strands [106]. The catalytic domain varies from 440 to 550 residues because of the variability in the loops at the C-terminal ends of the β-strands of the $(\beta / \alpha)_8$ barrel. Many quaternary structures of monomeric enzymes, dimers, tetramers, hexamers, octamers, and large aggregates are formed [103]. GH1 members have very similar active sites, which indicates that most of the interactions occur in analogous residues [107].

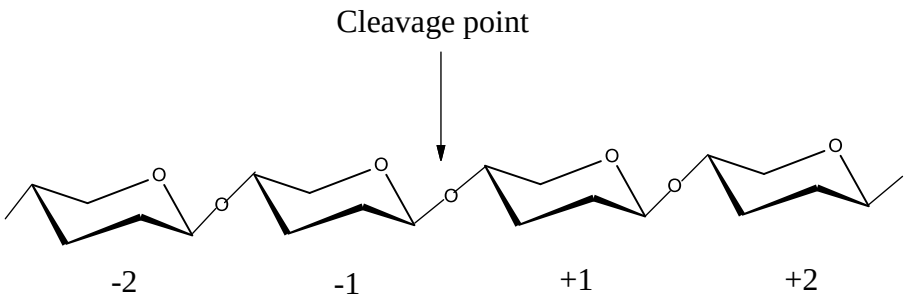


Figure 7. Nomenclature of sugar binding sites. The cleavage occurs between subsites -1 and +1.

The sugar binding subsites indicate the position of subsite relative to the cleavage point in glycoside hydrolases (Figure 7). The subsites on the non-reducing end are denominated -*n* and on reducing end +*n*, where *n* is an integer. In GH1 monosaccharides of the substrate bind to the -1 subsite, also called glycone subsite, whereas the aglycone binding region accommodates the remaining part of the substrate and may consist of several subsites, e.g. +1,+2,+3 [108]. The cleavage point is indicated by an arrow in (Figure 7) between +1 (glycone subsite) and -1 (aglycone binding region) [108]. In GH1 the numbers of subsites differ in the aglycone region and the *Streptomyces* species have the single subsite +1 [109]. *Hordeum vulgare* β-glycosidase has a large number of subsites, with four subsites (+1,+2,+3 and +4) [110], and all other β-glycosides have an intermediate number of subsites within this range [111]. The catalytic and binding site residues are generally well conserved in GH1 enzymes. There are still some differences which contribute to the functionality of enzymes and a few of them have been studied in **Paper**

I. The gate to the active site in GH1 is formed by four extended loops [112]. These loops (termed loop A-D) have been shown to be responsible for the overall shape of the aglycone binding pocket. Differences in the conformation of one or more of these loops would likely change the overall shape of the entrance to the active site [112].

In the present research, a molecular model of *Tn*Bgl1A was built (Figure 8A) **(Paper I)** based on the 3-D structure of β-glucosidase BglA from *T. maritima* (*Tm*GH1) [113, 114]. *Tm*GH1 is the closest homologoue with highest similarity: 90% identity, and 97% positives. The crystal structure of *Tn*Bgl1A was obtained recently by Kulkarni et al (Unpublished results) and was compared with, the *Tn*Bgl1A model, used in this work. The model was used as a template and aligned with the crystal structure. The model fitted very well with the crystal structure (Figure 8B).

The overall structure presented in the *Tn*Bgl1A model is the typical GH1 structure with $(ß/\alpha)_8$ barrel fold. The active site is a deep channel (18-21 Å) that narrows at the bottom, with a wider cleft at the entrance. Two conservative motifs (TLN**E**P and IT**E**NG) are situated opposite to each other inside the active site (Figure 8C), more precisely, they are at the ends of strands ß4 and ß7 containing the acid/base (E164) and the nucleophile (E349) residues (Figure 8D). The subsite -1 included conservative residues forming hydrogen bonds (Q18, H119, N163, E403) and hydrophobic interactions (W396, W404). It is the sugar/glycone recognition site and may be very specific for only one sugar or may recognize different sugar molecules and subtle change in this site can totally change which sugars are accommodated. The following sites are the +1 and +2 subsites which interact with aglycone and determine this specificity. Residues in these sites are less conserved in GH1 and depend on aglycone specificity. Mutation in these sites may also lead to changes in catalytic efficiency as has been seen in *Tn*Bgl1A **(papers I and V)**. The *Tn*Bgl1A aglycone +1 subsite is formed by the hydrophilic residues (N171, H178, N220) and mainly by aromatic and hydrophobic residues (W33, F36, W120, V167, V171, W322, A405, and F412). Inspection of residues surrounding the +2 subsite showed a non-conserved region at the ”floor” of the active site, more precisely at the end of β-strand 5 (F219, N220, N221, G222, Y223, F224). The non-conserved region consists of five

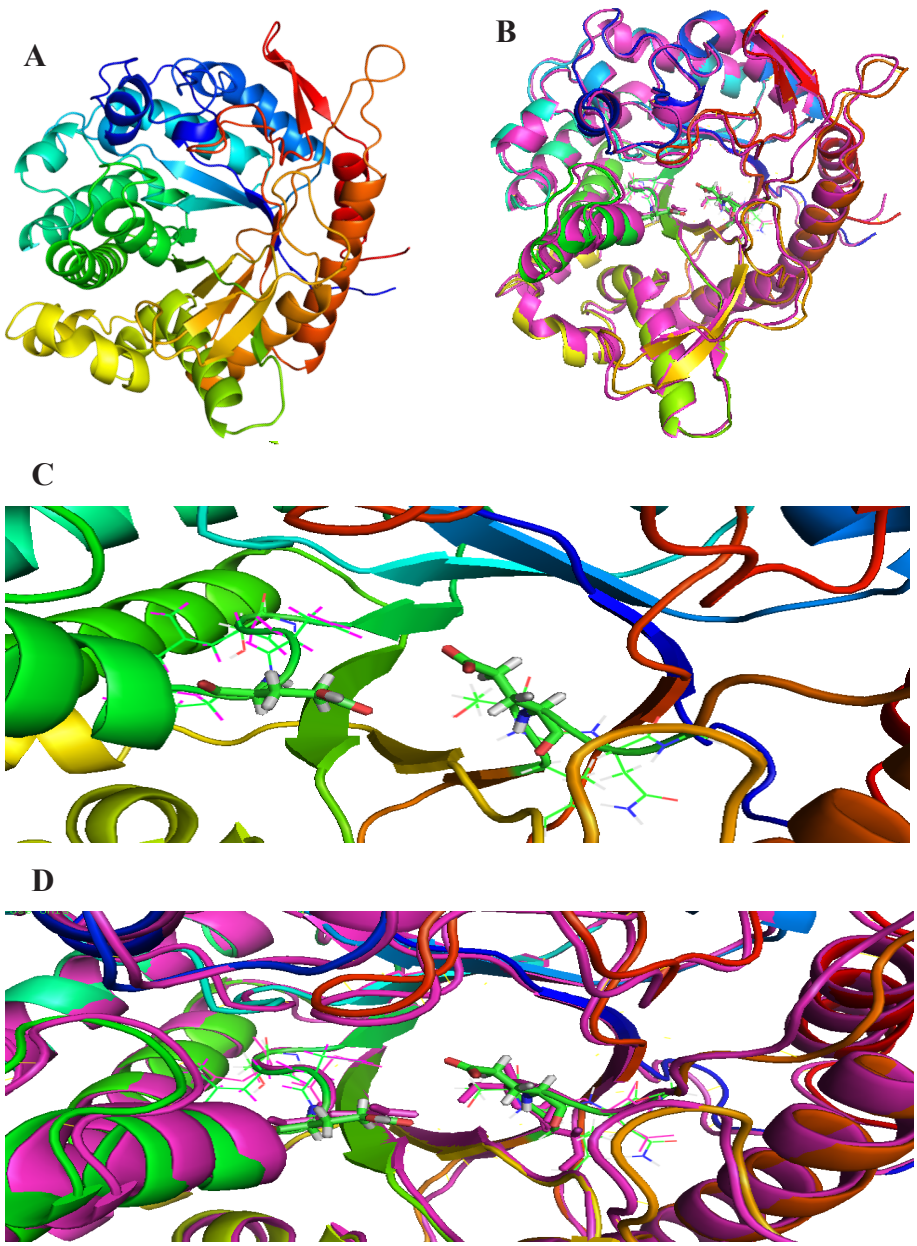


Figure 8. Structure of *Tn*Bgl1A. (A) *Tn*Bgl1A model, (B) *Tn*Bgl1A model and crystal structure aligned, (C) close view of the active site, the two conservative motifs (TLN**E**P and IT**E**NG) are shown as lines and (D) close view of the active site of the model with crystal structure aligned. The acid/base (E164) and nucleophile (E349) are shown as sticks.

more hydrophobic interactions, of which two were aromatic residues (W166, I170, V171, V176, and F310) **(Paper I)**.

4.6. Catalytic mechanism

Hydrolysis of glycosidic bonds occurs with two distinct mechanisms: either with inversion or with retention of chirality of the anomeric carbon [83, 115]. In both mechanisms a pair of carboxylic acid takes part in the active site. In inverting enzymes, one residue acts as a catalytic acid and the other as a catalytic base and both are placed at a distance of 10 Å to accommodate both substrate and a water molecule. In retaining enzymes, one residue acts as nucleophile while the other as an acid/base and both are only 5 Å apart. β-glucosidases from GH9 use an inverting mechanism, where a carboxylic amino acid acts a base and activates a water molecule, which makes a direct nucleophilic attack on anomeric carbon at the centre of the glycosidic bond. On the other hand, another carboxylic acid residue acts as a general acid and allows the departure of the leaving group by breaking the glycosidic linkage. The overall configuration of the anomeric centre is inverted (see Figure 9a) and results in an α-sugar from β-linkage and vice versa [69] .

β-glucosidases from GH1, GH3, GH30 and GH116 are retaining β-glucosidases and hydrolyse their substrate with retention of the anomeric centre. The reaction occurs in two steps via double displacement, as seen in Figure 9b [1, 116]. In the first step the carboxylic residue functions as a general acid and donates a proton to the leaving group, while the second carboxylic acid residue makes a nucleophilic attack on the anomeric carbon. Here an enzyme glycosyl intermediate is formed, with anomeric carbon forming a covalent bond with a nucleophilic residue with opposite configuration. In the second step, the deprotonated acid/base that acted as a general acid earlier now acts as a base and activates the incoming water molecule. In the second nucleophilic substitution, water oxygen attacks at the anomeric centre and releases sugar from the covalent bond to the enzyme.

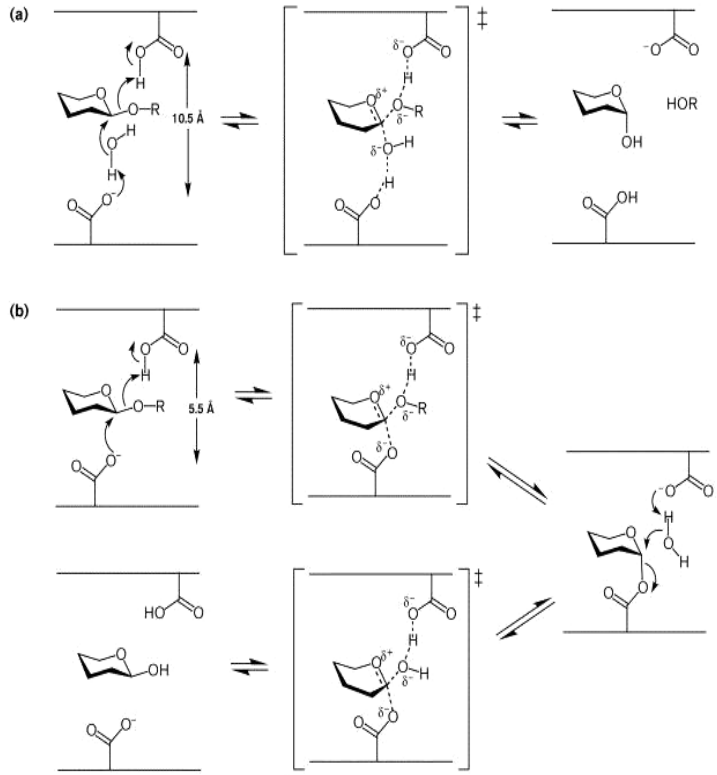


Figure 9. General glycosidase mechanisms for **(a)** an inverting β-glycosidase and **(b)** a retaining β-glycosidase proceeding through an intermediate with a $^4C_1$ conformation. Reprinted with permission from [117].

4.7. Substrate specificity in glycoside hydrolase family 1

The catalytic domain regions in GH1 are well conserved but the enzymes in the family have varying substrate specificities. Some of these enzymes are very specific for only one sugar and one aglycone and, on the other hand, others have a broad range of specificities for glycone or aglycone or for both. This makes GH1 an interesting model for studies of the relationship between structure and activity of the enzymes [118]. In general they are more specific for the recognition of the glycone and only the D-form of the glycone is recognised, whereas they vary with respect to aglycone identities [107].

In spite of their well-conserved catalytic residues, little information is available about the recognition and interactions of these enzymes with the substrates which define their functional diversity. To date 34 enzymes have been identified and characterisation of their variants are available with 3-D

structures, among them four from archaea, thirteen from bacteria, two from animals, one from fungus and fourteen from plants [75]. In GH1 the same β-glycosidase may hydrolyse substrates presenting different glycones such as glucose, galactose, fucose, mannose, xylose, 6-phospho-glucose and 6-phospho-galactose, but their catalytic activity is different for different glycones [111]. Minor changes in glycone structure result in a large variation in catalytic activity, as seen in *Spodoptera frugiperda* β-glycosidase, which hydrolyses fucosides 40 times faster than galactosides, which have the 6-OH group missing in fucosides [119]. The crystal structures of many β-glycosidases with ligand in -1 subsite provide very useful information about the interations in the -1subsite [106]. The specificity of β-glycosidases is also dependent on the aglycone moieties. Some glycosidases are active only on one substrate while others have broad aglycone specificity [103, 120]. In plants, the specificities of β-glucosidases aglycone differ a great deal. The enzymes such as ZmGlu1from maize and sorghum dhurrinase (SbDhr1) have 70% sequence similarity but different substrate specificity because of different aglycones in their natural substrate [69]. GH1 enzymes bind to sugar with the same conserved residues but it is still very difficult to assign these enzymes to a class such as galatosidases, β-glucosidases etc [69].

The β-glucosidase (*Tn*Bgl1A) from the hyperthermophile *T. neapolitana* has broad specificity towards aglycone and was used in the present research for the deglucosylation of flavonoid and isoflavonoids **(Papers I and V)**. The specificity of this enzyme is dependent on the aglycone moiety, type of sugar and linkage, and can hydrolyse only certain flavonoid glycosides, a property which is common in GHs [17, 121]. Some glucosidases classified under GH1 [75] are flavonoid-hydrolysing enzymes. In a previous work, Turner et al have shown that the oligosaccharide hydrolysing GH1 wild-type β-glucosidase *Tn*Bgl1A efficiently hydrolyses quercetin-4´-glucosides (Q4´) [17] but selectivity towards quercetin-3-glucoside is lower **(Papers I and V)** [17]. The activity of this enzyme was also high towards genestein-7-glucoside, diadzein-7-glucoside and kaempferol-7-O-glucoside compared to glucosylation at position 3 **(Paper V).** Hydrolysis was not observed when rhamnose or galactose was attached to the aglycone **(Paper V)**. To understand why the enzyme prefers substrates glucosylated at certain positions rather than at other, active site residues were studied by engineering of *Tn*Bgl1A.

4.8. Engineering GH1 enzymes

In nature, enzyme function has evolved optimally according to the conditions in living cells and may not be perfect in specific biotechnological applications. For optimal performance, the desired features are high specificity, high activity and high stability. Not all proteins necessarily meet these requirements. Poor specificity towards a particular substrate is thus a major drawback for enzymes in particular applications. Amino acids in the active site of *Tn*Bgl1A studied in this thesis are very important as is the control their overall structure function relationship **(Papers I and V)**.

Structural analysis was used to guide selection of residues affecting deglucosylation of flavonoids and isoflavonoids and efforts were made to point out residues that may influence interaction with specific flavonoid substrates **(Papers I and V)**. In our work a first mutation selection idea was based on the aligment of the *Tn*Bgl1A sequence with oligosaccharide, flavonoid, and isoflavonoid hydrolysing GH1 enzymes as well as GH1 enzymes active on other bulky phenol-containing substrates such as, for example, the alkaloid strictosidine and raucaffricine sequences **(Paper I)**. The *Tn*Bgl1A sequence shows variability at positions 219, 221 and 222 in the vicinity of the substrate binding pocket, which were targeted for mutagenesis. The amino acids (F219L, N221S, G222Q, G222M) were changed on the basis of residues found in enzymes hydrolysing the bulky phenol-containing substrates. Two mutations were designed at G222, one with a hydrophobic (M) and another with a hydrophilic (Q) residue. The enzymes specific for large and bulky substrates contain residues L and S at positions 219 and 221, respectively; so in our enzyme the changes F219L and N221S were made towards those enzymes and to generate space for better substrate accommodation according to prediction. Interesting results were obtained with the N221S mutant, where the K_M value was decreased for Q3 **(Paper II)**.

The decrease in K_M value for Q3 could be explained by modelling of Q3 in *Tn*Bgl1A **(Paper I)**. Examination of the binding interactions of N221 in the presence of Q3 shows that this residue is important for the shape of the protein at that part of the active site as it interacts with the residues in neighbouring strands of the protein (the backbone carbonyls of Y246 and N291) (see Figure 10). When N221 is mutated to S221, it loses interactions with N291, but docking studies suggest that a change occurs in the position of the backbone carbonyl of the 291 residue, that may allow hydrogen

bonding to the 5-OH of Q3 (located in a position corresponding to the +1 sub site) **(Paper I)**.

The decrease in K_M for flavonoid glucosides in the aglycone binding site **(Paper I)** raised the interesting question of studying residues in this site, and amino acids in the +1 and +2 sites were targeted for further mutation **(Paper V)**. The hCBG (human cytosolic beta-glucosidase) residue F225 has an important role in the aglycone binding site, as has been previously shown by mutagenesis [121]. In the present study, the corresponding residue N220 in *Tn*Bgl1A was subjected to mutagenesis and the result shows that the mutation has no significant effect on the K_M of *p*NP-glucoside but does cause a 2.6 times decrease in K_M for Q-3-glucoside, combined with a 4.2-fold increase in catalytic efficiency.

Similar studies have been conducted by other researchers with the aim of improving the activity in thermostable β-glucosidase from *Pyrococcus furiosus* toward galactoside residues present in *Lactococcus lactis* 6-phospho-β-galactosidase. Serine at position 417 in the active site has previously been incorporated by site-directed mutagenesis, and analysis of E417S resulted in a 5-fold increase in the hydrolysis of *o*-nitrophenol-α-D-galactopyranoside-6-phosphate, while a significant decrease occurred in the activity on nonphosphorylated sugars [122]. In human β-glucosidase (hCBG), the substrate preference was studied by mutational analysis in the aglycone binding site and residues V168, F225 and Y308 lining the walls of aglycone binding pocket were found to be be involved in the catalytic activity, as their mutation resulted in a drastic decrease in the enzyme efficiency [121] , which indicates that the aglycone moiety in hCBG determines substrate hydrolysis. In maize β-glucosidase (ZmGlu1) the role of single mutants F193, F200, W373 and F461 forming the aglycone binding region was studied for the recognition of aglycone, and it was found that all the mutations except F461 had a severe effect on the enzyme interactions with large substrate. In addition, the specificity shifted towards small, polar aglycones such as *p*NP [123]. ZmGlu1 β-glucosidase has broad substrate specificity and can hydrolyse many synthetic and natural substrates but cannot hydrolyse dhurrin, the natural substrate for sorghum β-glucosidase (Sb-Dhr1). The inactive double mutant of ZmGlu1-E191D/F198V shows that valine at position 198 drastically reduces the activity of Zm-Glu1 [124]. The same residue V198 in Sb-Dhr1 has been changed toward Zm-Glu1 (V198F), but this did not shift the enzyme substrate specificity towards maize enzyme

substrates [124]. This behaviour clearly shows that different residues are responsible for aglycone recognition in Dhr1. For example, in rice Bglu1, four residues I179, N190, L442 and N245, in the +1 site have previously been mutated to match this enzyme to barley BGQ60, but the kinetic data show no expected improvement for cellobiose, which leads to the conclusion that no single residue is responsible for the oligosaccharide preferences between these enzymes [118]. The kinetic data of hCBG indicate that W345 is an important residue responsible for aglycone binding [112].
The crystal structure with ligand and the computational analysis information suggest that GH1 enzymes now can be engineered to obtain an enzyme with high specificity, activity and thermostability by using site directed mutagensis [107].

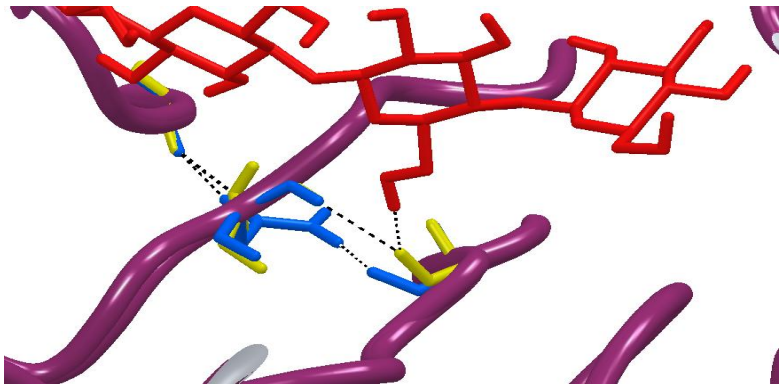


Figure 10. Interactions around position 221. *Tn*Bgl1A homology model with a cellotetraose ligand with and without residue N221 (blue) mutated to S (yellow). The S mutation allows for recognition of the 6-hydroxyl group of the glucose in the +1 subsite and of the 5-hydroxylgroup of quercetin, by hydrogen bonding to the backbone carbonyl oxygen of residue N291.

5. Enzymes immobilisation

An enzyme in any reaction system is usually said to be immobilised when the protein does not dissolve in the reaction medium, retains its catalytic activity and can be reused repeatedly and continuously. There are a number of practical benefits from immobilising an enzyme. The most obvious is that enzymes can be easily separated from the product mixture after a completed reaction and consequently can be reused in a batch-wise or in a continuous flow system. As a result, the substrate can be passed through immobilised enzymes and converted into a product during a pre-defined time depending on the enzymes stability [125-130]. The separation of enzymes from the product is essential in some cases such as with pharmaceuticals, where it is not acceptable for the enzymes to remain in the final product. In many cases enhancement of thermal stability and recycling of enzyme**s** is essential for the development of large-scale enzyme-catalysed reactions [131]. The enzyme studied in this thesis is a thermostable enzyme **(Papers I-V).** However, it has been found that the stability of thermophilic enzymes can be further stabilised by immobilisation [132].

At high reaction temperatures, quick deactivation is caused by partial unfolding of the enzyme molecule. Prevention of unfolding to provide a more rigid structure for the enzyme molecule can be accomplished by different immobilisation methods [133, 134]. Immobilised enzymes also enhance process robustness, allowing longer duration of activity and re-use in multiple cycles, which reduces production cost and can improve the cost-effectiveness of the process significantly [135].

The drawbacks to immobilisation are that it adds costs for carrier and other reagents required and often reduce the catalytic activity of the enzyme compared to processes using its soluble counterpart. This decrease in activity is partly due to the diffusion constraints of the substrate in the microenvironment compared to the macroenvironment, where the substrate is dispersed in the solution. The immobilisation process can also deactivate enzymes by modification of the active site and of the orientation of the enzyme, which limits its flexibility and can change or even remove the activity [135, 136]. However, immobilisation is especially important for expensive enzymes and has become an indispensable part of the use of industrial enzymes.

5.1. General properties of Immobilised Enzymes

Two things must be considered for any immobilised enzyme, independent of the method of preparation. One is the non-catalytic functions (NCF) and the second is the catalytic functions (CF). NCF are designed to aid separation by methods such as the isolation of the catalyst from the application environment, its reuse and control of the process. On the other hand, CF contribute to the conversion of the substrate into the product within the defined time and space. NCF are dependent on the physical and chemical nature of the carrier selected. Their selection depends in turn on aspects of the target application such as the reactor design, reaction medium, reaction system and the process conditions (temperature, pressure, and pH). The CF are linked to the catalytic properties of the biocatalyst such as activity, selectivity, durability of the catalyst, temperature and pH stability. Their selection is based on achieving the desired properties. To achieve an immobilised preparation which meets the desired CF and NCF, the hard part is the selection of carrier, enzyme and working conditions. It is very important to design a robust immobilised enzyme for a specific application which meets the requirements for both CF and NCF [135, 136].

Hundreds of enzymes have been immobilised; thousands of scientific articles, including patents, have been published and dozens of immobilised enzymes are currently in use in different industrial applications. However, the basic methods of enzyme immobilisations can be classified into a few methods such adsorption, entrapment, encapsulation, crosslinking and covalent immobilisation [137]. Today the problem is not how to immobilise enzymes, but how to get an immobilised with the desired performance in a specific application. Thus the immobilisation process is designed according to the requirements of each application and each individual application requires a specific solution [136].

The selection of immobilisation method is thus dependent on the application and on the enzyme stability. In our case, one of the aim was to use immobilised enzymes in on-line bioconversion of glucosylated antioxidants extracted by using hot water from yellow onion waste **(Paper IV)**. This type of processes can be carried out only at high temperatures. The choice of the method of immobilisation is therefore based on the specific conditions for this kind of application, and care must be taken in selecting the carrier and intrinsic properties of enzymes. Therefore, to obtain a stable immobilised preparation, covalent immobilisation was selected because the

immobilisation in this case is solid: the binding forces between support and enzyme are very strong and enzyme leaching is avoided. Large numbers of carriers and methods are available to activate these enzymes and one can achieve a stable immobilised preparation at the reaction condition by this method [129]. The drawback with this method is the difficulty in performing the experimental procedures and the 3-D structure of the enzyme is considerably modified, resulting in significant loss of the initial activity of the biocatalyst [9]. Efforts were made in the present study to set up an immobilisation method for the wild-type β-glucosidase from *T. neapolitana* as well as for an enzyme variant (N221S/P342L) with increased catalytic efficiency for hydrolysis of quercetin glucosides suitable for a high temperature process **(Papers III and IV)**.

5.2. Covalent immobilisation of *Tn*Bgl1A

Based on the reaction conditions involving high temperature, a covalent immobilisation method was selected **(Paper III)** as it creates a stable linkage. This method involves chemical reaction between different groups of the amino acids on the surface of enzymes with support materials. The amino acids which are expected to take part in immobilisation and which are exposed to aqueous environment are, for example lysine, serine and threonine residues. They are labelled in colours in Figure 11 and most of these residues are in *Tn*Bgl1A on the surface opposite the active site, and hence suitable for immobilisation with different groups of functionalised and activated supports **(Paper III)**. The covalent immobilisation method was thus chosen to obtain minimum protein leaching from the supports, and because of the absence of a barrier between the enzyme and the reaction medium. Furthermore covalent immobilisation can extend the lifetime of the enzyme by protecting its 3-D structure and can enhance stability [129, 138]. The thermostable β-glucosidase variants from *T. neapolitana* were covalently immobilised on Eupergit® C, Eupergit® C 250L and Cryogels, which are all epoxy- activated polyacrylic matrices **(Paper III)**. Two approaches were explored: (a) direct enzyme binding to the matrices via the oxirane groups, and (b) enzyme binding to the support via an ethylene diamine/glutaraldehyde spacer **(Paper III)**.

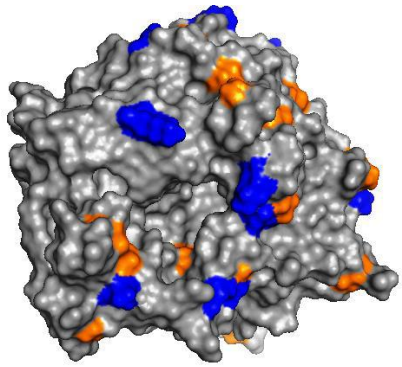


Figure 11. Surface view of the 3-D structure model of *Tn*Bgl1A. Lysine residues are labelled blue, Serine and Threonine residues in orange.

5.3. Support Materials

Epoxy (oxirane) activated support materials provide easy immobilisation of enzymes at both laboratory and industrial scale. They are stable during storage, chemically and mechanically, and stable at neutral pH [139, 140]. They are able to form stable covalent multipoint attachments with different protein groups (amino, thiol and phenolic) that are abundant on the enzyme, and easily cap the remaining epoxy groups during processes using a variety of reagents (BSA, mercaptoethanol, ethanolamine, glycine, etc.) to prevent any undesirable protein-support interaction [138, 141]. However, epoxy groups are hardly reactive under mild experimental conditions such as low ionic strength. In addition, immobilisation with epoxy groups is random, which might not promote conformational stabilisation, and some studies have even shown a decrease in stabilisation of the immobilised enzymes [142, 143]. Multi-covalent immobilisation can result in the distortion of the 3-D structure of enzyme, maintenance of which is crucial for enzyme activity, and such distortions often decrease enzyme activity [144].

Eupergit® C, Eupergit® C 250L and Cryogels are examples of epoxy-activated acrylic supports. Both Eupergit® C and Eupergit® C 250 L are micro-porous, epoxy-activated acrylic beads commercially available with a diameter of 100-250 μm and pore size 10-100 nm. Cryogels, on the other hand, are produced through gelation at subzero temperature and exhibit properties such as elastic and sponge-like morphology with pore size 1-100 μm [145]. They have high flow permeability, biocompatibility and high enzyme stability. Fast mass transfer characteristics have been reported for

enzymes immobilised on this material [145-147]. Ligands can be coupled directly or via spacer arms to a reactive group such as the epoxy group present on both Eupergit® C and cryogels [148, 149].
The suitability of the support material for application in a high temperature method was confirmed in our work by Scanning electron microscopy (SEM) images and thermogravimetric analysis (TGA) (Figure 12). The results confirm that the support materials were stable even above 150°C and can be used in future on-line flow systems at temperatures up to 150°C (above the current upper limit for protein stability) [150].

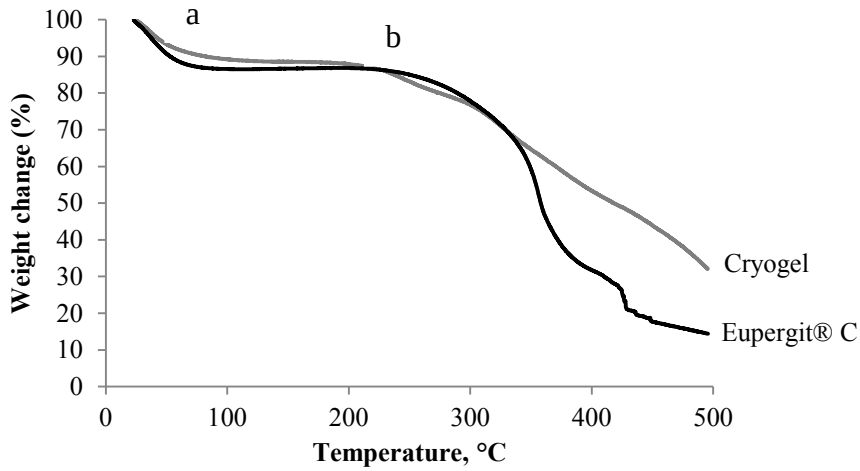


Figure 12. Thermogravimetric analysis of cryogel and Eupergit® C showing the change of weight of sample with temperature. (a) First weight loss due to loss of water and (b) second weight loss due to the decomposing of the polymer.

Comparing the specific activity (U/mg) obtained from enzymes immobilised on three support materials in the present study, a higher activity was reported for Eupergit® C 250L than for Eupergit® C and Cryogel. Both Eupergit® C and Eupergit® C 250L have the same composition. However, the lower specific activity on Eupergit® C could be due either to its smaller pore diameter, which may cause diffusion limitation [138], or to the higher content of the oxirane groups of Eupergit® C, which can result in the formation of multiple covalent bonds between the support and enzyme. This formation might in turn disturb the enzyme structure and cause some loss in activity [151]. The lower activity on the Cryogel compared to the Eupergit® supports is probably because of limiting the access of the substrate to the active site due to diffusion limitation. Higher activity on Cryogel was obtained when the

reactive groups in the support were secluded from its surface through short spacer arms. Thus, the spacer arms help in avoiding steric hindrance effects and improve activity, as they place the enzyme in a better position for catalytic action after immobilisation [152, 153]. The activity was higher from the enzyme immobilised on Cryogel when used in the on-line system where substrate passed through the enzyme and there was no problem of substrate access to enzyme **(Paper IV)**.

Many studies support the view that additives which bind to the active site of the enzyme or which coat unreacted groups on the support can improve the performance of the immobilised enzyme. Indeed, the addition of glucose resulted in a small improvement of the activity **(Paper III)** as it is believed to protect the active site of enzymes during immobilisation [154]. Similarly, the application of bovine serum albumin (BSA) after immobilisation enhanced the activity and it was possible to recover the same activity after a long time (tested after three months) by simply incubating the immobilised enzyme with BSA **(Paper III)** for 24 hours prior to use. Thus by immobilising the enzymes and add techniques to protect or recondition activity, a stable preparation can be obtained.

As mentioned above, our long term research goal is to develop sustainable methods for extraction and bioconversion of high value compounds (e.g. flavonoid antioxidants) from agricultural and industrial by-products. With an immobilised enzyme available, a step in this direction was to set up an on-line mini-flow system, combining the pressurised hot water extraction with biocatalysis without extensive cooling in-between the extraction and hydrolysis step. Such a system is predicted to maintain the low energy consumption and CO_2-equivalent release found for the batch method, and in addition to minimize manual handling. Several trials have been carried out in our study using immobilised thermostable β-glucosidases in combination with pressurised hot water extraction in the on-line mini-flow system for the conversion of quercetin glucosides to quercetin and carbohydrates **(Paper III and IV)**. The set-up was successful, showing that quercetin glucosides were extracted with hot water and passed through the immobilised enzymes, where the hydrolytic activity was monitored. The overall process time was however long in the current set-up (1.5-4h dependent on flow-rate) indicating need of optimisation. No change in enzyme activity was however observed for 5 hours, proving that the enzyme was stable and thus suitable for this kind of high temperature process **(Paper IV)**, and enzyme activity was not

significantly decreased during 48h using an 84°C process temperature. The cost of immobilised enzymes in industry is often an indicator of process viability and immobilisation often gives stability and durability to enzymes.

6. Concluding Remarks and Future Perspectives

The thesis reports on the characterisation of a β-glucosidase from the hyperthermophile (*T. neapolitana*), an extremely stable enzyme with unfolding temperature 101.9°C. We have shown that this enzyme can be successfully used in the modification of natural products such as quercetin-glycosides to free quercetin extracted from onion using hot water as extraction solvent. The high value compounds present in low value biomass from agricultural by-products and industrial waste were isolated by this environmentally sustainable method. In addition, these compounds were biocatalytically converted into a uniform product in a single step procedure and with shorter process time than for traditional methods **(Paper II)**. It was possible to improve the selectivity for the aglycone part of the flavonoid substrate by protein engineering using site-directed mutagenesis **(Paper I and V)**. Single mutations had negligible effect on the stability of the enzyme but an increased number of mutations were found to slightly decrease the stability **(Papers I and V)**.

This works shows that single changes in the active site are sufficient for changes in activity or specificity in this family of enzymes, and further studies of the enzyme at molecular level are motivated to further elucidate how enzymes in this family (with diverse specificities in related enzymes) evolve their functions. Co-crystallisation of the enzymes with quercetin glucosides or other related substrates would for example help in understanding the interactions. In addition, the enzymes can be further engineered for selective hydrolysis of substrates which are difficult to hydrolyse.

To produce the above desired variants, the chaperonin co-expression in *E. coli* cultures, shown to promote the active soluble enzyme fraction **(Paper VI)** is a valuable aid to obtain sufficient amounts of newly developed variants.

The wild-type and N221S/P342L mutant enzymes were immobilised on epoxy-activated support materials Cryogel and Eupergit® C, and specific activity was measured at different temperatures in the range 80-95°C using *p*NPGlc as model substrate **(Paper III and IV)**. The maximum activity was obtained when the enzyme was immobilised with Eupergit® C 250L. The storage stability was studied and it was found that the enzyme can be stored

for a long time, e.g. three months (the study period in our case). It was possible to recover the loss in activity upon storage by incubation with 1% BSA for 24 hrs. The study has also shown that glucose can improve the activity when used during immobilisation of the enzymes.

The immobilised enzymes were successfully used for the on-line bioconversion of glucosylated antioxidants present in onion when combined with a pressurised hot water extraction process **(Paper IV)**. The immobilised enzyme was also reused ten times without significant loss in activity using *p*NPGlc in a batch experiment **(Paper III).** The application was extended to onion samples in continuous mode and used for 5 hours without loss in activity **(Paper IV)**. Application trials on onion extracts show successful use of the enzyme, thus leading to further improvement of sustainability in the processing methodology **(Paper IV)**. The immobilised enzyme used in this study was stable at 84°C for 48 hours without loss of activity and can thus be used continuously for 48 hours at the reaction temperature, thereby proving that thermostable enzymes can be further stabilised by covalent immobilisation.

To continue this work in the future, other support materials may be tested to evaluate and hopefully improve the activity of immobilised enzymes. In addition, enzymes immobilised by the current methodology, can be used in optimisation trials, to further improve extraction yield, and decrease extraction times in the on-line flow system.

7. Bibliography

1. Davies, G. and B. Henrissat, *Structures and mechanisms of glycosyl hydrolases.* Structure, 1995. **3**(9): p. 853-859.
2. Cox, P.M., et al., *Acceleration of global warming due to carbon-cycle feedbacks in a coupled climate model.* Nature, 2000. **408**(6809): p. 184-187.
3. Slimestad, R., T. Fossen, and I.M. Vagen, *Onions: A source of unique dietary flavonoids.* J Agric Food Chem, 2007. **55**(25): p. 10067-10080.
4. Thomas, W., *Absorption and metabolism of flavonoids.* Free Radic Biol Med, 2004. **36**(7): p. 829-837.
5. Stalikas, C.D., *Extraction, separation, and detection methods for phenolic acids and flavonoids.* J Sep Sci, 2007. **30**(18): p. 3268-3295.
6. Ignat, I., I. Volf, and V.I. Popa, *A critical review of methods for characterisation of polyphenolic compounds in fruits and vegetables.* Food Chem, 2011. **126**(4): p. 1821-1835.
7. Naczk, M. and F. Shahidi, *Extraction and analysis of phenolics in food.* J Chromatogr A, 2004. **1054**(1-2): p. 95-111.
8. Hertog, M.G.L., P.C.H. Hollman, and D.P. Venema, *Optimization of a quantitative HPLC determination of potentially anticarcinogenic flavonoids in vegetables and fruits.* J Agric Food Chem, 1992. **40**(9): p. 1591-1598.
9. Villeneuve, P., et al., *Customizing lipases for biocatalysis: a survey of chemical, physical and molecular biological approaches.* J Mol Catal B-Enzym, 2000. **9**(4-6): p. 113-148.
10. Panke, S., M. Held, and M. Wubbolts, *Trends and innovations in industrial biocatalysis for the production of fine chemicals.* Curr Opin Biotechnol, 2004. **15**(4): p. 272-279.
11. Rozzell, J.D., *Commercial scale biocatalysis: Myths and realities.* Bioorg Med Chem 1999. **7**(10): p. 2253-2261.
12. Grattinger, M., et al., *Recombinant phosphoglycerate kinase from the hyperthermophilic bacterium Thermotoga maritima: Catalytic, spectral and thermodynamic properties.* J Mol Biol, 1998. **280**(3): p. 525-533.
13. Vieille, C., D.S. Burdette, and J.G. Zeikus, *Thermozymes*, in *Biotechnol Annu Rev*, M.R. El-Gewely, Editor 1996, Elsevier. p. 1-83.
14. J.B, H., *Plant polyphenols—XIV. : Characterization of flavonoid glycosides by acidic and enzymic hydrolyses.* Phytochemistry, 1965. **4**(1): p. 107-120.
15. Nuutila, A.M., K. Kammiovirta, and K.M. Oksman-Caldentey, *Comparison of methods for the hydrolysis of flavonoids and phenolic acids from onion and spinach for HPLC analysis.* Food Chem, 2002. **76**(4): p. 519-525.
16. Anastas, P. and N. Eghbali, *Green Chemistry: Principles and Practice.* Chem Soc Rev, 2010. **39**(1): p. 301-312.
17. Turner, C., et al., *Subcritical water extraction and beta-glucosidase-catalyzed hydrolysis of quercetin glycosides in onion waste.* Green Chem, 2006. **8**(11): p. 949-959.
18. Stojanović Dejan, B., et al., *Extremophiles: Link between earth and astrobiology.* Zbornik Matice Srpske za Prirodne Nauke, 2008. **2008**(114): p. 5-16.

19. Pinzon-Martinez, D.L., et al., *Thermophilic bacteria from Mexican thermal environments: isolation and potential applications.* Environ. Technol, 2010. **31**(8-9): p. 957-966.
20. Cardoso, A., et al., *Hunting for Extremophiles in Rio de Janeiro.* Front Microbiol, 2011. **2**.
21. Cavicchioli, R., et al., *Life and applications of extremophiles.* Environ Microbiol, 2011. **13**(8): p. 1903-1907.
22. Antranikian, G., *Extremophiles and Biotechnology*, in *eLS*,2001, John Wiley & Sons, Ltd.
23. Gross, M., *Life on the Edge: Amazing Creatures Thriving in Extreme Environments* 2001.
24. Macelroy, R.D., *Some comments on evolution of extremophiles.* Biosystems, 1974. **6**(1): p. 74-75.
25. Cary, S.C., T. Shank, and J. Stein, *Worms bask in extreme temperatures.* Nature, 1998. **391**(6667): p. 545-546.
26. Hough, D.W. and M.J. Danson, *Extremozymes.* Curr Opin Chem Biol, 1999. **3**(1): p. 39-46.
27. Stetter, K.O., *Extremophiles and their adaptation to hot environments.* FEBS Lett, 1999. **452**(1-2): p. 22-25.
28. Stetter, K.O., *Hyperthermophilic procaryotes.* FEMS Microbiol Rev, 1996. **18**(2-3): p. 149-158.
29. Unsworth Larry, D., J. van der Oost, and S. Koutsopoulos, *Hyperthermophilic enzymes stability, activity and implementation strategies for high temperature applications.* FEBS J, 2007. **274**(16): p. 4044-4056.
30. Huber, H. and K.O. Stetter, *Hyperthermophiles and their possible potential in biotechnology.* J Biotechnol, 1998. **64**(1): p. 39-52.
31. Atomi, H., T. Sato, and T. Kanai, *Application of hyperthermophiles and their enzymes.* Curr Opin Biotechnol, 2011.
32. Blochl, E., et al., *Pyrolobus fumarii, gen. and sp. nov., represents a novel group of archaea, extending the upper temperature limit for life to 113 degrees C.* Extremophiles, 1997. **1**(1): p. 14-21.
33. Kazem, K. and R.L. Derek, *Extending the Upper Temperature Limit for Life.* Science, 2003. **301**(5635): p. 934.
34. Bouzas, T.D., J. Barros-Velazquez, and T.G. Villa, *Industrial applications of hyperthermophilic enzymes: A review.* Protein Pept Lett, 2006. **13**(7): p. 645-651.
35. Joshi, G.K., et al., *Geothermal spring sites as excellent reservoir of novel microorganisms.* Environ. Conserv, 2011. **12**(1/2): p. 121-124.
36. Kristjánsson, J.K. and G.O. Hreggvidsson, *Ecology and habitats of extremophiles.* World J Microbiol Biotechnol, 1995. **11**(1): p. 17-25.
37. Brock, T.D., *Life at high temperatures - evolutionary ecological and biochemical significance of organisms living in hot springs is discussed.* Science, 1967. **158**: p. 1012-1019.
38. Bruins, M.E., A.E.M. Janssen, and R.M. Boom, *Thermozymes and their applications - A review of recent literature and patents.* Appl Biochem Biotechnol, 2001. **90**(2): p. 155-186.

39. Vieille, C. and G.J. Zeikus, *Hyperthermophilic enzymes: Sources, uses, and molecular mechanisms for thermostability.* Microbiol Mol Biol Rev, 2001. **65**(1): p. 1-43.
40. Bauer, M.W. and R.M. Kelly, *The family 1 beta-glucosidases from Pyrococcus furiosus and Agrobacterium faecalis share a common catalytic mechanism.* Biochemistry, 1998. **37**(49): p. 17170-17178.
41. Zeikus, J.G., C. Vieille, and A. Savchenko, *Thermozymes: biotechnology and structure–function relationships.* Extremophiles, 1998. **2**(3): p. 179-183.
42. Danson, M.J. and D.W. Hough, *Structure, function and stability of enzymes from the Archaea.* Trends Microbiol, 1998. **6**(8): p. 307-314.
43. Jannasch, H.W., et al., *Thermotoga-neapolitana sp-nov of the extremely thermophilic, eubacterial genus thermotoga.* Arch Microbiol, 1988. **150**(1): p. 103-104.
44. Belkin, S., C.O. Wirsen, and H.W. Jannasch, *A new sulfur-reducing, extremely thermophilic eubacterium from a submarine thermal vent.* Appl Environ Microbiol, 1986. **51**(6): p. 1180-1185.
45. Frock, A.D., J.S. Notey, and R.M. Kelly, *The genus Thermotoga: recent developments.* Environ Technol, 2010. **31**(10): p. 1169-1181.
46. Van Ooteghem, S., S. Beer, and P. Yue, *Hydrogen production by the thermophilic bacterium <i>Thermotoga neapolitana</i&gt.* Appl.Biochem. Biotechnol., 2002. **98-100**(1): p. 177-189.
47. Van Ooteghem, S.A., et al., *H-2 production and carbon utilization by Thermotoga neapolitana under anaerobic and microaerobic growth conditions.* Biotechnol. Lett., 2004. **26**(15): p. 1223-1232.
48. Childers, S.E., M. Vargas, and K.M. Noll, *Improved methods for cultivation of the extremely thermophilic bacterium Thermotoga-neapolitana.* Appl Environ Microbiol, 1992. **58**(12): p. 3949-3953.
49. Sunna, A., et al., *Glycosyl hydrolases from hyperthermophiles.* Extremophiles, 1997. **1**(1): p. 2-13.
50. Chhabra, S.R., et al., *Carbohydrate-induced differential gene expression patterns in the hyperthermophilic bacterium Thermotoga maritima.* J Biol Chem, 2003. **278**(9): p. 7540-7552.
51. Adams, M.W.W. and R.M. Kelly, *Finding and using hyperthermophilic enzymes.* Trends in Biotechnology, 1998. **16**(8): p. 329-332.
52. Henrissat, B., *A classification of glycosyl hydrolases based on amino-acid-sequence similarities.* Biochem. J, 1991. **280**: p. 309-316.
53. Baneyx, F. and M. Mujacic, *Recombinant protein folding and misfolding in Escherichia coli.* Nat Biotechnol, 2004. **22**(11): p. 1399-1408.
54. Hockney, R.C., *Recent developments in heterologous protein-production in Escherichia-coli.* Trends Biotechnol, 1994. **12**(11): p. 456-463.
55. Esposito, D. and D.K. Chatterjee, *Enhancement of soluble protein expression through the use of fusion tags.* Curr Opin Biotechnol, 2006. **17**(4): p. 353-358.
56. Loprasert, S., I. Urabe, and H. Okada, *Overproduction and single-step purification of Bacillus stearothermophilus peroxidase in Escherichia coli.* App Microbiol Biotechnol, 1990. **32**(6): p. 690-692.

57. Patchett, M.L., et al., *Heat treatment purification of thermostable cellulase and hemicellulase enzymes expressed in E. coli.* Enzyme Microb Technol, 1989. **11**(2): p. 113-115.
58. Retallack, D.M., et al., *Transport of heterologous proteins to the periplasmic space of Pseudomonas fluorescens using a variety of native signal sequences.* Biotechnol Lett, 2007. **29**(10): p. 1483-1491.
59. Hartl, F.U., *Molecular chaperones in cellular protein folding.* Nature, 1996. **381**(6583): p. 571-580.
60. Gragerov, A., et al., *Cooperation of groel/groes and dnak/dnaj heat-shock proteins in preventing protein misfolding in Escherichia-coli.* Proc Natl Acad Sci U S A, 1992. **89**(21): p. 10341-10344.
61. Nishihara, K., et al., *Overexpression of trigger factor prevents aggregation of recombinant proteins in Escherichia coli.* Appl Environ Microbiol, 2000. **66**(3): p. 884-889.
62. Nishihara, K., et al., *Chaperone coexpression plasmids: Differential and synergistic roles of DnaK-DnaJ-GrpE and GroEL-GroES in assisting folding of an allergen of Japanese cedar pollen, Cryj2 in Escherichia coli.* Appl Environ Microbiol, 1998. **64**(5): p. 1694-1699.
63. Fujiwara, S., *Extremophiles: Developments of their special functions and potential resources.* J Biosci Bioeng, 2002. **94**(6): p. 518-525.
64. Haki, G.D. and S.K. Rakshit, *Developments in industrially important thermostable enzymes: a review.* Bioresour Technol, 2003. **89**(1): p. 17-34.
65. Song, X., et al., *Comparison of Three Thermostable beta-Glucosidases for Application in the Hydrolysis of Soybean Isoflavone Glycosides.* J Agric Food Chem. , 2011. **59**(5): p. 1954-1961.
66. Wolfenden, R., X.D. Lu, and G. Young, *Spontaneous hydrolysis of glycosides.* J Am Chem Soc, 1998. **120**(27): p. 6814-6815.
67. Christopher, M., VanHolde, and Ahern, *Biochemistry 3rd edition.* 1999: p. 215-220.
68. Rouvinen, J., et al., *Three-dimensional structure of cellobiohydrolase II from Trichoderma reesei.* Science, 1990. **249**(4967): p. 380-386.
69. Cairns, J.R.K. and A. Esen, *β-Glucosidases.* Cell Mol Life Sci, 2010. **67**(20): p. 3389-3405.
70. Butters, T.D., *Gaucher disease.* Curr Opin Chem Biol, 2007. **11**(4): p. 412-418.
71. Wakarchuk, W.W., et al., *The molecular cloning and expression of a cellobiase gene from an Agrobacterium in Escherichia coli.* Mol Gen Genet, 1986. **205**(1): p. 146-152.
72. Painbeni, E., et al., *Purification and characterization of a bacillus-polymyxa beta-glucosidase expressed in Escherichia.coli.* J Bacteriol, 1992. **174**(9): p. 3087-3091.
73. Grabnitz, F., et al., *Structure of the beta-glucosidase gene bgla of clostridium-thermocellum, sequence analysis reveals a superfamily of cellulases and beta-glycosidases including human lactase phlorizin hydrolase.* Eur J Biochem, 1991. **200**(2): p. 301-309.

74. Gabelsberger, J., W. Liebl, and K.H. Schleifer, *Purification and properties of recombinant β-glucosidase of the hyperthermophilic bacterium thermotoga maritima.* App Microbiol Biotechnol, 1993. **40**(1): p. 44-52.
75. Cantarel, B.L., et al., *The Carbohydrate-Active EnZymes database (CAZy): an expert resource for Glycogenomics.* Nucleic Acids Res, 2009. **37**: p. D233-D238.
76. Bisaria, V. and S. Mishra, *Regulatory aspects of cellulase biosynthesis and secretion.* Crit Rev Biotechnol, 1989. **9**(2): p. 61-103.
77. Tomme, P., R.A.J. Warren, and N.R. Gilkes, *Cellulose Hydrolysis by Bacteria and Fungi*, in *Adv Microb Physiol*, R.K. Poole, Editor 1995, Academic Press. p. 1-81.
78. Lymar, E.S., B. Li, and V. Renganathan, *Purification and characterization of a cellulose-binding beta-glucosidase from cellulose-degrading cultures of phanerochaete-chrysosporium.* Appl Environ Microbiol, 1995. **61**(8): p. 2976-2980.
79. Ziya, G., *Flavor Enhancement in Fruit Juices and Derived Beverages by Exogenous Glycosidases and Consequences of the Use of Enzyme Preparations*, in *Handbook of Food Enzymology*2002, CRC Press.
80. Yasumoto, K., et al., *Isolation from rice bran of a bound form of vitamin-B6 and its identification as 5'-o-(beta-d-glucopyranosyl) pyridoxine.* Agric Biol Chem, 1977. **41**(6): p. 1061-1067.
81. Aziz, A.A., et al., *Absorption and excretion of conjugated flavonols, including quercetin-4 '-O-beta-glucoside and isorhamnetin-4 '-O-beta-glucoside by human volunteers after the consumption of onions.* Free Radic Res, 1998. **29**(3): p. 257-269.
82. *http://www.chem.qmul.ac.uk/iubmb/enzyme*. 4 Dec., 2011.
83. Sinnott, M.L., *Catalytic mechanisms of enzymatic glycosyl transfer.* Chem Rev, 1990. **90**(7): p. 1171-1202.
84. Henrissat, B. and A. Bairoch, *New families in the classification of glycosyl hydrolases based on amino-acid-sequence similarities.* J Biol Chem, 1993. **293**: p. 781-788.
85. Henrissat, B. and G. Davies, *Structural and sequence-based classification of glycoside hydrolases.* Curr Opin Struct Biol, 1997. **7**(5): p. 637-644.
86. Souza, J.N.S., et al., *Antioxidant capacity of four polyphenol-rich Amazonian plant extracts: A correlation study using chemical and biological in vitro assays.* Food Chem, 2008. **106**(1): p. 331-339.
87. DeGraft-Johnson, J., et al., *Ferric-reducing ability power of selected plant polyphenols and their metabolites: Implications for clinical studies on the antioxidant effects of fruits and vegetable consumption.* Basic clin pharmacol, 2007. **100**(5): p. 345-352.
88. Flora, S.J.S., *Structural, chemical and biological aspects of antioxidants for strategies against metal and metalloid exposure.* Oxid Med Cell Longev, 2009. **2**(4): p. 191-206.
89. Machlin, L.J. and A. Bendich, *Free-radical tissue-damage - protective role of antioxidant nutrients.* Faseb J, 1987. **1**(6): p. 441-445.

90. Corradini, E., et al., *Flavonoids: chemical properties and analytical methodologies of identification and quantitation in foods and plants.* Nat Prod Res, 2011. **25**(5): p. 469-495.
91. Nijveldt, R.J., et al., *Flavonoids: a review of probable mechanisms of action and potential applications.* Am J Clin Nutr, 2001. **74**(4): p. 418-425.
92. Larson, A.J., J.D. Symons, and T. Jalili, *Quercetin: a treatment for hypertension? - A review of efficacy and mechanisms.* Pharmaceuticals, 2010. **3**(1): p. 237-250.
93. Rice Evans, C.A., N.J. Miller, and G. Paganga, *Structure-antioxidant activity relationships of flavonoids and phenolic acids (vol 20, pg 933, 1996).* Free Radic Biol Med, 1996. **21**(3): p. 417-417.
94. Griffiths, G., et al., *Onions - A global benefit to health.* Phytother Res, 2002. **16**(7): p. 603-615.
95. Hollman, P.C. and a.I.C. Arts, *Flavonols, flavones and flavanols – nature occurrence and dietary burden.* J Sci Food Agric, 2000. **80**(7): p. 1081-1093.
96. Wood, T.M. and K.M. Bhat, *Methods for measuring cellulase activities*, in *Methods in Enzymology*, S.T.K. Willis A. Wood, Editor 1988, Academic Press. p. 87-112.
97. Murray, P., et al., *Expression in Trichoderma reesei and characterisation of a thermostable family 3 beta-glucosidase from the moderately thermophilic fungus Talaromyces emersonii.* Protein Expr Purif, 2004. **38**(2): p. 248-257.
98. Bhatia, Y., S. Mishra, and V.S. Bisaria, *Purification and characterization of recombinant Escherichia coli-expressed Pichia etchellsii beta-glucosidase II with high hydrolytic activity on sophorose.* App Microbiol Biotechnol, 2005. **66**(5): p. 527-535.
99. Yoon, J.J., K.Y. Kim, and C.J. Cha, *Purification and characterization of thermostable beta-glucosidase from the brown-rot basidiomycete Fomitopsis palustris grown on microcrystalline cellulose.* J. Microbiol, 2008. **46**(1): p. 51-55.
100. Korotkova, O.G., et al., *Isolation and properties of fungal beta-glucosidases.* Biochemistry (Mosc), 2009. **74**(5): p. 569-577.
101. Meyer, V.R., *Practical High-Performance Liquid Chromatography* 2004: 4th edition.
102. DA Skoog, F.H., TA Nieman *Principles of Instrumental Analysis*, ed. t. Edition1998.
103. Sanz-Aparicio, J., et al., *Crystal structure of beta-glucosidase A from Bacillus polymyxa: Insights into the catalytic activity in family 1 glycosyl hydrolases.* J Mol Biol, 1998. **275**(3): p. 491-502.
104. Sinnott, M., *Catalytic Mechanisms of Enzymic Glycosyl Transfer*. Chem. Rev, 1990. **90**: p. 1171-1202.
105. Burmeister, W.P., et al., *High resolution x-ray crystallography shows that ascorbate is a cofactor for myrosinase and substitutes for the function of the catalytic base.* J Biol Chem, 2000. **275**(50): p. 39385-39393.
106. Marana, S., *Molecular basis of substrate specificity in family 1 glycoside hydrolases.* IUBMB Life (International Union of Biochemistry and Molecular Biology: Life), 2006. **58**(2): p. 63-73.

107. Hill, A.D. and P.J. Reilly, *Computational analysis of glycoside hydrolase family 1 specificities.* Biopolymers, 2008. **89**(11): p. 1021-1031.
108. Davies, G.J., K.S. Wilson, and B. Henrissat, *Nomenclature for sugar-binding subsites in glycosyl hydrolases.* Biochem J, 1997. **321**: p. 557-559.
109. Fukuda, K., et al., *Identification of essential ionizable groups and evaluation of subsite affinities in the active site of beta-D-glucosidase F(1) from a Streptomyces sp.* Biosci Biotechnol Biochem, 2002. **66**(10): p. 2060-2067.
110. Hrmova, M., et al., *Catalytic mechanisms and reaction intermediates along the hydrolytic pathway of a plant beta-D-glucan glucohydrolase.* Structure, 2001. **9**(11): p. 1005-1016.
111. Marana, S.R., *Molecular basis of substrate specificity in family 1 glycoside hydrolases.* Iubmb Life, 2006. **58**(2): p. 63-73.
112. Tribolo, S., et al., *The Crystal Structure of Human Cytosolic β-Glucosidase Unravels the Substrate Aglycone Specificity of a Family 1 Glycoside Hydrolase.* J Mol Biol, 2007. **370**(5): p. 964-975.
113. Zechel, D.L., et al., *Iminosugar Glycosidase Inhibitors: Structural and Thermodynamic Dissection of the Binding of Isofagomine and 1-Deoxynojirimycin to β-Glucosidases.* J Am Chem Soc, 2003. **125**(47): p. 14313-14323.
114. Gloster, T.M., et al., *Glycosidase inhibition: An assessment of the binding of 18 putative transition-state mimics.* J Am Chem Soc, 2007. **129**(8): p. 2345-2354.
115. McCarter, J.D. and S.G. Withers, *Mechanisms of enzymatic glycoside hydrolysis.* Curr Opin Struct Biol, 1994. **4**(6): p. 885-892.
116. Koshland, D.E., *Stereochemistry and the mechanism of enzymatic reactions.* Biol Rev Camb Philos Soc, 1953. **28**(4): p. 416-436.
117. Rye, C.S. and S.G. Withers, *Glycosidase mechanisms.* Curr Opin Chem Biol, 2000. **4**(5): p. 573-580.
118. Chuenchor, W., et al., *Structural insights into rice BGlu1 beta-glucosidase oligosaccharide hydrolysis and transglycosylation.* J Mol Biol, 2008. **377**(4): p. 1200-1215.
119. Marana, S.R., et al., *Amino acid residues involved in substrate binding and catalysis in an insect digestive beta-glycosidase.* BBA-Proteins proteom, 2001. **1545**(1-2): p. 41-52.
120. Uhlig, C., et al., *Heterologous expression, refolding and functional characterization of two antifreeze proteins from Fragilariopsis cylindrus (Bacillariophyceae).* Cryobiology, 2011.
121. Berrin, J.G., et al., *Substrate (aglycone) specificity of human cytosolic beta-glucosidase.* Biochem. J, 2003. **373**: p. 41-48.
122. Kaper, T., et al., *Comparative structural analysis and substrate specificity engineering of the hyperthermostable beta-glucosidase CelB from Pyrococcus furiosus.* Biochemistry, 2000. **39**(17): p. 4963-4970.
123. Dopitova, R., et al., *Functional analysis of the aglycone-binding site of the maize beta-glucosidase Zm-p60.1.* FEBS J, 2008. **275**(24): p. 6123-6135.
124. Verdoucq, L., et al., *Structural determinants of substrate specificity in family 1 beta-glucosidases - Novel insights from the crystal structure of*

sorghum dhurrinase-1, a plant beta-glucosidase with strict specificity, in complex with its natural substrate. J Biol Chem., 2004. **279**(30): p. 31796-31803.
125. Rao, S.V., K.W. Anderson, and L.G. Bachas, *Oriented immobilization of proteins.* Microchim Acta, 1998. **128**(3): p. 127-143.
126. Hanefeld, U., L. Gardossi, and E. Magner, *Understanding enzyme immobilisation.* Chemical Society Reviews, 2009. **38**(2): p. 453-468.
127. Rusmini, F., Z.Y. Zhong, and J. Feijen, *Protein immobilization strategies for protein biochips.* Biomacromolecules, 2007. **8**(6): p. 1775-1789.
128. Linqiu, C., *Immobilised enzymes: science or art?* Curr Opin Chem Biol, 2005. **9**(2): p. 217-226.
129. Mateo, C., et al., *Improvement of enzyme activity, stability and selectivity via immobilization techniques.* Enzyme Microb Technol, 2007. **40**(6): p. 1451-1463.
130. Katchalski-Katzir, E. in *First enzyme engineering conference*. 1971. Henniker, New Hampshire: USA.
131. Torres-Bacete, J., et al., *Stabilization of penicillin V acylase from Streptomyces lavendulae by covalent immobilization.* J Chem Tech Biotechnol, 2001. **76**(5): p. 525-528.
132. Fernandezlafuente, R., D.A. Cowan, and A.N.P. Wood, *Hyperstabilization of a thermophilic esterase by multipoint covalent attachment.* Enzyme Microb Technol, 1995. **17**(4): p. 366-372.
133. Martinek, K., et al., *The principles of enzyme stabilization I. Increase in thermostability of enzymes covalently bound to a complementary surface of a polymer support in a multipoint fashion.* (BBA)-Enzymology, 1977. **485**(1): p. 1-12.
134. Klibanov, A.M., *Enzyme stabilization by immobilization.* Anal Biochem, 1979. **93**(1): p. 1-25.
135. Cao, L., L.v. Langen, and R.A. Sheldon, *Immobilised enzymes: carrier-bound or carrier-free?* Curr Opin Chem Biol, 2003. **14**(4): p. 387-394.
136. Linqiu, C., *Carrier-bound Immobilized Enzymes: Principles, Applications and Design*2005: WILEY-VCH Verlag GmbH & Co.
137. Klaus, M., *Immobilized enzymes.* Trends Biochem Sci, 1980. **5**(1): p. 1-3.
138. Wang, P.-Y., S.-W. Tsai, and T.-L. Chen, *Improvement of enantioselectivity and stability of Klebsiella oxytoca hydrolase immobilized on Eupergit C 250L.* J Chem Tech Biotechnol, 2008. **83**(11): p. 1518-1525.
139. Katchalskikatzir, E., *Immobilized enzymes- Learning from past successes and failures.* Trends Biotechnol, 1993. **11**(11): p. 471-478.
140. Hartmeier, W., *Immobilized biocatalysts - From simple to complex-systems.* Trends Biotechnol, 1985. **3**(6): p. 149-153.
141. Mateo, C., et al., *Increase in conformational stability of enzymes immobilized on epoxy-activated supports by favoring additional multipoint covalent attachment.* Enzyme Microb Technol, 2000. **26**(7): p. 509-515.
142. Alexander M, K., *Stabilization of Enzymes against Thermal Inactivation*, in *Adv Appl Microbiol*, I.L. Allen, Editor 1983, Academic Press. p. 1-28.
143. Gianfreda, L. and M.R. Scarfi, *Enzyme stabilization-State of the art.* Mol Cell Biochem 1991. **100**(2): p. 97-128.

144. Clark, D.S., *Can immobilization be exploited to modify enzyme activity.* Trends Biotechnol, 1994. **12**(11): p. 439-443.
145. Plieva, F.M., I.Y. Galaev, and B. Mattiasson, *Macroporous gels prepared at subzero temperatures as novel materials for chromatography of particulate-containing fluids and cell culture applications.* J Sep Sci, 2007. **30**(11): p. 1657-1671.
146. Massolini, G. and E. Calleri, *Immobilized trypsin systems coupled on-line to separation methods: Recent developments and analytical applications.* J Sep Sci, 2005. **28**(1): p. 7-21.
147. Josic, D., A. Buchacher, and A. Jungbauer, *Monoliths as stationary phases for separation of proteins and polynucleotides and enzymatic conversion.* J Chromatogr B-Analyt Technol Biomed Life Sci, 2001. **752**(2): p. 191-205.
148. Plieva, F., et al., *Macroporous polyacrylamide monolithic gels with immobilized metal affinity ligands: The effect of porous structure and ligand coupling chemistry on protein binding.* J Mol Recognit, 2006. **19**(4): p. 305-312.
149. Mallik, R. and D.S. Hage, *Affinity monolith chromatography.* J Sep Sci, 2006. **29**(12): p. 1686-1704.
150. Tanaka, T., et al., *Hyper-thermostability of CutA1 protein, with a denaturation temperature of nearly 150 °C.* FEBS Lett, 2006. **580**(17): p. 4224-4230.
151. Cho, Y.J., O.J. Park, and H.J. Shin, *Immobilization of thermostable trehalose synthase for the production of trehalose.* Enzyme Microb Technol, 2006. **39**(1): p. 108-113.
152. Luo, Q., et al., *High-performance affinity chromatography with immobilization of protein A and L-histidine on molded monolith.* Biotechnol Bioeng, 2002. **80**(5): p. 481-489.
153. Knezevic, Z., et al., *Immobilization of lipase from Candida rugosa on Eupergit C supports by covalent attachment.* Biochem Eng. J, 2006. **30**(3): p. 269-278.
154. Tu, M., et al., *Immobilization of beta-glucosidase on Eupergit C for lignocellulose hydrolysis.* Biotechnol Lett, 2006. **28**(3): p. 151-6.

Paper I

Khan et al. BMC Biochemistry 2011, 12:11
http://www.biomedcentral.com/1471-2091/12/11

BMC Biochemistry

RESEARCH ARTICLE Open Access

Aglycone specificity of *Thermotoga neapolitana* β-glucosidase 1A modified by mutagenesis, leading to increased catalytic efficiency in quercetin-3-glucoside hydrolysis

Samiullah Khan[1], Tania Pozzo[1], Márton Megyeri[1,2], Sofia Lindahl[3], Anders Sundin[3], Charlotta Turner[3], Eva Nordberg Karlsson[1*]

Abstract

Background: The thermostable β-glucosidase (*Tn*Bgl1A) from *Thermotoga neapolitana* is a promising biocatalyst for hydrolysis of glucosylated flavonoids and can be coupled to extraction methods using pressurized hot water. Hydrolysis has however been shown to be dependent on the position of the glucosylation on the flavonoid, and e.g. quercetin-3-glucoside (Q3) was hydrolysed slowly. A set of mutants of *Tn*Bgl1A were thus created to analyse the influence on the kinetic parameters using the model substrate *para*-nitrophenyl-β-D-glucopyranoside (*p*NPGlc), and screened for hydrolysis of Q3.

Results: Structural analysis pinpointed an area in the active site pocket with non-conserved residues between specificity groups in glycoside hydrolase family 1 (GH1). Three residues in this area located on β-strand 5 (F219, N221, and G222) close to sugar binding sub-site +2 were selected for mutagenesis and amplified in a protocol that introduced a few spontaneous mutations. Eight mutants (four triple: F219L/P165L/M278I, N221S/P165L/M278I, G222Q/P165L/M278I, G222Q/V203M/K214R, two double: F219L/K214R, N221S/P342L and two single: G222M and N221S) were produced in *E. coli*, and purified to apparent homogeneity. Thermostability, measured as T_m by differential scanning calorimetry (101.9°C for *wt*), was kept in the mutated variants and significant decrease (ΔT of 5 - 10°C) was only observed for the triple mutants. The exchanged residue(s) in the respective mutant resulted in variations in K_M and turnover. The K_M-value was only changed in variants mutated at position 221 (N221S) and was in all cases monitored as a 2-3 × increase for *p*NPGlc, while the K_M decreased a corresponding extent for Q3. Turnover was only significantly changed using *p*NPGlc, and was decreased 2-3 × in variants mutated at position 222, while the single, double and triple mutated variants carrying a mutation at position 221 (N221S) increased turnover up to 3.5 × compared to the wild type. Modelling showed that the mutation at position 221, may alter the position of N291 resulting in increased hydrogen bonding of Q3 (at a position corresponding to the +1 subsite) which may explain the decrease in K_M for this substrate.

Conclusion: These results show that residues at the +2 subsite are interesting targets for mutagenesis and mutations at these positions can directly or indirectly affect both K_M and turnover. An affinity change, leading to a decreased K_M, can be explained by an altered position of N291, while the changes in turnover are more difficult to explain and may be the result of smaller conformational changes in the active site.

* Correspondence: eva.nordberg_karlsson@biotek.lu.se
[1]Biotechnology, Dept of Chemistry, Lund University, P.O. Box 124, SE-221 00 Lund, Sweden
Full list of author information is available at the end of the article

Background

Glycoside hydrolases (GH) are enzymes that hydrolyse glycosidic bonds between two or more carbohydrates, or between a carbohydrate and a non-carbohydrate moiety. Carbohydrates are essential components of biomass, which is estimated to be produced in a quantity of about 60 Gt/year [1] and contain an array of structural and storage polysaccharides. To utilize these raw materials, microorganisms produce a wide variety of carbohydrate hydrolysing and modifying glycoside hydrolases. These enzymes can also be used as specific catalysts in industrial applications, *e.g.* in the food and feed industries, the paper and pulp, starch and textile industries, and in newly emerging sustainable processes [2,3] taking advantage of their specificity in selective preparations of carbohydrate-containing raw materials.

Antioxidants are bioactive compounds that have received great interest due to their potential as health beneficial agents. The action of antioxidants is to counteract oxidative stress imposed by reactive oxygen species shown to play a crucial role in the pathophysiology associated with neoplasia, atherosclerosis and neurodegenerative diseases [4,5]. Polyphenolic compounds show a wide range of antioxidant activities, and are thought to exert protective effects against the diseases specified above [5-8]. Flavonoids are polyphenolic compounds that are important antioxidative constituents of fruits and vegetables, but the type of compound is different in different sources. Fruits and vegetables rich in anthocyanins (*e.g.* strawberry, raspberry and red plum) show highest antioxidant activities, followed by those rich in flavonones (*e.g.* orange and grapefruit) or flavonols (*e.g.* onion, leek, spinach and green cabbage), while hydroxycinnamate-rich fruits (*e.g.* apple, tomato, pear and peach) exhibit lower antioxidant activities [5,9]. Flavonols are generally not found as free aglycones (*e.g.* quercetin and kaempferol), but rather as complex conjugates with sugar residues (*e.g.* glucose or rhamnose). GHs hydrolyse certain flavonoid glycosides, dependent on the aglycone moiety, type of sugar and linkage [3,10] and some glucosidases classified under GH family 1 (GH1) [11] are flavonoid-hydrolysing enzymes. In a previous work we have shown that the oligosaccharide hydrolysing GH1 β-glucosidase *Tn*Bgl1A, from the hyperthermophile *Thermotoga neapolitana*, efficiently hydrolyses quercetin-4'-glucosides (Q4') [3]. Q4' and quercetin-3,4'-diglucoside (Q3,4') are the most abundant species in yellow onion [12], and to obtain the quercetin aglycone both Q4' and Q3 species must be hydrolysed.

Most commonly quercetin and its glycosides have been extracted from yellow onion by simple liquid/solid extraction techniques (*e.g.* aqueous methanol) combined with chemically (*e.g.* HCl) catalyzed hydrolysis reaction of the extracted quercetin glucosides [12-16].

We have instead utilized pressurized hot water to extract the quercetin species from yellow onion, followed by biocatalytic conversion of the quercetin glucosides to quercetin and carbohydrates [3]. In this system, use of enzymes with high thermostability is essential, and the enzyme *Tn*Bgl1A from the hyperthermophile *T. neapolitana* is from this perspective a suitable biocatalyst [3,17]. It was also shown that the Q4' was much more efficiently hydrolysed than the Q3.

In this investigation, the β-glucosidase *Tn*Bgl1A, was mutated to investigate the influence of mutations on the enzyme kinetics (using the substrate *para*-nitrophenyl-β-D-glucopyranoside (*p*NPGlc)), thermostability, and selective hydrolysis of glucose at two positions (4' and 3) on the aglycone quercetin. We have utilized a strategy to design mutants based on bioinformatics and structural analysis, with an amplification protocol that allowed spontaneous mutations, in order to find residues that influence specificity of the enzyme.

Methods

Chemicals

All chemicals were of pro-analysis grade from Merck Eurolabs (Darmstadt, Germany) unless otherwise stated.

Cloning of *Tnbgl1A*

The gene encoding Bgl1A was PCR-amplified from genomic *Thermotoga neapolitana* (DSM strain 4359) DNA as described by Turner et al [3]. Primers (1 and 2 with restrictions sites for cloning, *Nde*I and *Xho*I, underlined, Table 1) were designed to amplify the coding sequence of *bgl1A (*previously termed *gghA* [18]) from the sequence deposited at the NCBI server [19] under the accession number AF039487. The gene was inserted in vector pET-22b(+) (Novagen, Madison, WI, USA) under control of the T7/lac promoter and incorporating the C-terminal hexa-histidine tag [3]. The resulting plasmids were transformed into *E. coli* Nova Blue cells (Novagen) and screened by colony PCR using the T7 forward and T7 reverse primers (3 and 4, Table 1) and *Taq* DNA polymerase. Positive clones were transformed into the *E. coli* expression host Tuner (DE3) (Novagen). The complete gene was sequenced at MWG Biotech (Ebersberg, Germany).

Mutagenesis

Mutagenesis was performed in order to introduce the following designed changes: F219L, N221S, and G222M/Q, respectively. *Taq* polymerase, which lacks proofreading, (Invitrogen Life Technologies) was used (with wild-type gene as template) to allow introduction of a few random additional mutations. Standard concentrations of $MgCl_2$ (1.5 mM) and dNTPs (200 μM) were used. In a first PCR (94°C 3 min; 35 cycles: 94°C 45 s, 55°C 30 s,

Table 1 Oligonucleotides used for cloning and mutagenesis

Primer	Direction	Mutation	Sequence[a,b]
1	Forward	N/A	5'-TATTCTTATCATATGAAAAAGTTTCCCGAAGGGTTC
2	Reverse	N/A	5'-TATTCTTATCTCGAGATCTGTTAGTCCGTTGTTTTTG
3	Forward	N/A	5'-AATACGACTCACTATAGG
4	Reverse	N/A	5'-CTAGTTATTGCTCAGCGG
5	Forward	F219L	5'GACGGAAAGATAGGGATTGTT**TTA**AACAACGGATACTTCGA
6	Forward	N221S	5'-GGGATTGTTTTCAAC**AGC**GGATACTTCGAACCTGC
7	Forward	G222M	5'-GATTGTTTTCAACAAC**ATG**TACTTCGAACCTGCAAG
8	Forward	G222Q	5'- GGATTGTTTTCAACAAC**CAA**TACTTCGAACCTGCAAGTGAGAGAG
9	Reverse	F219L	5'- GACGGAAAGATAGGGATTGTT**TTA**AACAACGGATACTTCGA
10	Reverse	N221S	5'- GCAGGTTCGAAGTATCC**GCT**GTTGAAAACAATCCC
11	Reverse	G222M	5'-CTCTCTCACTTGCAG**TAC**CGAAGTATTGGTTGTTGAA
12	Reverse	G222Q	5'- CTCTCTCACTTGCAG**GTT**CGAAGTATTGGTTGTTGAAAACAATCC

[a]The restriction sites for cloning are underlined. [b]Induced changes are shown in bold.

72°C 90 s, 72°C 10 min), a mutated gene fragment encoding the C-terminal part of the enzyme was constructed using forward mismatched primers (primers 5-8, mismatch in bold, Table 1) together with the reverse gene specific primer 2, encoding the C-terminal sequence of Bgl1A.

The N221S mutants were obtained by amplifying the full length gene in a second PCR (94°C 3 min; 35 cycles: 94°C 45 s, 60°C 30 s, 72°C 90 s, 72°C 10 min) using the product of the first PCR as a reverse "megaprimer" together with the forward primer 1 matching the start of the gene.

Mutations at position 219 (F219L), and 222 (G222M/Q) were constructed using overlap extension PCR of the mutated fragment and an overlapping gene fragment encoding the N-terminal part of Bgl1A. Reverse primers (9-12, Table 1) together with the forward gene specific primer 1 were used to create the overlapping fragments that were PCR-amplified (94°C, 3 min; 35 cycles: 94°C 45 s, 55°C 30 s, 72°C 90 s, 72°C 10 min) with the wild-type gene as a template. Overlap extension PCR reaction was then run in two steps. Firstly, extension without template at: 94°C 4 min; 10 cycles: 94°C 1 min, 47.5°C 1 min, 72°C 90 s, 72°C 7 min and the ramp between the annealing and extension changed from the default 3°C/s to 5°C/s. Secondly, amplification at standard conditions, using the gene-specific forward and reverse primers (1 and 2), and the product of the overlap extension PCR as template.

All mutated genes (inserts) were purified with QIAEXII Gel Extraction kit (Qiagen) after gel separation. Both insert and vector were digested and ligated as described under the cloning section. Resulting plasmids were transformed into *E. coli* Nova Blue cells (Novagen) and screened by colony PCR using the T7 forward and T7 reverse primers (3 and 4, Table 1). Selected mutant clones were fully sequenced at MWG Biotech.

Homology modelling and ligand binding

A *Tn*Bgl1A homology model was constructed utilizing the Schrodinger 2010 software suite [20]. Energy minimizations were performed with Macromodel, utilizing the OPLS-2005 force field and the GB/SA model for water solvation. Molecular dynamics was performed with Desmond, utilizing default settings.

To find homologes to the *Tn*Bgl1A sequence, a BLAST search was performed on proteins with X-ray diffraction data in the PDB data base. The amino acid sequence of the GH1 β-glucosidase from *Thermotoga maritima* provided the highest similarity, 90% identity, and 97% positives. The PDB structure 2WC4 of BglA from *T. maritima* with 3-imino-2-thio-(+)-castanospermine as ligand, had the highest resolution of these X-ray crystallographic models (1.7 Å), and was used as template to build a 3D model of *Tn*Bgl1A (Prime version 2.2.108). The Protein Preparation Wizard was utilized to add hydrogens, assign charges, optimize hydrogen bond networks, and to analyze the quality of the homology model. The RMSD (over all α-carbons) with the template was 0.2 Å after optimisation.

In the region of the active site, a cis-bond between W396 and S397 was found, however this cis-amide is present in all crystals of *T. maritima* β-glucosidase. Furthermore, energy minimization showed the cis conformer to have lower energy than the corresponding trans amide bond.

A low energy model of β-D-cellotetraose was constructed [21] and placed in the homology model so that the β-D-glucose at the non-reducing end superimposed with the six membered rings of the *T. maritima* BglA ligands in the pdb structures 2WC4, 3CMJ, 1QOX, 1E4I, and 1BGA. The glycosidic linkage torsion angle φ of this sugar was adjusted from -121° to 26° to make cellotetraose fit in the ligand binding pocket. The resulting structure was energy minimized to yield a cellotetraose binding pose

that is in agreement with the proposed mechanism of family 1 β-glucosidases [22]. The binding pose was verified as stable in 2 ns molecular dynamics simulation.

A conformational search was performed on Q3, and the lowest energy conformer was placed in the homology model using the same method as for cellotetraose. The resulting structure was energy minimized, and the binding pose, which was in agreement with the proposed mechanism of family 1 β-glucosidases, was verified with a 2 ns molecular dynamics simulation.

Figures of docking results were prepared within the Schrodinger 2010 software suite, including the program PyMol.

Expression and purification

The mutant and wild-type enzymes were produced in 2.5 L batch cultivations at 37°C, pH 7, using a defined medium [23] with 100 μg mL^{-1} ampicillin and a dissolved oxygen tension (DOT) above 40%. Expression was induced at $OD_{620\ nm}$ = 3, by the addition of 0.1-1 mM isopropyl-beta-D-thiogalactopyranoside (IPTG), and continued for 3 h. Production levels were analysed by SDS-PAGE (see below) in samples (1 mL) withdrawn hourly after induction. The cells were pelleted, resuspended in 300 μL 50 mM citrate-phosphate buffer, pH 5.6, ultrasonicated for 2 × 90 s with a UP400S equipped with a 3 mm titanium probe (Dr. Hielscher, Stahnsdorf, Germany) using a sound intensity of 60% and a cycle of 0.5 and thereafter centrifuged for 15 min at 13 000 × *g* to separate soluble proteins from insoluble proteins and cell debris.

The cells were harvested, separated from the cultivation medium by centrifugation (10000 × *g*, 10 min, 4°C), and dissolved in binding buffer (20 mM imidazole, 20 mM Tris-HCl, 0.75 M NaCl, pH 7.5). The ice-chilled cell suspension was lysed by sonication for 5 × 3 min using a 14 mm titanium probe sound intensity of 60% and a cycle of 0.5 (UP400 S, Dr. Hielscher), centrifuged (30 min, 39000 × *g*, 4°C), heat treated (70°C, 30 min) and again centrifuged. The supernatant was passed through a 0.45 μm Minisart high-flow filter (Sartorius, Göttingen, Germany) and purified on an ÄKTA prime system (Amersham Biosciences, Uppsala, Sweden) by immobilized metal ion affinity chromatography using copper as a ligand as described elsewhere [24]. The fractions containing the purified protein were pooled and dialyzed against 20 mM citrate phosphate buffer, pH 5.6, overnight using a Spectra/Por dialysis membrane with a 3500 Da molecular weigh cut-off (Spectrum laboratories, Rancho Dominguez, CA, USA). The dialysed protein fractions were stored at 4°C until use.

Protein analysis

The purity of each mutant and wild-type enzyme was analysed by SDS-PAGE according to [25]. Expression levels were also analysed by SDS-PAGE after separating insoluble and soluble proteins.

Total protein concentration was estimated at 562 nm by the BCA method (Sigma, Steinheim, Germany) using bovine serum albumin (Sigma-Aldrich) as standard.

Differential scanning calorimetry (DSC)

DSC analysis was made on a MicroCal differential scanning calorimeter (VP-DSC, MicroCal, Northampton, MA, USA) with the cell volume of 0.5072 mL. The samples (in 20 mM citrate phosphate buffer, pH 5.6) were concentrated to 1 mg mL^{-1} using Vivaspin (Sartorius AG, Goettingen, Germany) centrifuging tubes with a MWCO of 30,000 Da and were degassed before the scans. The samples were scanned at a rate of 1°C/min in the temperature range of 25-110°C.

Enzyme activity on *p*NPGlc, Q3 and Q4'

Enzyme activity, and kinetic parameters (K_M and k_{cat}) were determined at 80°C, pH 5.6 using *p*NPGlc (*para*-nitrophenyl-β-D-glucopyranoside, Figure 1A) as substrate in 20 mM citrate phosphate buffer, on a Shimadzu UV-1650 Visible spectrophotometer (Shimadzu, Duisburg, Germany). A volume of 980 μL of *p*NPGlc (in a concentration range from 0.09125 to 1 mM) was preheated for 10 min, where after 20 μL of the enzyme solution (12 μg mL^{-1}, 4.56 pmol) was added. Absorbance at 405 nm was measured and plotted by the Shimadzu UV probe 2.01 software as a function of time during 1 min. The extinction coefficient of *p*NP (*para*-nitrophenol) under these experimental conditions was determined as $\varepsilon_{80°C,\ 405\ nm} = 2.4639 \times 10^3$ mL $mmol^{-1}cm^{-1}$. The kinetic parameters were determined by applying the Wilkinson non-linear regression method using Enzpack (Biosoft, UK). The effect of glucose on activity in this system was evaluated by adding 10 mM glucose to the stock solution of 1 mM *p*NPGlc and the kinetic values K_M and k_{cat} were determined as above.

Enzyme activity of selected mutants of *Tn*Bgl1A were screened for two quercetin glucosides (Figure 1B), quercetin-3-glucoside (Q3) (Polyphenols Laboratories AB, Sandnes, Norway) and quercetin-4'-glucoside (Q4') (Polyphenols Laboratories AB) at 90°C. 200 nmol samples of Q3 or Q4' dissolved in methanol (in triplicate) were evaporated and 1.0 mL of 100 mM citrate phosphate buffer, pH 5.0, was added and the vials were heated at 90°C until substrate was dissolved. A 50 μL fraction was collected and added to 450 μL of mobile phase composed of methanol/water (50:50) and 0.13 M formic acid. The reaction was started by adding 20 pmol of enzyme and 5 min after addition of enzyme, 50 μL fractions were collected and added to 450 μL of mobile phase. Samples were analyzed by HPLC with UV detection (HPLC-UV). The conditions and methodolology for

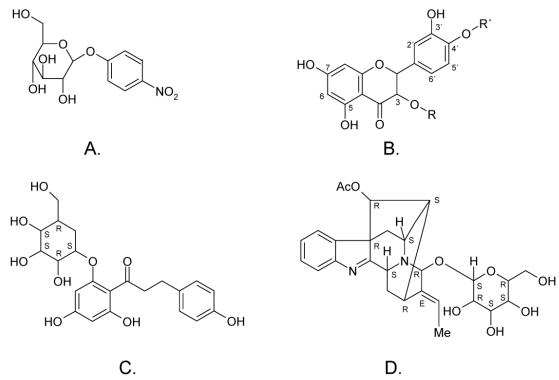


Figure 1 Schematic structures of the substrates used for enzyme activity analysis and the bulky alkaloid substrates hydrolysed by some GH1 representatives. Structures of *para*-nitrophenyl-beta-D-glucopyranoside (**A**.), quercetin species (**B**.), and the alkaloids strictosidine (**C**.), and raucaffricine (**D**.). Quercetin (R, R' = H), quercetin-3-glucoside (R = glucose and R' = H), quercetin-4'-glucoside (R = H and R' = glucose) and quercetin-3,4'-glucoside (R, R' = glucose).

the kinetic measurement of Q3 and Q4' was published in Lindahl et al [17]. In summary, 33-167 nmol Q3 and Q4' were dissolved per ml citrate-phosphate buffer pH 5.0. For Q3 hydrolysis 200 pmol *wt* and 20 pmol mutant N221S/P342L were used, and for Q4' hydrolysis 20 pmol *wt* and 10 pmol mutant N221S/P342L were used.

HPLC analysis

HPLC-UV analysis was performed using the chromatographic system UltiMate 3000 from Dionex (Germering, Germany). An Agilent Zorbax SB-C18 column (100 × 2.1 mm, 3.5 μm) was used for isocratic separation with a methanol:water (50:50) and 0.13 M formic acid mobile phase at a flow rate of 0.15 mL min^{-1}. The injection volume was 10 μL and detection was accomplished at 350 nm. Quantification of quercetin and glycosides was performed using a five-point calibration curve of a quercetin dihydrate standard (Sigma-Aldrich, Steinheim, Germany) and Q3 and Q4' standards at concentrations between 0.5 and 25 μg mL^{-1}. Each vial taken to analysis had a total volume of 500 μL.

Results

The different members of glycoside hydrolase family 1 (GH1), catalyse hydrolysis of a glucose molecule from a number of different substrates, including some hydrophobic substrates linked to glucose. Only a few members of the family are, however, commercially available, and many different analytical assays have thus utilized β-glucosidase from almond, which has been classified under GH1 [26], and which is often available in heterogeneous preparations. β-Glucosidase A from *Thermotoga neapolitana*, *Tn*Bgl1A, was chosen for this work as an interesting candidate based on previous promising results in biocatalytic conversion of quercetin glucosides to quercetin and carbohydrates in yellow onion extract, extracted using a pressurized hot water extraction technique [3]. This enzyme is thermostable, and hence suitable for application in hot water. Moreover, this enzyme belongs to a GH family with diverse substrate specificity (including enzymes active on both oligosaccharides and larger substrates), with many gene sequences available allowing comparison, and with necessary structural information available, including three-dimensional (3D) structures of the closely related enzyme *Tm*GH1 from *Thermotoga maritima* [27,28], allowing homology modelling of the 3D structure.

Structural considerations and mutation strategy

As noted elsewhere [3], the deduced amino acid sequence encoded by the *bgl1a*-gene used in this work has one change in primary sequence (G436V) compared to the deposited sequence (NCBI accession number AAB95492). Sequence alignments revealed V at position 436 to be conserved among several members of GH family 1, and the obtained sequence is hereafter referred to as wild type (*wt*) and designated *Tn*Bgl1A.

A molecular model of *T. neapolitana Tn*Bgl1A was generated based on the 3D structure of β-glucosidase BglA from *T. maritima* (here termed *Tm*GH1) [27,28]. *Tm*GH1 provided the highest similarity: 90% identity, and 97% positives. The homology detection structure

prediction server HHpred [29,30] also confirmed *Tm*GH1 as the best template. The overall structure presented in the *Tn*Bgl1A model is the typical $(\beta/\alpha)_8$ barrel fold characteristic of GH1 (Figure 2A). The active site is a deep channel (18-21Å) narrowed at the bottom with a wider cleft at the entrance. Two conservative motifs TLN**E**P and IT**E**NG are situated opposite to each other inside the active site, more precisely at the ends of strands β4 and β7 containing the acid/base (E164) and the nucleophile (E349) residues respectively.

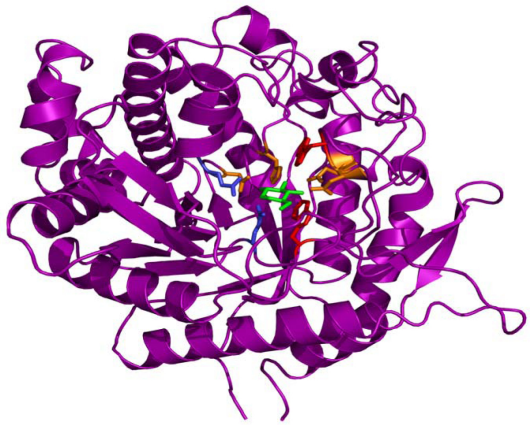
Superimposition of *Tm*GH1 in complex with 2-deoxy-2-fluoro-β-D-glucopyranoside (2GF) (PDB 1OIN) and *Tn*Bgl1A allowed identification of the residues forming the glycone binding site (or -1 subsite) (Figure 2A). Interacting residues included conservative residues forming hydrogen bonds (Q18, H119, N163, E403) and hydrophobic interactions (W396, W404).

Cellotetraose was modelled into the *Tn*Bgl1A model to analyse putative interactions at the +1 and +2 subsites. The β-D-glucopyranoside at the non-reducing end (-1 subsite) was placed by superimposing the six membered rings of the ligand (hexose rings or inhibitors) in 5 GH1 structures (see materials and methods for details). The sugars in the cellotetraose were positioned in a dynamically stable position in agreement with the proposed mechanism of family 1 β-glucosidases [22]. The same procedure was repeated for the glucosylated quercetin to see its position in relation to the sugar binding subsites. Both Q3 (Figure 2B) and Q4' could be fitted in positions relevant for hydrolysis, while the double glucoside (Q3,4') could not be accommodated with the 3-glucoside positioned at the -1 subsite (data not shown). This explains why hydrolysis at the 4'-position precedes hydrolysis of glucose bound at the 3-position in the double glucoside [17].

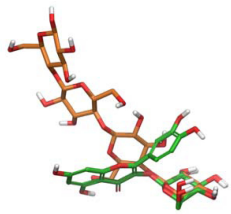
The *Tn*Bgl1A aglycone +1 subsite is formed by the hydrophilic residues (N171, H178, N220) and mainly by aromatic and hydrophobic residues (W33, F36, W120, V167, V171, W322, A405, and F412). A comparison with structure determined plant enzymes from GH1 (*Oryza sativa japonica*, *Zea mays)* showed these enzymes to also display aromatic and hydrophobic residues at this subsite but generally with longer hydrophobic residue side chains at the corresponding positions [31].

Inspection of residues surrounding the +2 subsite showed a non-conserved region at the "floor" of the active site, more precisely at the end of β-strand 5 (F219, N220, N221, G222, Y223, F224) (Figure 2C). The nonconserved nature of these residues were shown by analysing a multiple sequence alignment as well as by superimposition of known structures. Five more hydrophobic interactions including two aromatic residues (W166, I170, V171, V176, and F310) were also found. Superimposition of 3D structures of GH1 enzymes with varying substrate specificities (of different origin and thermostability such as *Pyroccoccus horikoshii* OT3 (PDB code, 1VFF), *Paenibacillus polymyxa* (PDB code 2Z1S), *Homo sapiens* (PDB code 2JFE), *Oryza sativa japonica* (PDB 2RGL) and *Zea mays* (PDB code 1E4N)) confirmed the variable area of β-strand 5 in the vicinity of the substrate pocket binding as well as variability in the loops. In general, thermophilic β-glucosidases presented shorter loops and more compact overall structures compared to plant counterparts, in line with previous results [32]. The variability at the end of β-strand 5 was also corroborated by the multiple sequence alignment of the *Tn*Bgl1A sequence (Figure 3) with sequences of oligosaccharide, flavonoid-, and isoflavonoid hydrolysing GH1 enzymes as well as GH1 enzymes active on other bulky phenol-containing substrates like *e.g.* the alkaloids strictosidine and raucaffricine (Figure 1C and 1D). The residues at position 219, 221 and 222 were targeted for mutagenesis because of the sequence variation between specificity groups at these sites combined with their location close to the cellotetraose +2 sugar residue. The changes (F219L, N221S, G222Q, G222M) (Figure 2C) were chosen based on residues found in enzymes hydrolysing the bulky phenol-containing substrates (Figure 3). In the case of G222 two mutations were designed, one with an hydrophobic (M) and another one with hydrophilic (Q) residue. Changing a G for M was made to increase hydrophobicity at the entrance of the active site but may exclude water molecules reducing cleavage of the glucosidic linkage in hydrolysis reactions. The G222Q mutation could instead result in substrate interactions via hydrogen bonds (*e.g.* with OH2 and OH3 of a carbohydrate substrate). The F219L and N221S mutations were selected based on residues found at corresponding positions in the enzymes specific for large and bulky substrates, and predicted to generate space for better substrate accommodation.

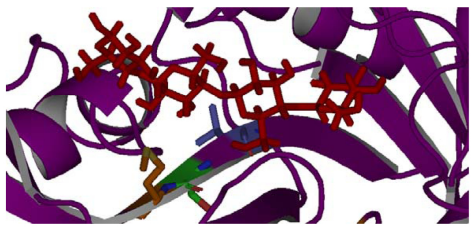
The residues were mutated in a protocol that introduced a few spontaneous mutations during the amplification procedure. Sequencing of the obtained genes showed that the designed mutations were obtained in all cases, and that one or two spontaneous mutation(s) were present in six clones. The selected clones included four genes with triple mutations: F219L/P165L/M278I, N221S/P165L/M278I, G222Q/P165L/M278I, G222Q/V203M/K214R, two with double mutations: F219L/K214R, N221S/P342L, and two with the single mutation G222M and N221S. The spontaneous mutations (totally five residues) in principle involved conserved changes (replacing a hydrophobic residue with another hydrophobic residue in two cases (V203M, M278I), replacing proline with a hydrophobic residue in two cases (P165L, P342L), and a charged basic residue in one case



A.



B.



C.

Figure 2 Overall structure of *Tn*Bgl1A, relative substrate positions and the cellotetraose position in relation to mutated residues. *Tn*Bgl1A homology model showed the typical $(\beta/\alpha)_8$ barrel fold (**A**.), a feature of the overall structure in GH1. The proton donor E164 and nucleophile E349 of the enzyme are illustrated in blue and shown in stick representation. In green the G2F inhibitor is shown in the -1 subsite, interacting with four residues in orange: (Q18, H119, N163, E403) by hydrogen bonds and two by hydrophobic interactions (W396, W404). In panel **B** the relative positions of the ligands cellotetraose (orange) and quercetin-3-glucoside (Q3, in green) when bound in the enzyme are shown. The matching binding of the glucopyranoside in the two substrates at the -1 subsite is shown to the right. The cellotetraose labelling from left to right correspond to subsites +3, +2, +1, -1. In panel **C** the cellotetraose (again positioned with the +3 subsite to the left), is displayed in the active site channel, and the selected residues close to the +2 subsite (from left to right:G222, N221 and F219) are shown in the mutated forms as M222, S221 and L219. The G222M was made to increase hydrophobicity at the entrance of the active site, while the G222Q mutation (not shown) was predicted to result in hydrogen bonding with the substrate. The F219L and N221S mutations were predicted to generate space for better substrate accommodation.

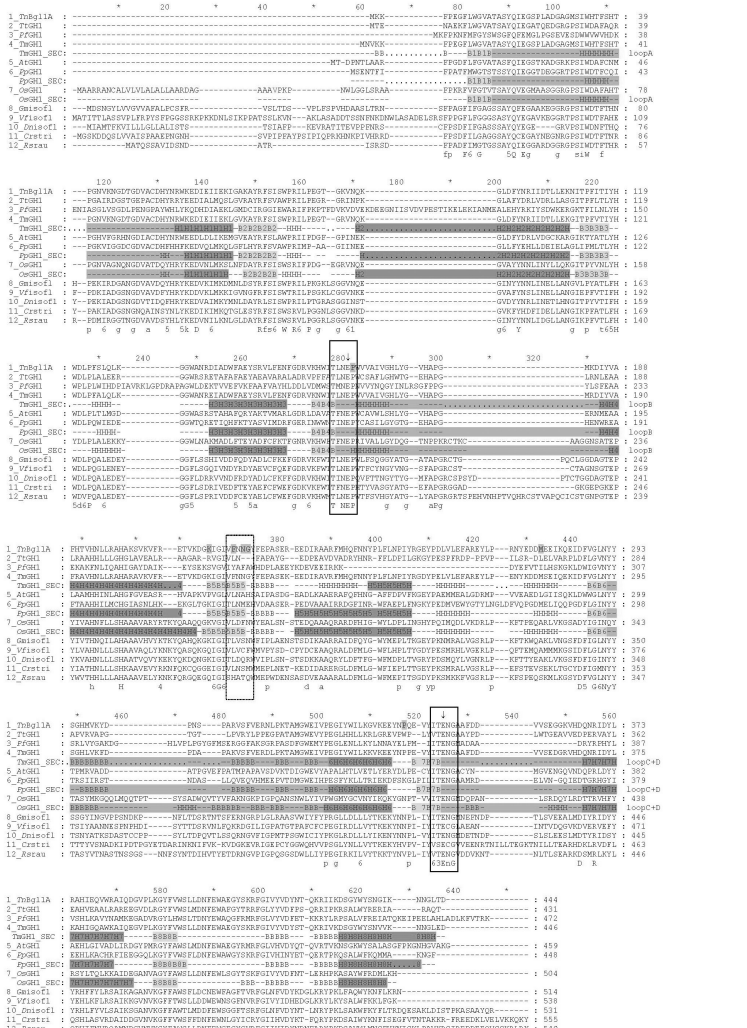


Figure 3 Multiple sequence alignment of GH1 representatives. A multiple sequence alignment of β-glucosidases and flavonoid glucosidases from GH1. *Tn*Bgl1A, *Thermotoga neapolitana* Bgl1A (this work); *Tm*GH1, *Thermotoga maritima* BglA (Q08638); *At*GH1, *Agrobacterium tumefaciens* glucosidase (Q7CV27); *Pp*GH1, *Paenibacillus polymyxa* BglB (P22505); OsGH1, *Oryza sativae* (rice) glucosidase (Q42975); *Gm*GH1, *Glycine max* (soy) isoflavonoid glucosidase (AB259819); *Vfiso*, *Viburnum_furcatum* isoflavonoid glucosidase (AB122081); *Dniso*, *Dalbergia nigrescens* isoflavanoid glucosidase (AY766303); *Crstri*, *Catharanthus roseus* strictosidine β-glucosidase (Q9M7N7); *Rsrau*, *Rauvolfia serpentina* raucaffricine β-glucosidase (Q9SPP9). The region selected for mutagenesis is marked by a dashed box, and the two conserved motifs are boxed. The catalytic residues are indicated by arrows. Mutated residues are shaded in grey. Secondary structures are indicated below structure determined enzymes. Helices and strands of the β/α_{8}-barrel are numbered and indicated in green and yellow, respectively. The sequence parts corresponding to the four loops (A-D) are indicated in cyan. A consensus sequence is shown in bold below the aligned sequences. Completely conserved resides are shown in upper cases, residues conserved in more than 80% of the sequences are shown in lower case, positions with related residues are indicated by numbers (1 = N, D; 3 = S, T; 4 = K, R; 5 = Y, F, W; 6 = I, L, V, M).

(K214R). Four of the five spontaneous mutations were located at the surface of the enzyme, opposite the active site. Replacement of proline with leucine at position 165 (P165L) located next to the catalytic acid/base was found in three of the four triple mutants, along with a surface located hydrophobic residue mutation (M278I).

Expression and purification

The wild-type and mutated variants of *T. neapolitana* β-glucosidase 1A were produced in *Escherichia coli* Tuner (DE3) as described by Turner et al [3]. The expression level was analysed by SDS-PAGE and showed all enzymes to have very similar production patterns (Figure 4A), leaving less than half of the produced protein in a soluble active form, despite use of inducer tuning (reducing the IPTG concentration from 1 to 0.1 mM in a lac-permease deficient strain). All mutated enzymes were screened for activity and found to hydrolyse *p*NPGlc (data not shown). Purification was accomplished by a two step protocol, including a heat treatment (70°C, 30 min) followed by immobilised metal ion affinity chromatography (IMAC) utilizing the C-terminal His-tag, which yielded a purity at or above 90% in all cases (Figure 4B).

Thermostability

Thermostability of all enzyme variants was evaluated by differential scanning calorimetry (DSC). A single transition peak was observed in all cases, which during unfolding resulted in aggregation (also manually observable in the sample after scanning). A repeated scan confirmed the denaturation to be irreversible in all cases (Figure 5). All enzyme variants kept unfolding temperatures above 90°C making them suitable as biocatalysts in applications requiring high thermostability. As expected, thermostability decreased with increasing number of mutations, and the triple mutants showed a decrease in the apparent unfolding temperature (ΔT_m) ranging from 5 - 10°C (Table 2). The single and double mutants did not change unfolding temperature to any large extent.

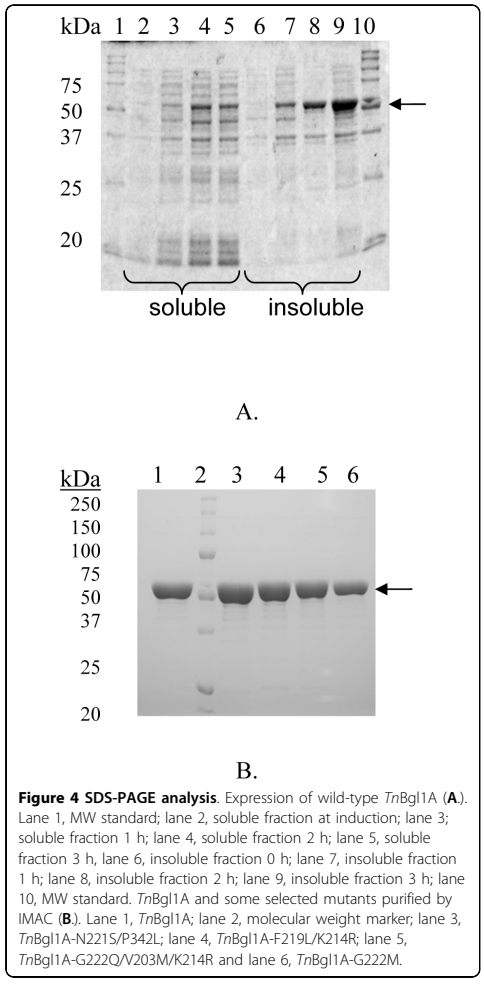


Figure 4 SDS-PAGE analysis. Expression of wild-type *Tn*Bgl1A (**A**.). Lane 1, MW standard; lane 2, soluble fraction at induction; lane 3; soluble fraction 1 h; lane 4, soluble fraction 2 h; lane 5, soluble fraction 3 h, lane 6, insoluble fraction 0 h; lane 7, insoluble fraction 1 h; lane 8, insoluble fraction 2 h; lane 9, insoluble fraction 3 h; lane 10, MW standard. *Tn*Bgl1A and some selected mutants purified by IMAC (**B**.). Lane 1, *Tn*Bgl1A; lane 2, molecular weight marker; lane 3, *Tn*Bgl1A-N221S/P342L; lane 4, *Tn*Bgl1A-F219L/K214R; lane 5, *Tn*Bgl1A-G222Q/V203M/K214R and lane 6, *Tn*Bgl1A-G222M.

Kinetic parameters in *p*NPGlc hydrolysis

The kinetic parameters for hydrolysis were monitored using the model substrate *p*NPGlc at pH 5.6 (Table 3). The pH was set based on previously published data taking into account the pH-range for the highest k_{cat}/K_M determined for the homologous *T. maritima* enzyme (pH optimum 5.8 ± 0.2) [24] combined with the three point screening of *Tn*Bgl1A (pH 3, 5 and 7, showing highest activity at pH 5) for Q4' hydrolysis [3]. The turnover number as well as the K_M values showed changes in some variants. The mutation at position 222 (in G222M, G222Q/P165L/M278I, and G222Q/V203M/K214R) resulted in a significant reduction in the turnover number (k_{cat}) both for the single G222M mutant and the triple mutants including the change G222Q as compared to the *wt*. Only a minor reduction in catalytic efficiency (k_{cat}/K_M), was however seen, as the variations in K_M in most cases counteracted the change in turnover. The triple mutant F219L/P165L/M278I showed a similar reduction in turnover, but this is likely an effect of added mutations, as the double mutant carrying the F219L change (F219L/K214R) showed parameters more similar to the wild type. Variants carrying P165L, located next to the catalytic acid/base, showed in all cases a lower turnover than other

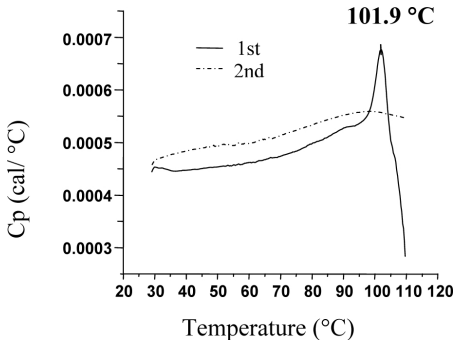


Figure 5 DSC analysis. Representative thermogram of wild-type *Tn*Bgl1A (1 mg mL^{-1} in 20 mM citrate phosphate buffer, pH 5.6).

variants carrying changes in either F219, N221 or G222, but lacking the change at position 165 (Table 3). All three mutants carrying the N221S change instead showed an increased turnover (3.5 ×, 2.4 × and 2.2 ×). The increase in turnover was however combined with a significant increase in K_M for all three mutants (N221S/P342L, N221S/P165L/M278I and N221S) which displayed 3.7 ×, 2.7 × and 1.8 × increases in K_M, respectively. As the increased K_M counteracted the increase in turnover only a minor change in catalytic efficiency was observed. The relatively more pronounced effect on the K_M however raised an interest for trying these variants in deglycosylation reactions of the quercetin glucosides, as glucose at the 3-position generally seems more difficult to hydrolyse with the GH1 enzymes (see below).

Previous work by Lindahl et al, and Yernool et al [17,18], have shown that glucose is acting as an activator in the *wt* enzyme. To assure that this effect is maintained after mutation, the kinetics of the *wt* and N221S/P342L variant in *p*NPGlc hydrolysis were determined in presence of 10 mM glucose. This resulted in an increase of the turnover (k_{cat}= 784 s^{-1} *wt*, k_{cat}= 2310 s^{-1} N221S/P342L) but no significant change in the K_M value (K_M = 0.25 ± 0.02 *wt* , K_M = 0.87 ± 0.08 N221S/P342L) leading to an increased catalytic efficiency in presence of glucose (k_{cat}/K_M = 3150 *wt*, k_{cat}/K_M = 2650 N221S/P342L). The activating effect of glucose is hence maintained to the same extent in the mutated enzyme, and no product inhibition upon glucose release is expected.

Table 2 The melting temperatures of the different expressed constructs measured by differential scanning calorimetry

Enzyme	T_m (°C)	ΔT_m (°C)
wild-type	101.9	
N221S/P342L	100.5	-1.4
N221S/P165L/M278I	95.5	-6.4
F219L/K214R	101.8	-0.1
F219L/P165L/M278I	93.4	-8.5
G222Q/V203M/K214R	96.8	-5.1
G222Q/P165L/M278I	94.8	-7.3
G222M	101.2	-0.7

Table 3 Michaelis-Menten constants

Enzyme	K_M (mM)	k_{cat} (s^{-1})	k_{cat}/K_M (s^{-1} mM^{-1})
	*p*NPGlc-hydrolysis		
Wild-type	0. 24 ± .04	485 ± 31	2000
N221S	0.43 ± 0.11	1170 ± 152	2720
N221S/P342L	0.89 ±0.19	1710 ± 237	1910
N221S/P165L/M278I	0.66 ±0.18	1070 ± 171	1630
F219L/K214R	0.24 ±0.04	594 ± 36	2510
F219L/P165L/M278I	0.18 ±0.006	253 ± 3	1440
G222Q/V203M/K214R	0.10 ± .008	154 ± 3	1470
G222Q/P165L/M278I	0.27 ± .05	181 ± 15	678
G222M	0.17 ±0.03	254 ± 16	1470
	Q4'-hydrolysis*		
Wild type	0.06 ±0.03	7.2 ± 1.2	122
N221S/P342L	0.02 ±0.015	6.2 ± 0.95	281
	Q3-hydrolysis*		
Wild type	0.13 ±0.06	7.2 ± 1.8	55
N221S/P342L	0.05 ±0.02	7.1 ± 1.1	154

*Data from Lindahl et al [17].

For *p*NPGlc hydrolysis measurements were made at pH 5.6, 80°C, and for the quercetin glucoside hydrolysis (*wt* and one mutant) at pH 5.0, 90°C.

Quercetin-glucoside hydrolysis

Although hydrolysis of different quercetin-glucosides by enzymes from GH1 has been reported [10], hydrolysis of the Q3 glucoside appears to be more unusual. In the case of *Tn*Bgl1A *wt*, it has been shown that hydrolysis of Q3 is possible but slow [3,17]. This motivated screening of obtained mutants in Q3 hydrolysis, to monitor improvements in the hydrolysis of this substrate using a fixed concentration of enzyme and substrate (Figure 6). Of the positions selected at the +2 site the mutation N221S led to the highest increase in conversion (from 11 to 35%) of Q3 to Q (Figure 7). Repeated trials with the N221S single mutant, showed that the second mutation P342L in the double mutant, had no major role in this increase (data not shown). The G222M mutation also led to increased Q3-conversion (27%). These improvements may be a result of improved substrate accomodation, and indeed the modelling of Q3 in the N221S variant showed that an additional hydrogen bonding to the substrate (5-OH on the quercetin backbone) can occur via the backbone carbonyl of N291 as a consequence of interaction changes caused by the mutation (Figure 7). Increased hydrophobicity could improve

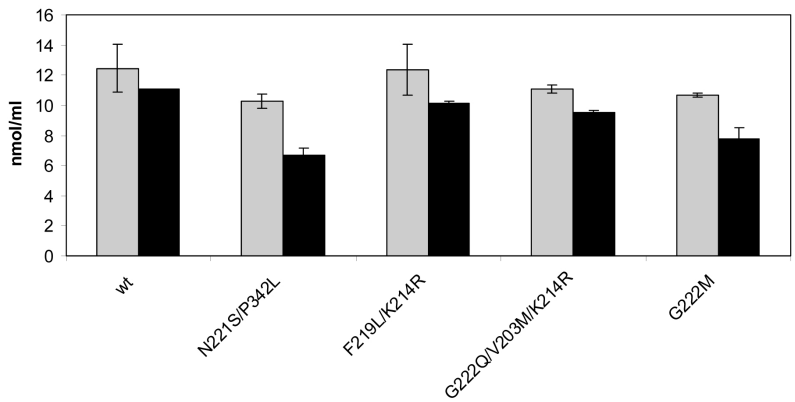


Figure 6 HPLC analysis. Analysis of quercetin-3-glucoside (Q3). Grey bars represent the Q3 concentration at time 0 min, and black bars remaining Q3 concentration after 5 min incubation with the enzyme.

interactions between the substrate and enzyme, and may be the case for G222M which likely has its sidechain pointing into the catalytic cleft. For Q4' all enzyme variants completely hydrolysed the substrate within the reaction time of 10 min.

The (N221S/P342L) mutant resulted in the highest conversion of Q3 and was selected for determination of the kinetic parameters for conversion of the Q4' and Q3 glucosides to Q (Table 3, [17]). Largest improvement was due to a decrease in K_M using both Q4'and Q3, and this was especially pronounced using the Q3 substrate (Table 3), which can be explained by the added substrate interaction (Figure 7).

Discussion

The thermostable β-glucosidase (*Tn*Bgl1A) from *T. neapolitana* has been used as a biocatalyst for conversion of quercetin-glucosides to quercetin [3] extracted using a hot water extraction method, shown to be beneficial from an environmental perspective [17]. It was shown

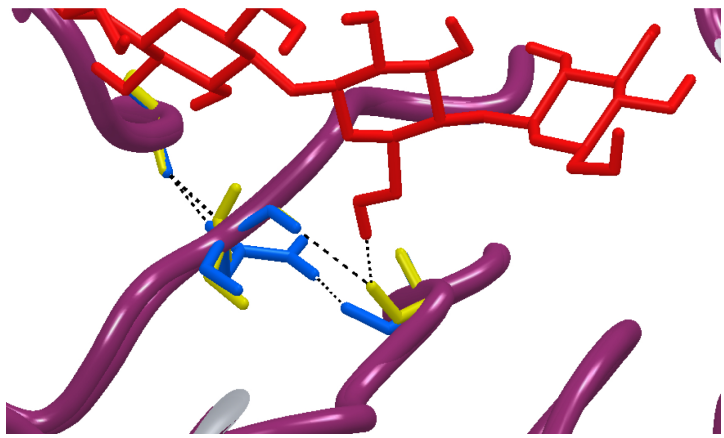


Figure 7 Interactions around position 221. The *Tn*Bgl1A homology model with a cellotetraose ligand with and without residue N221 (blue) mutated to S (yellow). The S mutation allows for recognition of the 6-hydroxyl group of the glucose in the +1 subsite and of the 5-hydroxylgroup of quercetin, by hydrogen bonding to the backbone carbonyl oxygen of residue N291.

that the enzyme completely converted the Q4' species to quercetin, and that also the glucoside connected to the 3-position of the flavonoid backbone (Q3) was converted, although with lower efficiency. This shows that the accessibility to the active site differs between enzymes in GH1, as previous work on flavonoid hydrolysing enzymes in GH1, have shown that enzymes capable of converting the 4'-glucoside (e.g. human β-glucosidase (*h*CBG)) are not capable of hydrolysing the flavonoid 3-glucosides [10,33]. The *Tn*Bgl1A enzyme is thus an interesting target to study by molecular modelling, as well as by site-directed mutagenesis to analyse how changes of amino acids affect the kinetic parameters and the conversion of the quercetin-3-glucoside.

Interactions with inhibitors in the -1 subsite of the homologous *T. maritima* enzyme have been extensively studied [27,28,34,35], and have shown this site to be important for the selection of the sugar to be hydrolysed. In a study on a *Sulfolobus solfataricus* representative of GH1 by Corbett et al, [36] it was for example shown that mutagenesis of substrate interacting residues in the -1 site led to a shift in affinity towards xylose, or from glucose to mannose, dependent on the residue chosen. Here we are instead focusing on interactions closer to the entrance of the active site, and close to the +2 site, which has not been targeted to the same extent. Comparison of the structure of human CBG, which cannot hydrolyse the flavonoid 3-glucoside [10,33] with the model of *Tn*Bgl1A, shows a difference in the shape of the active site (Figure 8A and 8B). It can be clearly seen that *h*CBG forms a regular oval shaped pocket, while the corresponding pocket in *Tn*Bgl1A (as well as in *Tm*BglA used as template for modelling) has a wider conformation, which may facilitate accomodation of the 3-linked substrate. The active site entrance in GH1 is formed by four extended loops [30]. These loops (termed loop A-D) have been defined as being responsible for the overall shape of the aglycone binding pocket, and differences in the conformation of one or more of these loops would likely change the overall shape of the entrance [33]. In *h*CBG, loops B (residue 173-187, *h*CBG numbering) and D (residue 378-385, *h*CBG numbering) are short, which is claimed to result in a small entrance to the pocket. Comparison with *Tn*Bgl1A show that loop length of loop B is the same as in *h*CBG, and despite low sequence similarity in this area both loop A and loop B superimpose very well between the two enzymes. Loop D is however longer in *Tn*Bgl1A, in accordance with the suggestion that this may contribute to a wider active site entrance of this enzyme (Figure 8B and 8C). However, parts of loop C do not superimpose between the two enzymes and in *Tn*Bgl1A this loop is significantly shorter than in the human enzyme (Figure 8C). It is actually close to this shorter loop C that we see a significant widening of the *Tn*Bgl1A active site compared to *h*CBG, indicating that the longer loop in the human enzyme is closing the active site entrance. Moreover, the mutations introduced at the +2 site are located close to loop C in the structure and may aid in a further widening of the active site entrance simplifying binding of Q3.

In addition to interactions at the +2 subsite, the modelling of Q3 into the active site of *Tn*Bgl1A showed that +1 subsite binding was also affected as a consequence of changes in the interactions between residues. The side chain of residue N221 selected here, which upon mutagenesis led to a decrease in K_M for Q3 as well as increased catalytic efficiency, is not pointing towards the catalytic cleft in our model. A close look at the binding interactions of N221 in presence of Q3 shows that this residue is interacting with residues in neighbouring strands of the protein (the backbone carbonyls of Y246 and N291) (Figure 7), and N221 may thus have importance for the shape in that part of the active site. The introduced S221 leads to loss of the interaction with N291, but also a change in the position of the backbone carbonyl of the residue allowing hydrogen bonding to the 5-OH of Q3 (located in a position corresponding to the +1 sub site). Such a bond would explain the observed decrease in K_M. Calculation of the $\Delta\Delta G$ (= $-RT \ln([k_{cat}/K_M]_{wt}/[k_{cat}/K_M]_{mut})$) for Q3 corresponded to a free energy change of 3 kJ/mol. A hydrogen bond interaction to an uncharged amino acid is in the range 2-6 kJ/mol [37] showing that the change in K_M is likely the result of an affinity increase. Changes in turnover are more difficult to explain, and may be caused by conformational changes caused by indirect changes in interactions between residues. Other explanations include changes in the position of the side chain of the neighbouring residue (N220, interacting with the +1 site in *Tn*Bgl1A) pointing into the catalytic site. A similar position of the sidechain is found for the corresponding residue in the homologous human glucosidase (F225 in the *h*CBG structure). The *h*CBG residue F225 was shown by mutagenesis to affect the aglycone specificity [10]. Mutagenesis of the neighboring residue (N221) in *Tn*Bgl1A may change the position of N220 in the active site, or lead to changes in the local environment that promotes an affinity change. Mutation of F219 (also with the side chain located away from the catalytic cleft), preceding N220, did however not lead to any corresponding or significant changes in affinity or turnover.

All spontaneous mutations, except one, involved residues located at the surface of the enzyme. Only P165L of the spontaneous mutations, is located in the active site next to the catalytic acid/base. This change from proline to leucine should introduce more flexibility. The effect of this change on the activity is, based on the

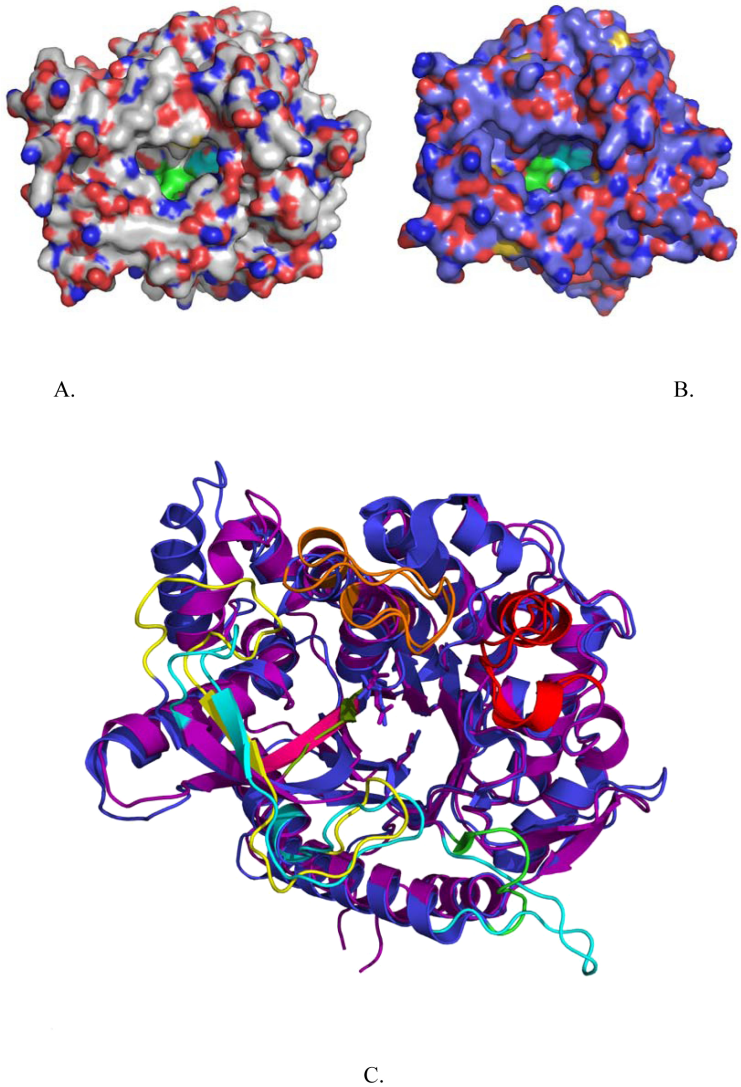


Figure 8 Structural comparison of *h*CBG and *Tn*Bgl1A. Panel A and B show a surface view of *h*CBG (**A**.) and *Tn*Bgl1A (**B**.). The smaller entrance at the *h*CBG active site and wider entrance of the *Tn*Bgl1A active site are clearly visible. The overall similarity of the structures is shown by superimposition (**C**.) of the *Tn*Bgl1A (purple) and *h*CBG (blue) structures. The active site residues are shown as sticks. The fours loops surrounding the active site are indicated. Loops A (red), and B (orange) do not show differences but loop C (yellow in *h*CBG, cyan in *Tn*Bgl1A) and loop D (green in *h*CBG, cyan in *Tn*Bgl1A) presented big differences. Loop C, around the active site entrance, seems bigger in *h*CBG compared to *Tn*Bgl1A. Loop D on the other hand is long for *Tn*Bgl1A compared to the small and compact loop in *h*CBG. The β-strand close to the active site area (pink in *Tn*BglA, green in *h*CBG) was chosen for mutagenesis.

activity data, however not completely clear. It may lead to a minor decrease in activity compared to clones where this mutation is lacking, but its effect appears to be small. The additional surface located mutations, appear to mostly affect stability (slight destabilisation), but it is only in triple mutants that the transition temperature is affected to any large extent (> 5°C).

Conclusions

In conclusion, the mutation study done in this work shows that relatively small residue variations in the enzyme, made at or close to the +2 site, may modify the interactions in the active site, leading to increased substrate interactions as well as conformational changes that allow increased hydrolysis of a sterically differently attached glucose on the quercetin backbone. In addition, effects on the turnover of the introduced mutations were often counteracted by a change in K_M, leading to smaller differences in catalytic efficiency, than in the separated K_M and k_{cat} parameters.

List of abbreviations

DSC: differential scanning calorimetry; G2F: 2-deoxy-2-fluoro-beta-D-glucoside; GH: Glycoside Hydrolase; hCBG: human cytosolic beta-glucosidase; HPLC: high pressure liquid chromatography; K_M: The Michaelis constant; k_{cat}: turnover number; PCR: polymerase chain reaction; *p*NPGlc: para-nitrophenyl-beta-D-glucopyranoside; *p*NP: para-nitrophenol; *Tn*Bgl1A: *Thermotoga neapolitana* beta-glucosidase A from Glycoside hydrolase family 1; Q: quercetin; Q3: quercetin-3-glucoside; Q4': quercetin-4'-glucoside; SDS-PAGE: sodium dodecyl sulfate polyacrylamide gel electrophoresis; 3D: three dimensional.

Acknowledgements

Catherine J Paul is thanked for help with the bioinformatics analysis. Funding from Formas (2006-1346 (Sustain-X-Enz), and 2009-1527 (SuReTech)) is greatly appreciated. The authors wish to thanks Higher Education Commission of Pakistan (HEC) for the scholarship awarded to Samiullah Khan.

Author details

[1]Biotechnology, Dept of Chemistry, Lund University, P.O. Box 124, SE-221 00 Lund, Sweden. [2]Institute of Enzymology, Hungarian Academy of Sciences, H-1113 Budapest, Karolina út 29, Hungary. [3]Organic Chemistry, Lund University, P.O. Box 124, SE-221 00 Lund, Sweden.

Authors' contributions

SK: Took part in the construction and sequence analysis of mutants, took part in enzyme production, purification and kinetics, and in DSC-data collection. Took part in writing the manuscript and commenting on the manuscript.
TP: Took part in constructing mutants. Did the homology modelling and structure analysis together with AS, Took part in writing the manuscript and commenting on the manuscript.
MM: Took part in the construction and sequence analysis of mutants, in the enzyme production and purification, and in DSC-data collection. Took part in writing the manuscript.
SL: Did the hydrolysis and HPLC-analysis of quercetin and quercetin-3-glucosides, Took part in writing and commenting on the manuscript.
AS: Did the major part of the modelling (together with TP), and wrote some parts related to this work in the manuscript.
CT: Took part in the interpretation of HPLC-data. Took part in writing and commenting on the manuscript
ENK: Selected enzyme and mutation strategy. Took part in the bioinformatics analysis and the DSC-data collection. Took part in overall analysis and interpretation of results. Did the overall layout of manuscript, took part in writing and commenting on the manuscript.
All authors read and approved the final manuscript.

Received: 30 March 2010 Accepted: 23 February 2011
Published: 23 February 2011

References

1. Cox PM, Betts RA, Jones CD, Spall SA, Totterdell IJ: **Acceleration of global warming due to carbon-cycle feedbacks in a coupled climate model.** *Nature* 2000, **408**:184-187.
2. Turner P, Mamo G, Nordberg Karlsson E: **Potential and utilization of thermophiles and thermostable enzymes in biorefining.** *Microb Cell Fact* 2007, **6**:9.
3. Turner C, Turner P, Jacobson G, Waldebäck M, Sjöberg P, Nordberg Karlsson E, Markides K: **Subcritical water extraction and β-glucosidase-catalyzed hydrolysis of quercetin in onion waste.** *Green Chem* 2006, **8**:949-959.
4. Lee KW, Lee HJ: **The roles of polyphenols in cancer chemoprevention.** *Biofactors* 2006, **26**:105-121.
5. Boudet AM: **Evolution and current status of research in phenolic compounds.** *Phytochem* 2007, **68**:2722-2735.
6. Rice-Evans CA, Miller NJ, Bolwell PG, Bramley PM, Pridham JB: **The relative antioxidant activities of plant-derived polyphenolic flavonoids.** *Free Radic Res* 1995, **22**:375-383.
7. Lin LZ, Lu S, Harnly JM: **Detection and quantification of glycosylated flavonoid malonates in celery, chinese celery, and celery seed by LC-DAD-ESI/MS.** *J Agric Food Chem* 2007, **55**:1321-1326.
8. McGhie TK, Walton MC: **The bioavailability and absorption of anthocyanins: Towards a better understanding.** *Mol Nutr Food Res* 2007, **51**:702-713.
9. Proteggente AR, Pannala AS, Paganga G, Van Buren L, Wagner E, Wiseman S, Van De Put F, Dacombe C, Rice-Evans CA: **The antioxidant activity of regularly consumed fruit and vegetables reflects their phenolic and vitamin C composition.** *Free Radic Res* 2002, **36**:217-233.
10. Berrin JG, Czjzek M, Kroon PA, McLauchlan WR, Puigserver A, Williamson G, Juge N: **Substrate (aglycone) specificity of human cytosolic β-glucosidase.** *Biochem J* 2003, **373**:41-48.
11. Cantarel BL, Coutinho PM, Rancurel C, Bernard T, Lombard V, Henrissat B: **The Carbohydrate Active-Enzymes Databases (CAZy): an expert resource for Glycogenomics.** *Nucleic Acids Res* 2009, **37**:D233-D238.
12. Price KR, Rhodes MJC: **Analysis of the major flavonol glycosides present in four varieties of onion (*Allium cepa*) and changes in composition resulting from autolysis.** *Sci Food Agric* 1997, **74**:331-339.
13. Hakkinen SH, Karenlampi SO, Heinonen IM, Mykkanen HM, Torronen AR: **HPLC method for screening of flavonoids and phenolic acids in berries.** *Sci Food Agric* 1998, **77**:543-551.
14. Harborne JB: **Plant polyphenols-XIV: Characterization of flavonoid glycosides by acidic and enzymic hydrolyses.** *Phytochem* 1965, **4**:107-120.
15. Nuutila AM, Kammiovirta K, Oksman-Caldentey KM: **Comparison of methods for the hydrolysis of flavonoids and phenolic acids from onion and spinach for HPLC analysis.** *Food Chem* 2002, **76**:519-525.
16. Hertog MGL, Hollman PCH, Venema DP: **Optimization of a quantitive HPLC determination of potentially anticarcinogenic flavonoids in vegetables and fruits.** *J Agric Food Chem* 1992, **40**:1591-1598.
17. Lindahl S, Ekman A, Khan S, Wennerberg C, Borjesson P, Sjoberg PJR, Nordberg Karlsson E, Turner C: **Exploring the possibility of using a thermostable mutant of β-glucosidase for rapid hydrolysis of quercetin glucosides in hot water.** *Green Chem* 2010, **12**:159-168.
18. Yernool DA, McCarthy JK, Eveleigh DE, Bok JD: **Cloning and characterization of the glucooligosaccharide catabolic pathway beta-glucan glucohydrolase and cellobiose phosphorylase in the marine hyperthermophile *Thermotoga neapolitana*.** *J Bacteriol* 2000, **182**:5172-5179.
19. **The National Center for Biotechnology Information advances science and health by providing access to biomedical and genomic information.** [http://www.ncbi.nlm.nih.gov/].
20. Schrodinger, LLC, New York, NY; 2009.

21. French AD, Johnson GP: **Advanced conformational energy surfaces for cellobiose.** *Cellulose* 2004, **11**:449-462.
22. Vallmitjana M, Ferrer-Navarro M, Planell R, Abel M, Ausín C, Querol E, Planas A, Pérez-Pons J-A: **Mechanism of the family 1 β-glucosidase from *Streptomyces* sp: catalytic residues and kinetic studies.** *Biochemistry* 2001, **40**:5975-5982.
23. Ramchuran SO, Holst O, Nordberg Karlsson E: **Effect of postinduction nutrient feed composition and use of lactose as inducer during production of thermostable xylanase in *Escherichia coli* glucose-limited fed-batch cultivations.** *J Biosci Bioeng* 2005, **99**:477-484.
24. Turner P, Holst O, Nordberg Karlsson E: **Optimized expression of soluble cyclomaltodextrinase of thermophilic origin in *Escherichia coli* by using a soluble fusion-tag and by tuning of inducer concentration.** *Prot Express Purif* 2005, **39**:54-60.
25. Laemmli UK: **Cleavage of Structural Proteins during the Assembly of the Head of Bacteriophage T4.** *Nature* 1970, **227**:680-685.
26. He SM, Withers SG: **Assignment of sweet almond β-glucosidase as a family 1 glycosidase and identification of its active site nucleophile.** *J Biol Chem* 1997, **272**:24864-24867.
27. Zechel DL, Boraston AB, Gloster T, Boraston CM, Macdonald JM, Tilbrook DMG, Stick RV, Davies GJ: **Iminosugar glycosidase inhibitors: Structural and thermodynamic dissection of the binding of isofagomine and 1-deoxynojirimycin to β-glucosidases.** *J Am Chem Soc* 2003, **125**:14313-14323.
28. Gloster TM, Meloncelli P, Stick RV, Zechel D, Vasella A, Davies GJ: **Glycosidase inhibition: An assessment of the binding of 18 putative transition-state mimics.** *J Am Chem Soc* 2007, **129**:2345-2354.
29. Soding J: **Protein homology detection by HMM-HMM comparison.** *Bioinformatics* 2005, **21**:951-960.
30. **Bioinformatics Toolkit Max-Planck-Institute for Development Biology.** [http://toolkit.tuebingen.mpg.de/].
31. Chuenchor W, Pengthaisong S, Robinson RC, Yuvaniyama J, Oonanant W, Bevan DR, Esen A, Chen CJ, Opassiri R, Svasti J, Cairns JRK: **Structural insights into rice BGlu1 β-glucosidase oligosaccharide hydrolysis and transglycosylation.** *J Mol Biol* 2008, **377**:1200-1215.
32. Marana SR: **Molecular basis of substrate specificity in family 1 glycoside hydrolases.** *IUBMB Life* 2006, **58**:63-73.
33. Tribolo S, Berrin JG, Kroon PA, Czjzek M, Juge N: **The crystal structure of human cytosolic β-glucosidase unravels the substrate aglycone specificity of a family 1 glycoside hydrolase.** *J Mol Biol* 2007, **370**:964-975.
34. Gloster TM, Macdonald JM, Tarling CA, Stick RV, Withers SG, Davies GJ: **Structural, thermodynamic, and kinetic analyses of tetrahydrooxazine-derived inhibitors bound to β-glucosidases.** *J Biol Chem* 2004, **279**:49236-49242.
35. Gloster TM, Roberts S, Perugino G, Rossi M, Moracci M, Panday N, Terinek M, Vasella A, Davies GJ: **Structural, kinetic, and thermodynamic analysis of glucoimidazole-derived glycosidase inhibitors.** *Biochem* 2006, **45**:11879-11884.
36. Corbett K, Fordham-Skelton AP, Gatehouse JA, Davis BG: **Tailoring the substrate specificity of the β-glycosidase from the thermophilic archaeon *Sulfolobus solfataricus*.** *FEBS Lett* 2001, **509**:355-360.
37. Sierks MR, Bock K, Refn S, Svensson B: **Active site similarities of glucose dehydrogenase, glucose oxidase and glucoamylase probed by deoxygenated substrates.** *Biochem* 1992, **31**:8972-8977.

doi:10.1186/1471-2091-12-11
Cite this article as: Khan *et al.*: Aglycone specificity of *Thermotoga neapolitana* β-glucosidase 1A modified by mutagenesis, leading to increased catalytic efficiency in quercetin-3-glucoside hydrolysis. *BMC Biochemistry* 2011 **12**:11.

Paper II

Exploring the possibility of using a thermostable mutant of β-glucosidase for rapid hydrolysis of quercetin glucosides in hot water

Sofia Lindahl,[a] Anna Ekman,[c] Samiullah Khan,[b] Christina Wennerberg,[b] Pål Börjesson,[c] Per J. R. Sjöberg,[a] Eva Nordberg Karlsson[b] and Charlotta Turner*[a]

Received 11th June 2009, Accepted 28th September 2009
First published as an Advance Article on the web 4th November 2009
DOI: 10.1039/b920195p

The antioxidant quercetin was extracted from yellow onion waste and converted to its aglycone form by a combination of subcritical water extraction and enzymatic hydrolysis. The hydrolytic step was catalysed by a double residue (N221S, P342L) mutant of the thermostable β-glucosidase (*Tn*Bgl1A), isolated from the thermophile *Thermotoga neapolitana* and cloned and produced in *E. coli*. The activity of wt *Tn*Bgl1A was shown to be dependent on the position of the glucosylation on the quercetin backbone, favouring hydrolysis of quercetin-4′-glucoside over quercetin-3-glucoside. The mutated variant of the enzyme harboured a mutation in the +2 sub-site (N221S) and showed increased catalytic efficiency in quercetin-3-glucoside hydrolysis and also to a certain extent hydrolysis of quercetin-4′-glucoside. The mutated enzyme was used directly in yellow onion extracts, prepared by subcritical water extraction, resulting in complete hydrolysis of the glucosylated flavonoids quercetin-3,4′-diglucoside, quercetin-4′-glucoside, quercetin-3-glucoside, isorhamnetin-4′-glucoside and isorhamnetin-3,4′-diglucoside. To complete hydrolysis within five minutes, 3 mg of *Tn*Bgl1A_N221S was used per gramme of onion (dry weight). A life cycle assessment was done to compare the environmental impact of the new method with a conventional solid–liquid extraction-and-hydrolysis method utilising aqueous methanol and hydrochloric acid. Comparison of the methods showed that the new method is preferable regarding primary energy consumption and global warming potential. Another advantage of this method is that handling of toxic chemicals (methanol and HCl) is avoided. This shows that combined subcritical water extraction/enzyme hydrolysis is both a fast and sustainable method to obtain quercetin from onion waste.

Introduction

The agricultural, forestry, and food industries produce tonnes of waste materials and by-products every year, such as onion skin, apple peel, carrot waste and birch bark. In the case of onion, the amount of onion waste annually produced within the European Union is almost half a million tons.[1] The produced waste and by-products are today used in animal feed, or for composting, incineration or anaerobic digestion. However, these waste materials and by-products contain high-value compounds such as antioxidants that can be extracted before the waste reaches its final destination. Waste and by-products from fruits and vegetables are rich in antioxidants such as polyphenols, which have uses as additives in pharmaceuticals, food products and cosmetics. A large group of polyphenolic compounds are flavonoids. Quercetin (Q) is an example of a high-value flavonoid, and can for example be extracted from onion.[2,3] Hence, it is of interest to be able to extract quercetin and other flavonoids from waste and by-products since they possess antioxidative properties – studies have shown that they have positive effects against, for example, cancer[4] and neurodegenerative diseases.[5] These compounds are mainly present in glycosylated forms in fruits and vegetables. In onion, they mainly occur as quercetin-4′-glucoside (Q-4′) and quercetin-3,4′-diglucoside (Q-3,4′) (Fig. 1), but other types of quercetin glucosides as well as the aglycone form (Fig. 1) are also present, but at much lower concentrations.[6]

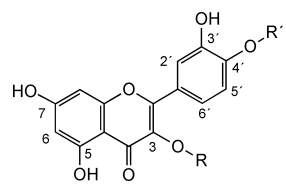


Fig. 1 Chemical structures of quercetin (R, R′ = H), quercetin-3-glucoside (R = glucose and R′ = H), quercetin-4′-glucoside (R = H and R′ = glucose) and quercetin-3,4′-diglucoside (R, R′ = glucose).

There are a number of different methods published on how to extract quercetin and quercetin glycosides from onion, of which the most commonly used technique is solid–liquid extraction with aqueous methanol as extraction solvent.[6,7] The extraction

[a]*Uppsala University, Department of Physical and Analytical Chemistry, Analytical Chemistry, P.O. Box 599, SE-751 24, Uppsala, Sweden. E-mail: Charlotta.Turner@kemi.uu.se*
[b]*Lund University, Department of Biotechnology, P.O. Box 124, SE-221 00, Lund, Sweden*
[c]*Lund University, Department of Technology and Society, P.O. Box 118, SE-221 00, Lund, Sweden*

of quercetin and quercetin glycosides is often combined with acid hydrolysis of the extracted glycosides. To catalyse the hydrolysis reaction, hydrochloric acid at high concentration is commonly used.[7-10] Deglycosylation of the quercetin glycoside mixture obtained after extraction is advantageous for two reasons. Firstly, the glycosides are present in an extract at various concentrations, and by hydrolysing them to the quercetin aglycone, the quantitative determination is simplified. Secondly, the antioxidising power is in general higher for the deglycosylated form,[11] which we have also measured for quercetin species using the DPPH (2,2-diphenyl-1-picrylhydrazyl) method (data not shown). Hence, by hydrolysing the quercetin glucoside mixture, the potential value of the material is increased. A problem with acid-catalysed hydrolysis is that the conditions are harsh, resulting in degradation of compounds such as phenolic acids in the extract, and even degradation of the target molecule quercetin.[8] Another drawback is that the acidity of the sample can cause problems in the analysis step. This is commonly performed by high performance liquid chromatography (HPLC), and the HPLC-column may be damaged by the extremely low pH of the sample. To overcome these problems, enzymatic hydrolysis is an alternative.[3,10,12] β-Glucosidases catalyse hydrolysis of glucose linked to either a carbohydrate or non-carbohydrate moiety, and these enzymes have previously been shown to hydrolyse quercetin glucosides.[3,10,13]

Turner *et al.*[3] have described a subcritical water extraction method, followed by hydrolysis by thermostable β-glucosidase from *Thermotoga neapolitana* that yielded quercetin in the aglycone form from a yellow onion extract. The pressurised liquid extraction with water as a solvent was run at 120 °C and 50 bar, and the following enzymatic hydrolysis in water at 90 °C and pH 5.0. This method gave the same yield as a conventional solid–liquid extraction method, although one limitation in the hydrolysis step was that the enzyme had a low conversion rate of quercetin-3-glucoside (Q-3) to the aglycone form (Q), which was speculated to be due to steric hindrance (Fig. 1). It is well known that water at elevated temperature and pressure is an efficient, less polar solvent, similar to methanol.[14,15] Use of a thermostable enzyme may allow a combination of the extraction with the hydrolysis step, without extensive cooling in-between. Moreover, the hydrolysis reaction will go faster due to increased diffusion rates. Other advantages using subcritical water extraction combined with enzymatic hydrolysis compared to conventional methanol extraction and acid hydrolysis include: (i) lower environmental load, as water is a more sustainable solvent than methanol, and hence no production or destruction facilities are needed; (ii) use of the onion waste as animal feed or as feedstock in biogas production after extraction and hydrolysis is still possible, since only a small fraction of the biomass is removed, and (iii) a significantly shorter total processing time per sample.

However, it is important to not just analyse qualitative parameters, like the exchange of a solvent with another, but to also quantify the environmental load of the method. For this purpose a standardised life cycle assessment (LCA) is an excellent tool allowing quantification of primary energy consumption and emissions.[16] In LCA, fluxes of materials and energy are followed from the extraction of raw materials, production and handling of the product when it has become waste. The flows are quantified and characterised to different environmental impact categories, of which Global Warming Potential (GWP) is one example. LCA can be used for several different purposes such as product and process development and for comparison of the environmental impact of different products and manufacturing processes. LCA is standardised according to ISO 14044.[17]

A long-term aim of our research is to develop a sustainable on-line extraction-biocatalysis system with high efficiency that can be used directly at the site of an industry that produces the waste, or alternatively at a plant taking care of waste products. As a first step, the aim of this study was to optimise the hydrolysis reaction with quercetin glucosides at an analytical level, and to improve the conversion of Q-3 to quercetin. A mutated β-glucosidase (*Tn*Bgl1A_N221S) was utilised in the hydrolysis and compared to the wild type (wt) enzyme, *Tn*Bgl1A. The optimised subcritical water extraction and enzymatic hydrolysis method was subsequently applied to a yellow onion extract, and the environmental performance of the developed processes, including greenhouse gas performance and energy efficiency, was assessed from a life cycle perspective.

Results and discussion

This paper focuses on use of the thermostable β-glucosidase, wt *Tn*Bgl1A, and a mutated variant of the same enzyme (termed *Tn*Bgl1A_N221S), as catalysts for deglucosylation of quercetin glucosides in a high-temperature step, coupled off-line to an environmentally sustainable, pressurised hot water method to extract the quercetin species. Previous work showed that the wild type, *Tn*Bgl1A, and a differently folded glucosidase (*Tn*Bgl3B), both efficiently hydrolysed the major quercetin glucoside in onion extracts, glucosylated at the 4′-position of the quercetin molecule (Q-4′).[3] However, the Q-3 in the extract was not efficiently deglucosylated under the conditions used.[3] To further investigate the possibility of the enzyme to hydrolyse quercetin species glucosylated at the 3-position (Fig. 1), this work involves deglucosylation studies of standards of quercetin glucosides (Q-4′, Q-3,4′, and Q-3, see Fig. 1) over time with a special emphasis on Q-3, using both the wt and a variant of *Tn*Bgl1A mutated at two positions (residues 221 (N→S) and 342 (P→L). Only one (residue 221) of the two mutated residues was judged to be important for specificity, as it is located at the substrate binding site (+2 sub-site), while the other (residue 342) is found on the surface on the reverse side (opposite the active site) of the enzyme.[18] This variant was chosen from a set of mutants, as it both showed the highest turnover of the substrate *para*-nitrophenyl-β-D-glucopyranoside (*p*NPG), and based on a simple screening experiment, also higher conversion of Q-3 to quercetin.[18] Finally, the mutated variant is utilized coupled with subcritical water extraction of onion waste, and the environmental performance of the method is evaluated.

HPLC-UV analysis

In order to follow the hydrolysis process, high selectivity in the detection method is needed. Direct spectrophotometric detection could be employed if (for example) in the hydrolysis of *p*NPG to *p*-nitrophenol (*p*NP), the substrate and products have clearly different absorption maxima. For quercetin and

quercetin glucosides there is no such difference in absorption maxima, and a separation step prior to detection with UV is necessary (although this is time-consuming).

The number of samples to be analysed with HPLC-UV is high, and to minimise the analysis time of each sample an isocratic elution profile was selected, with methanol–water (50 : 50) containing 0.13 M (0.5 vol%) formic acid as mobile phase. The isocratic elution profile used resulted in base-line separation of the analytes: Q-3,4′; Q-3; Q-4′; and quercetin (see Fig. 5b).

Hydrolysis of quercetin glucosides by wt *Tn*Bgl1A

In Fig. 2a,b, the relatively fast conversion of Q-4′ by wt *Tn*Bgl1A is clearly shown. Using a 1 : 100 enzyme:substrate molar ratio, Q-4′ was completely converted to Q within 5 min, while the same conditions for Q-3 resulted in only 70% conversion within 60 min. Enzymatic hydrolysis of the double glucosylated form (Q-3,4′), also showed that the 4′-glucosylation was completely hydrolysed within the 5 min interval, after which the remaining Q-3 was slowly deglucosylated to yield approximately 70% in the aglycone form (Fig. 2c). This shows that hydrolysis of the 4′-glucosylation is clearly favoured by wt *Tn*Bgl1A. Moreover, presence of Q-3 is not disturbing the deglucosylation at the 4′-position, and finally hydrolysis of Q-3 is possible but slow, which may be due to low affinity or steric hindrance for efficient binding of Q-3 in the active site. The homologous human β-glucosidase (hCBG), which has high specificity for Q-4′ and quercetin-7-glucoside, does not hydrolyse 3-linked flavonoid glucosides,[19] hence hydrolysis of this substrate by *Tn*Bgl1A suggests an interesting difference in the catalytic site.

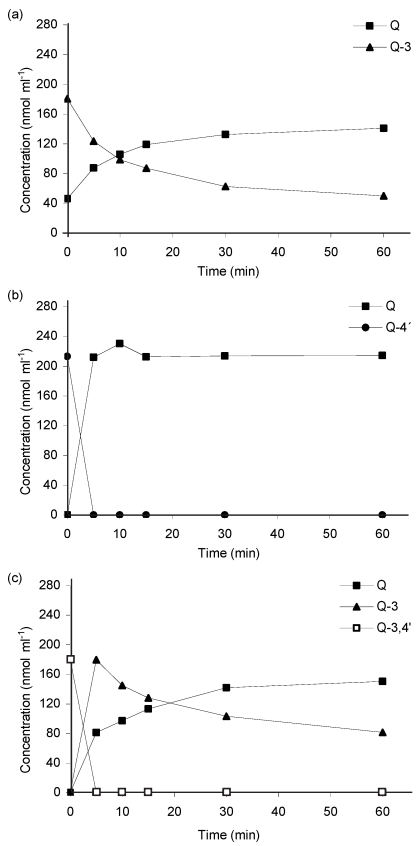


Fig. 2 Catalysis by *Tn*Bgl1A with (a) Q-3, (b) Q-4′ and (c) Q-3,4′. 20 pmol of *Tn*Bgl1A and 200 nmol of substrate were used. Relative standard deviation (RSD) of the concentration values at the different sampling time points ranges between 0.3 and 20.6% ($n = 3$).

Thermal stability and solubility of quercetin species

Self-hydrolysis of quercetin glucosides at high temperature was studied, and could be excluded since incubations of the quercetin glucosides at 95 °C for 1 h (the highest expected temperature for the enzyme catalysis), showed no thermal degradation of the glucosides (data not shown), and the biocatalyst was hence necessary.

A drawback with production of the more potent antioxidant Q (the deglucosylated form), is a decreased water solubility, and initial hydrolysis data showed that this decrease led to losses of the produced quercetin (data not shown). To avoid this, solubility tests were made, and the results revealed that concentrations of maximum 25 nmol ml^{-1} of quercetin were soluble in the buffer at 95 °C. The highest concentration of quercetin species used in the remaining work was thus limited to 17 nmol ml^{-1}, to ensure full solubility of the produced quercetin.

Enzymatic hydrolysis of the different quercetin glucosides

To obtain quantitative data, a number of enzyme:substrate ratios for the wt *Tn*Bgl1A and mutated variant, *Tn*Bgl1A_N221S, were tested, using the substrates, Q-3 and Q-4′, to find optimal ratios for determination of kinetic parameters. The detection method used requires sampling followed by quantification; hence the challenge was to find ratios where the initial reaction is neither too fast nor too slow. Indeed, the obtained results showed that different enzyme:substrate ratios were necessary for different enzyme/substrate combinations: for Q-3, 1 : 250-1250 and 1 : 2500-12500 for wt *Tn*Bgl1A and *Tn*Bgl1A_N221S, respectively, and for Q-4′, 1 : 2500-12500 and 1 : 5000-250000 for wt *Tn*Bgl1A and *Tn*Bgl1A_N221S, respectively. The higher concentration required of the wt was in accordance with the higher turnover and specific activity reported for *Tn*Bgl1A_N221S (1700 s^{-1} at 80 °C) as compared to the wt (485 s^{-1} at 80 °C) for hydrolysis of *p*NPG.[18]

In Table 1 all kinetic data are summarised. As shown in Fig. 1, the glucose on Q-4′ is situated at a position resulting in an extended molecule, while for Q-3 the glucose is in the middle of the flavonoid backbone resulting in a different overall shape, which may be sterically more difficult to accommodate in the active site. This is reflected in the higher K_m of Q-3 than for Q-4′. Furthermore, the K_m (Michaelis–Menten constant) value of *Tn*Bgl1A_N221S is lower for both Q-3 and Q-4′ compared to the wt (Table 1), which indicates that the mutation has increased

Table 1 Kinetic studies of enzymes *Tn*Bgl1A and *Tn*BglA_N221S, and the substrates Q-3 and Q-4′ ($n = 3$). K_m – Michaelis–Menten constant, V_{max} – maximal conversion rate, k_{cat} – turnover number and k_{cat}/K_m – catalytic efficiency.

Enzyme	Substrate	K_m/mM	V_{max}/μmol min^{-1} mg^{-1}	k_{cat}/s^{-1}	(k_{cat}/K_m)/s^{-1} M^{-1}
*Tn*Bgl1A	Q-3	0.129 ± 0.059	8.3 ± 2.1	7.2	55
	Q-4′	0.059 ± 0.026	8.3 ±1.4	7.2	122
*Tn*Bgl1A_N221S	Q-3	0.046 ± 0.020	8.1 ±1.3	7.1	154
	Q-4′	0.022 ± 0.015	7.1 ± 1.1	6.2	281

the affinity for both quercetin substrates. The effect is, however, especially pronounced for Q-3. The smaller size of the serine residue at position 221 may have reduced the steric hindrance of this molecule, leading to a better fit in the binding site resulting in higher catalytic efficiency. The maximal conversion rate (V_{max}) and the turnover number (k_{cat}) were in the same range for all enzyme and quercetin-substrate combinations, but as a consequence of the change in K_m the catalytic efficiency (k_{cat}/K_m) was higher for the mutant (see Table 1).

Despite the improved affinity for Q-3 using *Tn*Bgl1A_N221S, experiments with this variant at 95 °C showed a hydrolysis pattern with decreased deglucosylation rate at increasing incubation times (Fig. 3, 0–30 min). A second addition of the enzyme after 30 min restarted the reaction (Fig. 3, 30–65 min). To understand if this was a result of deactivation or product inhibition, thermal deactivation experiments of the enzyme in the absence of substrate, as well as incubations of the enzyme in the presence of products, were made (see below).

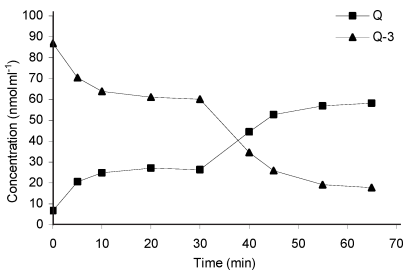


Fig. 3 A second addition of 20 pmol *Tn*Bgl1A_N221S 35 min after the first addition of 20 pmol *Tn*Bgl1A_N221S. RSD of the concentration values at the different sampling time points ranges between 0.4 and 16.9% ($n = 3$).

Thermal stability of the β-glucosidases *Tn*Bgl1A and *Tn*Bgl1A_N221S

The occurrence of irreversible thermal deactivation at 85, 90 and 95 °C was investigated to establish the tolerance of the enzymes to prolonged exposure to these temperatures (Table 2). Almost no enzymatic activity was lost for wt *Tn*Bgl1A during the first 30 min of incubation at 85 and 90 °C. Prolonged incubations up to 24 h fitted a first-order deactivation rate during the complete incubation period, resulting in a deactivation constant (k) of 0.027 h^{-1} (at 90 °C). The half-life of the enzyme ($t_{1/2}$) could thus be calculated to exceed 10 h at 90 °C. A further 5 °C increase in temperature led to significant activity loss, giving a $t_{1/2}$ of 28 min (0.46 h, Table 2). The mutant *Tn*Bgl1A_N221S showed

Table 2 Irreversible thermal deactivation of *Tn*Bgl1A and *Tn*Bgl1A_N221S. k is the deactivation constant and $t_{1/2}$ is the half-life of the enzymes

	*Tn*Bgl1A		*Tn*Bgl1A_N221S	
T/°C	k/h^{-1}	$t_{1/2}$/h	k (h^{-1})	$t_{1/2}$/h
85	0.018	20	0.029	7.0
90	0.027	11	0.040	6.4
95	0.73	0.46	1.51	0.23

a slightly higher deactivation rate, but both the wt and mutant were judged to be stable at temperatures up to 90 °C ($t_{1/2}$ of the mutant was approximately 6 h).

A larger difference was seen at 95 °C, where the deactivation constant (k) for the mutant was doubled, and its half-life reduced to 13 min (0.23 h, Table 2). Thermal transitions by differential scanning calorimetry have shown the apparent transition temperature (T_m) from the folded to unfolded state at 102.5 °C (wt) and 100.5 °C (mutant), respectively.[18] However, inspection of the chromatograms showed that the starting point of the transition occurred around 95 °C (data not shown), explaining the reduced stability at this temperature. This sets a limitation in the time frame applicable for the hydrolysis reaction if a temperature of 95 °C is used, and should only be applied if the overall reaction time is short (≤10 min). This relatively high temperature was used in the kinetic study and the total reaction time was less than 5 min. A temperature of 95 °C was also used in the hydrolysis of the onion extract (with a total reaction time of 5 min) to minimise the temperature difference between extraction and hydrolysis temperatures in the process.

However, in longer studies of the deglucosylation, deactivation will be significant at temperatures above 90 °C. This means that the slow deglucosylation rate of Q-3 seen within the 30 min interval for the wt cannot be a consequence of deactivation, as this was run at 90 °C (Fig. 2), while in the reaction performed with the mutated variant at 95 °C, deactivation should be significant (Fig. 3).

Activation/inhibition of the enzyme

An experiment was set up to evaluate if the mutated enzyme was product-inhibited. In this experiment, glucose and quercetin, respectively, were added to *Tn*Bgl1A_N221S 18 min prior to adding a hydrolysable substrate (in this case Q-4′). Product inhibition acts on the enzyme, and the presence of substrate can hence be used to detect if inhibition has occurred after the pre-incubation. To avoid thermal deactivation, the temperature was set to 90 °C. The results in Fig. 4 show that most of the Q-4′ was hydrolysed after 10 min with either quercetin or the control (just the buffer). Interestingly, incubations in the presence of glucose

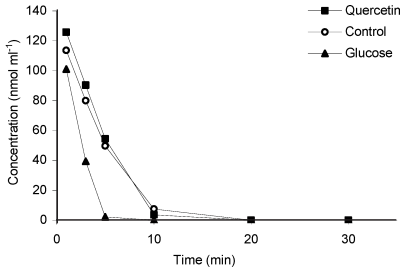


Fig. 4 The effect of incubation of *Tn*Bgl1A_N221S with glucose, quercetin and buffer (control) on the hydrolysis of Q-4′. RSD of the concentration values at the different sampling time points ranges between 0.3 and 6.1% ($n = 3$).

led to complete Q-4′ hydrolysis after just 5 min. The conclusion is that there is no inhibition effect by either quercetin or glucose on the enzyme, but rather that glucose is an activator of the enzyme. An activating effect of glucose on the β-glucosidases from the glycoside hydrolase family 1 (GH1) is in accordance with results reported by other researchers,[20-22] and has previously been reported for wt *Tn*Bgl1A.[18,21] This work shows that the activating effect of glucose also works for the mutated variant of the enzyme.

The activating effect of glucose may also explain the improved conversion of Q-3 to Q seen after the second addition of *Tn*Bgl1A_N221S (Fig. 3), where glucose released after the first enzyme addition is present in the reaction mixture. Currently, one can only speculate that glucose may prevent non-productive binding of the Q-3 or Q-4′ substrate, for example by blocking a high affinity sub-site. To obtain proof of this suggestion, detailed studies on the binding affinity of the subsites of *Tn*Bgl1A will be necessary.

Application to an onion sample

Whole yellow onion was chopped and both the skin and the pulp was mixed and extracted using subcritical water at 120 °C and 50 bar as solvent. The result of the HPLC-UV analysis of the extract was that seven flavonoids were identified in the yellow onion extract – Q-4′, Q-3, Q-3,4′, quercetin, isorhamnetin-4′-glucoside (I-4′), isorhamnetin-3,4′-diglucoside (I-3,4′) and isorhamnetin. Quercetin and the three quercetin glucosides were identified by comparing the retention time and elution order with standards, and isorhamnetin, I-4′ and I-3,4′ were identified based on the results by Turner *et al.*[3] One of the peaks labelled "unknown" is most likely kaempferol, and this assumption is based on a previous study.[3] Q-4′ and Q-3,4′ are the most abundant flavonoids identified in the yellow onion extract and this corresponds well to other publications of flavonoid content in yellow onion.[3,6] Three different amounts of *Tn*Bgl1A_N221S were tested (20 pmol, 100 pmol and 3 nmol) (data not shown) resulting in the following method optimised for 1.2 ml of onion extract corresponding to 0.4 g of onion (50 mg dw); 3 nmol of *Tn*Bgl1A_N221S, reaction temperature of 95 °C, pH 5.0 and 5 min reaction time. Chromatograms of yellow onion extracts before and after enzymatic treatment at 95 °C can be seen in Fig. 5b and 5c. In Fig. 5a the amounts of quercetin and

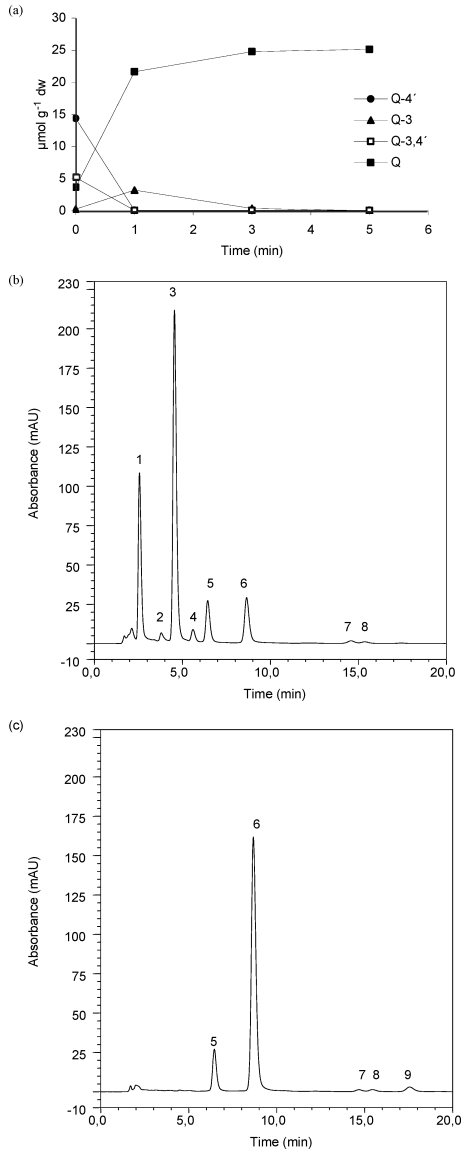


Fig. 5 HPLC-UV analysis of a yellow onion extract before and after enzyme treatment. Concentration change of Q-4′, Q-3, Q-3,4′ and Q during enzymatic reaction (a), RSD of concentration at the different sampling time points ranges between 0.8 and 10.6%, within the calibration curve range ($n = 3$). HPLC-UV chromatogram of yellow onion extract before addition of enzyme (b) and 5 min after addition of *Tn*Bgl1A_N221S (c). Peak labels; 1 (mix of Q-3,4′ and I-3,4′), 2 (Q-3), 3 (Q-4′), 4 (I-4′), 5 (morin (IS)), 6 (Q), 7-8 (unknowns) and 9 (isorhamnetin).

Table 3 Primary energy consumption and emission of greenhouse gases relating to the manufacture of the compounds used in the extraction methods per 10 g of quercetin. Ranges of values are in brackets

Compounds	Primary energy (MJ kg^{-1} solvent/catalyst)	GWP (kg CO_2-eq kg^{-1} solvent/catalyst)	Primary energy/MJ	GWP (kg CO_2-eq)
Methanol[a]	1.5	17	46 (24–83)	550 (270–960)
HCl[b]	27	1.6	86 (46–160)	5.1 (2.7–9.4)
Enzyme[c]	110	9	0.33 (0.22–0.66)	0.027 (0.018–0.054)

[a] Furnander.[23] [b] Boustead.[24] [c] Nielsen *et al.*[25]

Table 4 Primary energy consumption and emissions of greenhouse gases relating to the heating needed in the extraction methods per 10 g of quercetin. Ranges of values are in brackets

Extraction method	Heating/MJ	Extraction/kJ	Total energy/MJ	GWP (kg CO_2-eq)
Traditional method	13 (9.3–25)	70 (60–80)	13 (9.4–25)	1.3 (0.94–2.5)
New method	8.4 (4.2–15)	11 (8–13)	8.4 (4.2–15)	0.85 (0.42–1.5)

quercetin glucosides have been plotted as a function of sampling time points. Almost all the Q-4′ and Q-3,4′ have been hydrolysed after 1 min, while Q-3 concentration has increased, and after 5 min only trace amounts of glucosides are detected (Fig. 5c). Within the few minutes it takes to conduct the reaction of the yellow onion extract, *Tn*Bgl1A_N221S maintains its activity, even though the reaction is done at 95 °C. The quercetin aglycone concentration increased at the same rate as the concentration of glucosides decreased. The increase in Q-3 concentration (after 1 min) shows that the presence of the glucosylation at this position does not limit the activity of the enzyme, but that accommodation of the substrate favours hydrolysis of glucose bound at the 4′-position. By comparing the peaks in Fig. 5b and 5c it is seen that after 5 min of reaction no I-4′ or I-3,4′ peaks are detected, but instead an isorhamnetin peak can be found. This indicates that *Tn*Bgl1A_N221S also catalyses the hydrolysis of isorhamnetin glucosides to isorhamnetin. However, to confirm if it the enzyme catalyses the reaction, a study using isorhamnetin species standards should be done.

Environmental performance

The results from the environmental impact assessment are shown in Table 3 and 4. The traditional method with organic solvents leads to a significantly higher contribution to the GWP than the extraction with subcritical water. This is mainly due to the use of fossil methanol as a solvent, but also due to the larger amount of solvents needed per gramme of onion for the traditional method. Although clear differences in energy consumption can be seen for the two methods, energy for heating gives relatively small contributions to the overall environmental impact. The major input of primary energy is in the manufacture of HCl and methanol.

The environmental impact is summarised in Fig. 6. There are several uncertainities involved in the environmental assessment performed, such as the linear scaling-up of the laboratory methods. Furthermore, depending on the primary fuel used for producing heat (in this case natural gas), the contribution to the GWP will differ. However, despite these uncertainties, there are some "hot-spots" identified, supporting the conclusion that subcritical water extraction is more favourable. These "hot-spots" are the effect of the use of fossil-fuel-derived methanol on GWP, and in addition the effect of the use of HCl on the primary energy consumption.

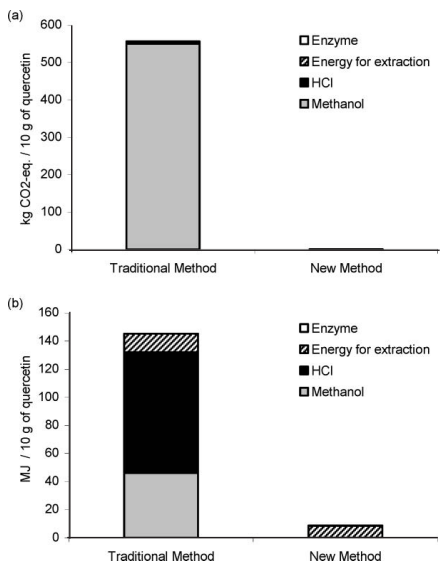


Fig. 6 Total contribution to the GWP expressed as kg CO_2-equivalents per 10 g of quercetin, for the two extraction methods (a), and as total primary energy consumption, expressed as MJ per 10 g of quercetin, for the two extraction methods (b).

Experimental

Materials

Methanol and formic acid were purchased from Merck (Darmstadt, Germany). Quercetin dihydrate, quercetin

3-β-D-glucoside (Q-3), citric acid monohydrate, disodium hydrogen phosphate and *p*NPG were purchased from Sigma-Aldrich (Steinheim, Germany). Quercetin 4′-*O*-β-glucopyranoside (Q-4′) and quercetin 3,4′-di-*O*-β-glucopyranoside (Q-3,4′) were purchased from Polyphenols Laboratories AB (Sandnes, Norway). Ultrapure water (MilliQ) was used at all times.

Enzyme production and purification

The gene encoding wt *Tn*Bgl1A was cloned in pET22b(+), under control of the T7*lac*-promoter, as described by Turner *et al.*[3] The *Tn*Bgl1A_N221S was obtained by site-directed mutagenesis using megaprimer PCR.[18] The fragment encoding the mutated *Tn*Bgl1A was cloned in pET22b(+) (Novagen, Madison, WI, USA) under control of the T7*lac*-promoter, and transformed to *E. coli* Bl21(DE3) for production.

Both enzymes were produced in 2.5 l batch cultivations, in strain Bl21(DE3) after induction with isopropyl-1-thio-β-D-galactopyranoside (IPTG) (2 h), as described by Abou-Hachem *et al.*[26] The cells were harvested by centrifugation (10000*g*, 10 min, 4 °C). Cell pellets were resuspended in buffer (20 mM citrate–phosphate pH 5.6 for a single step heat treatment (80 °C, 30 min, used only for the irreversible deactivation studies)/or binding buffer (20 mM Tris-HCl, 20 mM imidazole, 0.75 M NaCl, pH 7.5) for a two-step procedure involving heat treatment (70 °C, 30 min) and immobilized metal ion affinity chromatography). The cell pellets were disintegrated by sonication (60 W cm^{-2}, 0.5 cycle, for 5 × 3 min) with a UP400S sonicator (Dr. Hielscher, Stahnsdorf, Germany) while cooling in an ice–water mixture. The sonicated extracts were centrifuged at 25000*g* (20 min, 4 °C), and the resulting supernatants stored at 4 °C until purified.

After the heat treatment, the samples were again centrifuged (30 min at 25000*g* and 4 °C) for removal of denatured aggregated proteins. The supernatants obtained after heat treatment at 80 °C were used directly to study irreversible deactivation, while those treated at 70 °C were filtered (0.4 µm pore size) and further purified by immobilised metal-ion affinity chromatography (IMAC).[27]

Enzyme assays

The total protein concentrations were estimated using the bicinchoninic acid method (Sigma-Aldrich, Steinheim, Germany) or the Bradford protein assay (BioRad, Hercules, CA, USA) according to the manufacturers' instructions. Bovine serum albumin (BSA) was used as standard (in concentrations from 0.2 to 1.0 mg ml^{-1}). The absorbance was measured at 562 nm (bicinchoninic acid method) and 595 nm (Bradford method) using an Ultrospec 1000 (Pharmacia, Uppsala, Sweden). To estimate the purity of the enzyme, SDS-PAGE according to Laemmli[28] was used.

For thermal inactivation tests, the β-glucosidase activity was measured as the release of *p*NPG after 5 min incubations at 85 °C. To 960 µl preheated substrate (2.92 mM *p*NPG dissolved in citrate–phosphate buffer pH 5.6), 40 µl of enzyme solution was added, and the assay was performed as described in Turner *et al.*[29] Absorbance was measured at 405 nm using Ultrospec 1000 (Pharmacia). One Unit (U) for this substrate corresponds to the amount of enzyme that will release 1 µmol *p*-nitrophenol per minute under the described conditions.

Enzymatic hydrolysis of quercetin standards

In an initial activity screening, 200 nmol sample of Q-3, Q-4′ and Q-3,4′ dissolved in methanol (in triplicate) were evaporated, 1.0 ml of 100 mM citrate–phosphate buffer, pH 5.0, was added and the vials were heated at 90 °C until substrate dissolved. A 50 µl fraction was collected and added to 450 µl of mobile phase composed of methanol–water (50 : 50) and 0.13 M (0.5 vol%) formic acid (0 min fraction in Fig. 2). The reaction took place at 90 °C and the reaction was started by adding 20 pmol of wt *Tn*Bgl1A. 50 µl fractions were collected and added to 450 µl of mobile phase 1–60 min after addition of enzyme. Samples were analysed by HPLC-UV. Collection of fractions and enzyme addition was done with a 50 µl GC-syringe. The fraction-collection procedure and enzyme addition were similar for all experiments described below, unless otherwise noted.

Thermal stability of the quercetin standards

Samples (150 nmol) of Q-3, Q-4′ and Q-3,4′ dissolved in methanol were evaporated, 1.5 ml of 100 mM citrate–phosphate buffer was added (triplicate samples for each glucoside) and the vials were heated at 95 °C until substrate was dissolved. Fractions were collected 10, 30 and 60 min after the first fraction collection. Fractions were collected and diluted as above. The samples were analysed by HPLC-UV.

Enzyme kinetics of the hydrolysis of Q-3 and Q-4′

To determine the kinetics of the hydrolysis reactions, 50–250 nmol samples of Q-3 and Q-4′ dissolved in methanol were evaporated and re-dissolved as described above, and the volume of buffer used was 1.5 ml. Fractions were collected at 0, 1, 3, and 5 min after enzyme addition. The enzyme amounts used were 200 pmol of wt *Tn*Bgl1A and 20 pmol of *Tn*Bgl1A_N221S to Q-3 and 20 pmol of wt *Tn*Bgl1A and 10 pmol of *Tn*Bgl1A_N221S to Q-4′. Samples were analysed by HPLC-UV. The enzyme activity using Q-3 and Q-4′ is expressed in Units (U). One U corresponds to the amount of enzyme required to release 1 µmol quercetin per minute under the described conditions. Obtained data was analysed using Enzpack (Biosoft, UK), using a time-range allowing fit to the Michaelis–Menten curve. Kinetic constants were calculated using the Wilkinson non-linear regression method, supplied in the program.

Repeated enzyme additions at the applied conditions

A sample of Q-3 (triplicate of 150 nmol) in methanol was evaporated and re-dissolved as above at 95 °C. Fractions were collected at 0, 5, 10, 20 and 30 min after addition of 20 pmol *Tn*Bgl1A_N221S. 35 min after the first enzyme addition, the same amount of *Tn*Bgl1A_N221S was added and fractions were collected 40, 45, 55 and 65 min after the first enzyme addition. Samples were analysed by HPLC-UV.

Thermal stability of the enzymes in absence of substrate

Thermal stability of the enzymes was measured as irreversible deactivation. The heat-treated wt (*Tn*Bgl1A, 1.5 mg ml^{-1}, 29 μM) and mutant (*Tn*Bgl1A_N221S, 2.0 mg ml^{-1}, 38 μM) glucosidase were aliquoted into microcentrifuge tubes. These samples were incubated at 85, 90 and 95 °C, at time intervals ranging from 0 to 24 h. During the first 30 min, duplicate samples were withdrawn every 5 min, then samples were taken after 1, 2, 3, 4, 5, 18 and 24 h, and were chilled on ice (at least 30 min), centrifuged (13 000*g*, 30 min, 4 °C) and stored at 4 °C until activity analysis was performed. The inactivation kinetics of the different constructs was analysed assuming a first-order reaction rate. For a constant temperature:

$$A/A_0 = e^{-kt} \quad (1)$$

$$\ln A - \ln A_0 = -kt \quad (2)$$

where A_0 is the initial activity expressed in U ml^{-1}, A is the activity at time t, expressed as above, t (h) is the time, and k is the inactivation rate constant (h^{-1}). The linear regression line of the natural logarithm of the residual activity *versus* incubation time were obtained using Excel (Microsoft Office 2003). The half-life ($t_{1/2}$) at the respective temperature was obtained at $A = 0.5A_0$.

Product inhibition studies

Quercetin samples (triplicates of 100 nmol each) in methanol were evaporated and re-dissolved in 750 μl of 100 mM citrate–phosphate buffer pH 5.0, and the vials were heated at 90 °C until the quercetin had dissolved. Glucose (triplicates of 100 nmol) dissolved in 750 μl 100 mM citrate–phosphate buffer, pH 5.0, and controls containing 750 μl of 100 mM citrate–phosphate buffer pH 5.0, (triplicate) were also heated at 90 °C, and treated in the same way as the quercetin samples. *Tn*Bgl1A_N221S (10 pmol) was added to the vials and the samples were heated at 90 °C. Then, 18 min after addition of enzyme, pre-heated 100 nmol of Q-4′ dissolved in 750 μl of 100 mM citrate–phosphate buffer, pH 5.0, was added to the vials. Fractions were collected as above, 1, 3, 5, 10, 20 and 30 min after addition of Q-4′. The first point in Fig. 4 is the 1 min fraction. Samples were analysed by HPLC-UV.

Subcritical water extraction

15–18 g of onion samples, in triplicate, were weighed into 33 ml stainless steel extraction cells containing a cellulose filter at the bottom. Extractions were performed on a Dionex ASE®-200 subcritical water extraction system (Dionex, Sunnyvale, CA, USA) using water as the only solvent as described by Turner *et al.*[3] The pressure was set to 50 bar and the initial heating lasted for 5 min. Extraction temperature was 120 °C and extraction time 15 min (three extraction cycles of 5 min each). Purging between extractions was performed with nitrogen. The extract was collected in 100 ml clear glass vials, cooled down to room temperature and diluted with degassed water to a total volume of 50 ml.

Enzymatic hydrolysis of onion extract

A 1.2 ml aliquot of onion extract, prepared by subcritical water extraction, was mixed with 300 μl of 100 mM citrate–phosphate buffer, pH 5.0, and heated at 95 °C until the sample components dissolved. Fractions were collected at 0, 1, 3, 5 and 10 min after addition of 3 nmol of *Tn*Bgl1A_N221S. Samples were analysed by HPLC-UV.

HPLC analysis

HPLC-UV analysis was performed using the chromatographic system UltiMate 3000 from Dionex (Germering, Germany). An Agilent Zorbax SB-C18 column (100 × 2.1 mm, 3.5 μm) was used for isocratic separation with a methanol–water (50 : 50) and formic acid (0.13 M, 0.5 vol%) mobile phase at a flow rate of 0.15 ml min^{-1}. The injection volume was 10 μl and detection was accomplished at 350 nm. Quantification of quercetin and its glucosides was performed using a five-point calibration curve of a quercetin dihydrate, Q-3, Q-4′ and Q-3,4′ standards at concentrations between 0.5 and 25 μg ml^{-1}. Each vial taken to analysis had a total volume of 500 μl.

Environmental impact assessment

The global warming potential (GWP) and primary energy consumption for the new and a traditional extraction method were assessed from a gate-to-gate perspective.[16,30] The calculations are based on a functional unit of 10 g of quercetin. The new, subcritical water extraction method is described by Turner *et al.*[3] (see Subcritical water extraction, above), while the enzymatic reaction method is described in this study (see Enzymatic hydrolysis of onion extract, above). The traditional method using methanol extraction and HCl-catalysed hydrolysis (80 °C for 2 h) is described by Hertog *et al.*[7] and Nuutila *et al.*[8] Extractions were scaled up linearly to produce 10 g of quercetin. The dry matter content in onion is around 13.5% and the quercetin content in onion is, on average, 13 mg g^{-1} dry matter, varying between 7 and 25 mg g^{-1} (see Table 5).[3]

Energy demand is calculated for heating and heat losses during extraction according to Welty *et al.*[31] Extraction is set to be carried out in vessels of stainless steel insulated with mineral wool. Production of extraction equipment and generation of pressure are not considered in this study. A natural gas burner with 60% efficiency is set to be used for heating. Emissions due to

Table 5 Data for the extraction methods both producing 10 g of quercetin. Traditional method: 80 °C for 2 h. New method: extraction at 120 °C for 15 min and biocatalysis at 95 °C for 5 min. Ranges of values are in brackets

Extraction method	Onion/kg	Amount of solvent/kg	Amount of catalyst/kg	Total volume (dm^3)
Traditional method	5.7 (3.0–10.6)	Methanol: 31 (16–56)	HCl: 3.2 (1.7–5.9)	87 (41–150)
New method	5.7 (3.0–10.6)	Water: 13 (6.5–23)	Enzyme: 0.003 (0.002–0.006)	21 (9.4–33)

production and combustion of natural gas are from Uppenberg *et al.*[32] The environmental impact of water is assumed to be insignificant.

Conclusions

In this paper, we have developed analytical methods to study the kinetics of a thermostable β-glucosidase (wt *Tn*Bgl1A), and a mutant of the same enzyme (*Tn*Bgl1A_N221S) towards Q-3 and Q-4′. We have also developed an off-line biocatalysis method, in which the hydrolysis of all quercetin glucosides present in yellow onion extract is completed within 5 min at 95 °C using only 3 nmol of enzyme per 1.3 μmol of quercetin species. This study shows that *Tn*Bgl1A_N221S has higher activity towards both Q-3 and Q-4′ compared to the wt *Tn*Bgl1A, *i.e.* a mutation of residue 221 at the +2 binding site of the enzyme does significantly affect substrate affinity, but this effect is higher for the Q-3 substrate. Another conclusion that can be drawn is that deactivation is significant at 95 °C, and to avoid this, the temperatures should be kept below 95 °C, if reactions are more than a few minutes long. To increase the hydrolysis rate of quercetin glucosides, addition of glucose can be considered. Future aspects include immobilisation of *Tn*Bgl1A_N221S onto a support material allowing recycling of the enzyme, and the use of an on-line flow system of hot water extraction and enzymatic hydrolysis of quercetin and quercetin glucosides from yellow onion waste and other potential by-products and waste. The immobilisation of the enzymes may also improve the thermostability of the enzymes.

From an environmental point of view, extraction with subcritical water followed by an enzymatic hydrolysis is to be preferred regarding primary energy consumption and global warming potential. Shorter extraction time and smaller volumes makes the process more energy-efficient, although a higher temperature and pressure have to be generated. Furthermore, avoiding the use of organic solvents from fossil origin, such as methanol, makes the subcritical water extraction more resource efficient and favourable from a greenhouse gas perspective. Handling of large quantities of toxic chemicals, such as HCl and methanol, is also avoided.

Abbreviations

Q	Quercetin
Q-3	Quercetin-3-glucoside
Q-4′	Quercetin-4′-glucoside
Q-3,4′	Quercetin-3,4′-diglucoside
*p*NPG	*p*-Nitrophenyl-β-D-glucopyranoside
*p*NP	*p*-Nitrophenol
BSA	Bovine serum albumin
GWP	Global warming potential
LCA	Life cycle assessment
I-4′	Isorhamnetin-4′-glucoside
I-3,4′	Isorhamnetin-3,4′-diglucoside
IPTG	Isopropyl-1-thio-β-D-galactopyranoside
dw	Dry weight
DPPH	2,2-Diphenyl-1-picrylhydrazyl
RSD	Relative standard deviation
IMAC	Immobilised metal-ion affinity chromatography
wt	Wild type
T_m	Transition temperature
GH1	Glycoside hydrolase family 1

Acknowledgements

The Swedish Research Council for Environment, Agricultural Sciences and Spatial Planning (FORMAS, 2006-1346); the Swedish Research Council (VR, 2006-4084 and 2006-6048); the Swedish International Development Cooperation Agency (SIDA, SWE-2005-473); the Swedish Foundation for Strategic Research (SSF, 2005 : 0073/13) and the Foundation for Strategic Environmental Research (MISTRA, Greenchem) are greatly acknowledged.

References

1 A. Schieber, F. C. Stintzing and R. Carle, *Trends Food Sci. Technol.*, 2001, **12**, 401–413.
2 M. G. L. Hertog, P. C. H. Hollman and M. B. Katan, *J. Agric. Food Chem.*, 1992, **40**, 2379–2383.
3 C. Turner, P. Turner, G. Jacobson, K. Almgren, M. Waldeback, P. Sjöberg, E. N. Karlsson and K. E. Markides, *Green Chem.*, 2006, **8**, 949–959.
4 A. Murakami, H. Ashida and J. Terao, *Cancer Lett.*, 2008, **269**, 315–325.
5 K. Ono, T. Hamaguchi, H. Naiki and M. Yamada, *Biochim. Biophys. Acta, Mol. Basis Dis.*, 2006, **1762**, 575–586.
6 K. R. Price and M. J. C. Rhodes, *J. Sci. Food Agric.*, 1997, **74**, 331–339.
7 M. G. L. Hertog, P. C. H. Hollman and D. P. Venema, *J. Agric. Food Chem.*, 1992, **40**, 1591–1598.
8 A. M. Nuutila, K. Kammiovirta and K. M. Oksman-Caldentey, *Food Chem.*, 2002, **76**, 519–525.
9 S. H. Hakkinen, S. O. Karenlampi, I. M. Heinonen, H. M. Mykkanen and A. R. Torronen, *J. Sci. Food Agric.*, 1998, **77**, 543–551.
10 J. B. Harborne, *Phytochemistry*, 1965, **4**, 107–120.
11 Y. K. Do, J. M. Kim, S. M. Chang, J. H. Hwang and W. S. Kim, in *International Conference on Biorefinery*, Elsevier Science Bv, Beijing, JAPAN, 2007, pp. 173-178.
12 F. Siewek and R. Galensa, *J. Chromatogr., A*, 1984, **294**, 385–389.
13 N. Lambert, P. A. Kroon, C. B. Faulds, G. W. Plumb, W. R. McLauchlan, A. J. Day and G. Williamson, *Biochim. Biophys. Acta, Protein Struct. Mol. Enzymol.*, 1999, **1435**, 110–116.
14 E. Ibanez, A. Kubatova, F. J. Senorans, S. Cavero, G. Reglero and S. B. Hawthorne, *J. Agric. Food Chem.*, 2003, **51**, 375–382.
15 K. Hartonen, K. Inkala, M. Kangas and M. L. Riekkola, *J. Chromatogr., A*, 1997, **785**, 219–226.
16 A. Azapagic, *Handbook of Green Chemistry and Technology*, Blackwell Science, Oxford, 2002, 62-85.
17 I. S. Organization, *Enviromental management - Life cycle assesment - Requirements and guidelines*, ISO 14044, 2006.
18 S. Khan, T. Pozzo, C. Paul, S. Lindahl, C. Turner, and E. Nordberg Karlsson, 2009, unpublished work.
19 J. G. Berrin, W. R. McLauchlan, P. Needs, G. Williamson, A. Puigserver, P. A. Kroon and N. Juge, *Eur. J. Biochem.*, 2002, **269**, 249–258.
20 R. M. Wright, M. D. Yablonsky, Z. P. Shalita, A. K. Goyal and D. E. Eveleigh, *Appl. Environ. Microbiol.*, 1992, **58**, 3455–3465.
21 D. A. Yernool, J. K. McCarthy, D. E. Eveleigh and J. D. Bok, *J. Bacteriol.*, 2000, **182**, 5172–5179.
22 J. A. Perezpons, X. Rebordosa and E. Querol, *Biochim. Biophys. Acta, Protein Struct. Mol. Enzymol.*, 1995, **1251**, 145–153.
23 Å. Furnander, *Life Cycle Assessment of Dimethyl Ether as a Motor Fuel*, Chalmers University of Technology, Gothenburg, 1996.
24 I. Boustead, *Eco-Profiles of the European Plastic Industry - Hydrogen Chloride*, PlasticEurope - Association of Plastics Manufacturers, 2005.

25 P. H. Nielsen, K. M. Oxenboll and H. Wenzel, *Int. J. Life Cycle Assess.*, 2007, **12**, 432–438.
26 M. Abou-Hachem, F. Olsson and E. N. Karlsson, *Extremophiles*, 2003, **7**, 483–491.
27 P. Turner, D. Svensson, P. Adlercreutz and E. N. Karlsson, *J. Biotechnol.*, 2007, **130**, 67–74.
28 U. K. Laemmli, *Nature*, 1970, **227**, 680–685.
29 P. Turner, A. Pramhed, E. Kanders, M. Hedstrom, E. N. Karlsson and D. T. Logan, *Acta Crystallogr., Sect. F: Struct. Biol. Cryst. Commun.*, 2007, **63**, 802–806.
30 A. E. V. Petersson, L. M. Gustafsson, M. Nordblad, P. Borjesson, B. Mattiasson and P. Adlercreutz, *Green Chem.*, 2005, **7**, 837–843.
31 J. R. Welty, C. H. Wicks, R. E. Wilson and G. Rorrer, *Fundamentals of momentum, Heat and Mass Transfer*, John Wiley & Sons, Inc., 2001, 201-218.
32 S. Uppenberg, M. Almemark, M. Brandel, L.-G. Lindfors, H.-O. Marcus, H. Stripple, A. Wachtmeister and L. Zetterberg, *Miljöfaktabok för Bränslen Del 2*, IVL Swedish Enviromental Research Institute Ltd., 2001.

Paper III

Immobilization of thermostable β-glucosidase variants on acrylic supports for biocatalytic processes in hot water

Samiullah Khan[a], Sofia Lindahl[b], Charlotta Turner[b] and Eva Nordberg Karlsson[a*]

[a] *Biotechnology, Dept of Chemistry, Lund University, P.O. Box 124, SE-221 00 Lund, Sweden*

[b] *Centre for Analysis and Synthesis, Dept of Chemistry, Lund University, P.O. Box 124, SE-221 00 Lund, Sweden*

[*]*To whom correspondence should be addressed*

E-mail:*sami.khan@biotek.lu.se;sofia.lindahl@organic.lu.se;charlotta.turner@organic.lu.se;*

eva.nordberg_karlsson@biotek.lu.se (corresponding author)

Tel: +46 46 222 4626
Fax: +46 46 222 4713

Abstract
Two variants of the thermostable β-glucosidase *Tn*Bgl1A (*wt* and N221S/P342L) from *Thermotoga neapolitana* were immobilized on acrylic supports (Eupergit® C, Eupergit® C250L, and cryogel) and evaluated at conditions close to the boiling point of water. Thermo-gravimetric analysis showed the supports to be stable < 250°C. Both *wt* and N221S/P342L were covalently bound to oxirane-groups respectively via glutaraldehyde spacers, and for coupling reactions 26 Lys and 20 Ser/Thr were surface-located. Immobilized enzymes were active on all supports in the temperature range 80 - 95°C, but the observed specific activity was low (≤19 U mg^{-1}) using the cryogel. More than 91% of the initial activity was maintained after ten times recycling, and the same was recovered after 3 months storage at 4°C for Eupergit® supports by simply incubating the preparation with bovine serum albumin. No storage loss was detectable on cryogels. The glutaraldehyde spacer improved activity on cryogels, but not on Eupergit® supports. Immobilization on Eupergit® C250L yielded the highest observed specific activity (254 U mg^{-1} for N221S/P342L) in a procedure including blocking of free oxirane-groups by BSA. This biocatalyst was used for on-line hydrolysis of quercetin-glucosides in a yellow onion extract at 80°C, proving the immobilized biocatalyst to be promising in on-line systems for extraction and hydrolysis using hot pressurised water.

Key Words: Immobilization, glycoside hydrolase family 1, Eupergit® C, Eupergit® C250L, Cryogel

1. Introduction

Biocatalysts are attractive alternatives to chemical catalysts in industry for many reasons, including high substrate specificity, an ability to operate at high temperature (using extremophilic enzymes), and production of stereo-specific products. Consequently, much research has been undertaken to develop enzyme-mediated reactions [1, 2]. Enzymes have the potential to make biomass processes environmentally friendly, being renewable and biodegradable (compared to chemical catalysts) and able to catalyze specific conversions. In order to make it economically feasible to use enzymes on a large scale, a recycling system for the catalyst can be developed which will keep down costs as well as simplify separation of the biocatalyst from the obtained products.

In many processes, it can be an advantage to use enzymes that are operationally stable at high temperatures as many processes require a rise in temperature for more efficient substrate release/conversion [3]. When renewable resources such as agricultural crops or wood are utilized (for extraction of high value products, or intermediates for direct bioconversion into chemicals, commodities and fuels) thermostable enzymes have an obvious advantage as higher temperatures often promote better enzyme penetration and cell-wall disorganization of the raw materials [4-6]. Much research has been aimed at investigating the use of thermostable enzymes in applications where a rise in temperature could be beneficial. In Turner *et al* [3] thermostable β-glucosidases from *T. neapolitana* were for example utilized directly on a yellow onion extract (at 80-90° C) allowing conversion of glucosylated species of quercetin from yellow onion waste to a uniform aglycone form. Quercetin is an antioxidizing compound with electron scavenging properties that may slow down or prevent the development of cancer [7]. Variants of the same enzyme have also been developed to further increase conversion efficiency of quercetin glucosides [8]. β-glucosidases hydrolyze the glycosidic bond between two glucose residues or between a glucose residue and non-carbohydrate aglycone, and the latter reaction is utilized in the quercetin conversions. The possibility to use immobilized enzymes in these applications is not yet investigated, and is the focus of the current work.

Immobilization, or attachment of the enzymes to inert carrier supports, allow them to be held in place throughout the reaction, facilitating enzyme recycling through multiple cycles of batch wise hydrolysis and during use in flow systems. In addition, enzyme immobilization frequently leads to improved thermostability, improved shelf-life and resistance to shear inactivation,

resulting in increased process robustness, and a longer duration of the activity of the biocatalyst [9, 10].

Three methods for immobilization are traditionally used: adsorption, entrapment or crosslinking. Immobilization via crosslinking, by covalent bonds, is usually the method of choice, if harsh conditions, such as high temperature are used in the following process. Covalent crosslinking leads to minimal protein leaching from the supports, absence of a barrier between the enzyme and the reaction medium and rational control of the enzyme properties may be achieved [11]. Epoxy (oxirane)-activated supports allow easy immobilization of enzymes under mild conditions at both laboratory and industrial scale, and are highly stable during storage, chemically and mechanically, over a pH range from 0 to 14. They form very stable covalent multipoint attachment with amino, hydroxyl, thiol and phenolic groups of amino acid side chains on the enzyme surface [12] and capping of the remaining epoxy groups can be made to prevent undesirable protein–support interactions using a variety of reagents (BSA, mercaptoethanol, ethanolamine, glycine, etc.) [11, 13]. The epoxy group can be used directly for covalent attachment. However, because this sometimes results in steric hindrance of the enzyme, a connecting spacer can be added, for example by modification of the support to an amine form by reacting the epoxy group with reagents such as ethylenediamine followed by reaction with a dialdehyde (e.g. glutaraldehyde) to produce an aldehyde activated support [14].

In this investigation the thermostable β-glucosidase and a variant of the same enzyme from *T. neapolitana* were covalently immobilized with, as well as without, a spacer arm, on three different epoxy-activated polyacrylic matrices (Eupergit® C, Eupergit® C250L and cryogel), and the properties of the immobilized derivatives were studied. Thermostability of the support materials, reuse of the immobilized enzymes, long term storage stability, and post-immobilization stabilization were studied, and finally their application in flavonoid hydrolysis was tried.

2. MATERIALS AND METHODS

2.1. Chemicals

All the chemicals were of pro-analysis grade from Merck Eurolabs (Darmstadt, Germany) unless otherwise stated.

2.2. Preparation of support materials

2.2.1. Epoxy-activated supports. The two epoxy-group-containing matrices were Eupergit® C (purchased from Sigma (Sigma Aldrich)) and Eupergit® C250L (kindly provided by Evonik Röhm GmbH).

Cryogels containing epoxy-group were prepared from monomers (1.316 g acrylamide, 0.354 g N,N methylene bisacrylamide and 0.28 mL of allyl glycidyl ether) dissolved in deionized water (final concentration 10%) and the solutions were degassed under vacuum to remove dissolved oxygen. The samples were kept in an ice bath and tetramethylethylenediamine (26 µL) was added and the reaction was initiated by addition of an aqueous solution of ammonium persulphate (19.3 mg/mL). Portions of 0.5 mL of the reaction mixture was poured into cold glass tubes (diameter 7 mm) and was frozen at -12˚C, for 16 h in ethanol using a cryostat (Lauda RK20KP, Königshofen, Germany). The samples were defrosted at room temperature, washed with deionized water and stored at 4˚C for further use [15].

2.2.2. Glutaraldehyde extension onto epoxy activated supports. Glutaraldehyde was used to extend the epoxy containing matrices (Eupergit® C, Eupergit® C250L and the cryogel monolith). All three matrices were washed with de-ionized water and then with 0.2 M Na_2CO_3 solution**.** Ethylenediamine (0.5 M in 0.2 M Na_2CO_3) was applied to supports in a tube and mixed on a shaking table for 4h. After washing with water until the pH was close to neutral the supports were washed with 0.1 M sodium-phosphate buffer, pH 7.2. A solution of glutaraldehyde (5% V/V) in 0.1 M sodium-phosphate buffer, pH 7.2 was applied to Eupergit® C, Eupergit® C250L and cryogels in a tube and the mixture was left under gentle shaking for 5 h (Method 2 step I and II, Fig. 1). The support materials were washed with 0.1 M sodium phosphate buffer pH 7.2 [15].

2.3. Scanning electron microscopy and thermogravimetic analysis of supports

Cryogel was sliced into thin slices using a scalpel, approx 2 mm. Two mL of citrate phosphate buffer pH 5.0 was added to one slice of cryogel or 20 mg of Eupergit® C and heated at 95°C for 1 h. The buffer was removed and the cryogel and the Eupergit® C beads were mounted on holders using double sided tape and stored in an exicator overnight. The samples were sputtered with gold and SEM images were collected using the scanning electron microscope LEO 1530 (Cambridge, UK) at 3 kV.

Approximately 1 mg of cryogel (chopped) or Eupergit® C was weighed in an alumina pan and loaded into the graphite furnace of a TGA Q500 instrument (TA Instruments, New Castle, USA). The temperature program was set to 10°C/min from room temperature up to 500°C and the measurement was done under air.

2.4 Expression and purification of the *Tn*Bgl1A variants

The genes encoding the mutant (N221S/P342L) and wild-type (*wt*) glucosidases were cloned in the expression plasmid pET-22b under control of the T7/lac promoter, including a C-terminal hexa-histidine tag [3]. The resulting clones were produced in 2.5 L *Escherichia coli* batch cultivations at 37°C, pH 7, using a defined medium [16] with 100 μg/mL ampicillin and a dissolved oxygen tension (DOT) above 40%. Expression was induced at OD_{620nm} = 3, by the addition of 1 mM isopropyl-beta-D-thiogalactopyranoside (IPTG), and continued for 3 h.

The target proteins were purified from the *E. coli* cultures by immobilised metal ion affinity chromatography (IMAC) using the following conditions. The cells were harvested, separated from the cultivation medium by centrifugation (10000 × *g*, 10 min, 4°C), and dissolved in binding buffer (20 mM imidazole, 20 mM Tris-HCl, 0.75 M NaCl, pH 7.5). The ice-chilled cell suspension was lysed by sonication for 5×3 min using a 14 mm titanium probe with a sound intensity of 60% and a cycle of 0.5 (UP400S, Dr. Hielscher), centrifuged (30 min, 39000 × *g*, 4°C), heat treated (70°C, 30 min) and heat-labile proteins of *E. coli* were removed by centrifugation (30 min, 39000 × *g*, 4°C). The supernatant was passed through a 0.45 μm Minisart high-flow filter (Sartorius, Göttingen, Germany) and purified on an ÄKTA prime system (Amersham Biosciences, Uppsala, Sweden) by IMAC using copper as a ligand as described elsewhere [17]. The fractions containing the purified protein were pooled and dialyzed against 0.1M sodium phosphate buffer, pH 7.2, overnight using a Spectra/Por dialysis membrane with a

3500 Da molecular weight cut-off (Spectrum laboratories, Rancho Dominguez, CA, USA). The dialysed protein fractions were stored at 4°C until use.

2.5. Protein and enzyme analyses

2.5.1. Surface exposed amino acids for immobilization: The three-dimensional structure model of *Tn*Bgl1A constructed by Khan et al, (2011) was used to display surface exposed amino acids. The model of *Tn*Bgl1A was constructed using the atomic coordinates of β-glucosidase BglA from *T. maritima (Tm*GH1 (PDB code 2WC4)) as template. Figures were prepared using PyMol (Delano Scientific LLC).

2.5.2. Total protein estimation. The concentration of protein was estimated by the bicinchoninic acid (BCA) method (Sigma, Steinheim, Germany) according to the manufacturer's instruction, using bovine serum albumin (0.2-1 mg/mL, Sigma-Aldrich) as standard.

The amount of protein bound to the support was calculated by using the following formula: (amount of β-glucosidase before immobilization) - (amount of β-glucosidase recovered in the filtrate).

2.5.3. Estimation of β-glucosidase activity. Enzyme activity was estimated every 5°C in the interval 80-95°C using *p*NPGlc (*para*-nitrophenyl-β-D-glucopyranoside) as substrate in 20 mM citrate phosphate buffer pH 5.6. The reaction measures the release of *p*-nitrophenol and was determined spectrophotometrically at 405 nm. A certain quantity of immobilized enzyme beads equivalent to 2.4 µg of *wt* and 1.2 µg of N221S/P342L (Eupergit® C or Eupergit® C250L) or a piece of cryogel (42.2 mg) was pre-incubated at the specific reaction temperature for 10 min and in parallell the *p*NPGlc solution (1.46 mM) was pre-heated at the same conditions. Pre-heated substrate solution was added to the pre-incubated immobilized enzyme beads (1 mL) or cryogel piece (10 mL). The sample was incubated at the selected temperature for 1 min and put on ice for 5 min. After 5 min, the samples were removed from the ice and left at room temperature for another 5 min, and then absorbance was measured spectrophotometrically at 405 nm (Perkin Elmer). The extinction coefficient of *p*NP (*para*-nitrophenol) under these experimental conditions was determined to. $\varepsilon_{80°C,\ 405nm} = 2.4639\times10^3$ mL/mM cm , $\varepsilon_{85°C,\ 405nm} = 2.5588\times10^3$ mL/mM cm , $\varepsilon_{90°C,\ 405nm} = 2.9527\times10^3$ mL/mM cm , $\varepsilon_{95°C,\ 405nm} = 3.0965\times10^3$ mL/ mM cm. The specific activity (U/mg) of the immobilized enzyme was determined and all measurements were

performed in triplicate. Enzyme activity was expressed as μmol of *p*-nitrophenol formed per min at the respective reaction temperature.

2.6. Immobilization of the *Tn*Bgl1A variants

In order to immobilize β-glucosidase variants on Eupergit® C, Eupergit® C250 L and the cryogel, two immobilization methods were used. In the first method enzymes were immobilized directly via oxirane groups and in the second method a derivatized matrix with functional aldehyde groups was used for immobilization (Fig. 1).

2.6.1. Immobilization via oxirane groups (Method 1, Fig. 1). Purified β-glucosidase (0.350 mg/mL) in 0.1 M sodium phosphate buffer pH 7.2 (30 mL) was added to 2.0 g Eupergit® C or 15 mL of the above enzyme solution was added to 8 pieces of cryogel where one piece of cryogel approximately weighs 42.2 mg. The support (Eupergit® C or cryogel) and enzyme solution were gently incubated at room temperature on a shaking table for 24 h. After incubation, the Eupergit® C beads were collected by vacuum filtration using a porous glass filter and rinsed thoroughly on the same filter with plenty of MilliQ water and then with 20 mM citrate phosphate buffer, pH 5.6. In a similar fashion, the cryogel pieces were collected, washed thoroughly with plenty of MilliQ water and then with 20 mM citrate phosphate buffer, pH 5.6 [18].

2.6.2. Immobilization onto the glutaraldehyde linker (Method 2 Fig. 1). The derivatized matrix with functional aldehyde groups was used for immobilization of a *Tn*Bgl1A variant. A 65 ml solution of β-glucosidase (0.225 mg/mL) in 0.1 M sodium phosphate buffer, pH 7.2, was mixed with 20 g of either Eupergit® C or Eupergit® C250L, or with 72 pieces of cryogel and incubated at room temperature for 24 h. Finally, 70 mL of freshly prepared $NaBH_4$ solution, 0.1 M in 0.2 M sodium carbonate buffer pH 9.2, was applied to the cryogels for 3 h to reduce the Schiff-base formed between the enzyme and the aldehyde containing matrix [15].

2.7. Addition of glucose and bovine serum albumin in the immobilization procedure

As glucose is the hydrolysis product in β-glucosidase catalyzed reaction and can be bound to the active site, the effect of its presence during the immobilization procedure was evaluated by the addition of 1% glucose (w/v) to the enzyme solution prior to the addition of support (Eupergit® C, Eupergit® C250L or cryogel).

The importance of blocking unreacted epoxy-groups was evaluated by incubating freshly washed Eupergit® C, Eupergit® C250L and cryogels after the enzyme binding step with 1% (w/v) bovine serum albumin (BSA) in 0.1 M phosphate buffer pH 7.2 at room temperature for 24 h.

2.8. Recycling of the immobilized enzymes

Recycling of the immobilized preparation was assessed in the temperature interval 80-95°C by repeated incubations in the presence of *p*NPGlc followed by spectrophotometric measurement of released *p*NP under the same conditions as in the above activity assay (see 2.5.3). After each activity measurement, the bound enzyme was washed three times with 20 mM citrate phosphate buffer, pH 5.6. The Eupergit® C and Eupergit® C250L beads were then centrifuged at 13000 rpm for 5 min, the supernatant was decanted and the recycled supports subjected to the activity assay in a second cycle and so on. The washing step and activity measurement were the same as above for the cryogel pieces except that the centrifugation step was replaced by manually pressing the gel against a tissue paper followed by washing with buffer. Finally the absorbance of wash fractions from both Eupergit® and cryogel were measured spectrophotometrically at 405 nm (Perkin Elmer).

2.9. Pressurized hot water extraction

Yellow onion was bought from a local supermarket in Sweden and the onion was chopped, both skin and bulb, and kept in a freezer (-20°C). Approximately 15 g of chopped onion was weighed into 33 mL stainless steel extraction vessels, with a glass fiber filter in the bottom, and loaded into a Dionex ASE®-200 instrument (Dionex, Sunnyvale, CA, USA). Water was used as extraction solvent as described by Turner et al. (Turner et al 2006). The extraction was done at a pressure of 50 bars, at a temperature of 120 °C, with an initial heating of 6 min and the extraction time was 15 min (5 min x 3 cycles). After each extraction cycle the cell was purged with N_2 (g) and the extract was collected in 100 mL clear glass vials. The extraction was repeated six times and the collected extracts were pooled and diluted to 500 mL with MilliQ water

2.10. On-line hydrolysis of yellow onion extract and quantification of quercetin

A home built continuous flow system was used for the on-line hydrolysis of the yellow onion extract. The system was composed of one HPLC-pump (HP1050, Agilent Technologies,

Walbronn, Germany), a GC-oven (HP 5890, Agilent Technologies, Walbronn, Germany), a fraction collector (Fraction Collector III, Waters, Milford, US) and 1/16´´ stainless steel tubings were used. Upstream the cell with enzyme, a 2 m coil of tubing was used, to ensure enough time to heat up the solvent before reaching the enzymes. A 0.2 g portion (with added glucose during the procedure and with BSA blocking) of Eupergit® C250L with immobilized N221S/P342L, (according to Method 2, Fig 1) was loaded into a home built peek holder and the peek holder was loaded into a stainless steel guard column from Agilent Technologies and placed in the GC-oven. Twenty mL of yellow onion extract was mixed with 20 mL of 40 mM citrate-phosphate buffer, pH 5.5 (Solvent A). The temperature of the GC-oven was set to 80°C and the flow of the pump was 1 mL/min. A gradient between 20 mM citrate-phosphate buffer, pH 5.5, (Solvent B) and the onion-buffer solution (Solvent A) was used; 0-5 min 100% B, 5-5.5 min 100% A, 5.5-15.5 min 100% A, 15.5-16 min 100% B and 16-21 min 100% B. Between 0-5 min and 16-21 min the fraction collection time was 2.5 min/sample and between 5-16 min the fraction collection time was 1 min per sample. To the collected fractions, morin (internal standard), methanol and formic acid were added, to a final concentration of 0.03 g/L of morin, 40 % of methanol and 0.5 % of formic acid. Methanol and formic acid were added to avoid precipitation of quercetin (Q) and to lower the pH to ensure that hydroxyl groups of Q are protonated. 500 µL of the collected fractions were diluted with 1000 µL of mobile phase (see below) and analyzed with HPLC-UV.
The chromatographic system used was an UltiMate 3000 from Dionex (Germering, Germany) and the column used was an Agilent Zorbax SB-C18 column (100 x 2.1 mm, 3.5 µm), with an Agilent Zorbax SB-C8 pre-column (2.1 x 12.5 mm, 5 µm). Isocratic separation was used, with a methanol–water (40:60) and formic acid (0.13 M, 0.5 vol%) mobile phase at a flow rate of 0.15 mL/min. The injection volume was 10 mL and detection was accomplished at 350 nm. Quantification of Q, quercetin glucosides (Q-3, Q-4´ and Q-3,4´) and morin (internal standard) was done using a five-point calibration curve (standards of the respective compound at concentrations between 0.5 and 25 mg mL^{-1}). Each vial taken to analysis had a total volume of 500 mL.

3. Results

3.1. Thermostability of support materials

Two different experiments were set up to confirm the heat stability of the cryogel and Eupergit® C support materials at temperatures relevant for the process conditions. In the first experiment, the materials were heated in buffer at 95°C for 1 h and SEM images were collected of non heat-treated material and heat-treated material. Fig. 2 shows that there is no obvious conformation change of the support materials after heat treatment.

The second experiment the temperature range was extended, to study at which temperature the materials start to decompose, using thermogravimetric analysis (TGA) (Fig. 3). The result of the TGA shows that Eupergit® C and cryogel starts to decompose around 250°C (Fig. 3). These results confirm that the support materials are stable for use in systems operating up to the current upper limit for protein stability (150°C) (Tanaka et al, 2006).

3.2. Surface-exposed amino acid side-chains on *Tn*Bgl1A variants

Immobilization via epoxy- and aldehyde groups allow direct chemical coupling to amino-groups ($-NH_2$), and via the epoxy groups also hydroxyl-groups (-OH) on the amino acid side-chains [12]. Lys-residues are the most commonly utilized anchoring points as they are typically present on the exterior of proteins [12, 19]. Inspection of the 3D-structure model of *Tn*Bgl1A showed that 26 of the total 27 Lys residues in the sequence were located on the surface (Fig 4 a-c), evenly spread, and available for covalent interactions with aldehyde- and epoxy groups on the support materials. Five of the Lys-residues surrounded the active site (Fig. 4a), which can lead to covalent bonds that obstruct activity of the immobilized enzyme. In addition, serine (Ser) and threonine (Thr) residues were analyzed. Ten of the 18 Ser and 10 of totally 15 Thr were surface located, of which 3 Ser-residues flank the active site (Fig. 4a).

3.3. Immobilization of *Tn*Bgl1A wt and N221S/P342L

Three epoxy-activated acrylic polymer matrices with different texture properties were used and included particulate beads of Eupergit® C and Eupergit® C250L (differing in pore size and oxirane number) and a monolithic cryogel. Firstly, the enzyme variant N221S/P342L was used, and selected immobilization conditions repeated for the *wt* enzyme. The N221S/P342L was immobilized both directly onto the epoxy group (Method 1, Fig. 1) and via an aldehyde spacer

arm (Method 2, Fig. 1). It was found that the aldehyde had little effect on specific activity using Eupergit® C (1.3× specific activity decrease was observed using the spacer arm, Table 1). However the spacer arm had a drastic effect on the cryogel support, where the specific activity was approximately 50 times lower compared to Eupergit® C when the enzyme was directly linked to the epoxy group, but increased one order of magnitude upon introduction of the spacer arm (Table 1). In this set of experiments, no effort was made to protect the active site or to block oxirane groups remaining after incubation with the enzyme.

3.3.1. Additives

According to Tu *et al.* [20], application of 1% glucose and 1% BSA in the immobilization procedure of a β-glucosidase improves the immobilization efficiency. To investigate this for the N221S/P342L variant, glucose (1% w/v) was added to the enzyme solution prior to immobilisation by direct linkage to the epoxy groups of Eupergit® C (Method 1, Fig. 1). Glucose additions led to a slight improvement in specific activity (see Table 1) while addition of bovine serum albumin (BSA) after immobilization of the enzyme to block unreacted oxirane or aldehyde groups showed stronger effect on the Eupergit® support materials. In these experiments Eupergit® C 250L was used instead of Eupergit® C, as the latter was no longer available at the supplier. Moreover, the glutaraldehyde spacer arm was used, as complementary experiments had shown this to be better for long term use (see below). In accordance with the results by Tu and coworkers (Tu et al., 2006)), blocking by BSA resulted in a 5 times increase of the specific activity of the thermostable *Tn*Bgl1A variants (Table 1). A contribution to this effect may however also be the larger pore size of Eupergit® C250L which previously has been suggested to improve enzyme diffusion into the macroporous matrix [21]. However, when BSA was added to a stored preparation immobilized of Eupergit® C with the lower pore size (see below), also in this case specific activity was increased, indicating that BSA is indeed needed to reduce substrate/support interactions.

Neither glucose nor BSA had any significant effect on the specific activity using the cryogel support (Table 1). In all cases, the specific activity of the immobilized enzyme is significantly lower than that of free enzyme in solution. Immobilization leads to an approximately 5 times decrease using Eupergit® C250L, while on the cryogel the specific activity decrease is almost two orders of magnitude, using the current evaluation methodology.

3.4. Recycling and storage of immobilized *Tn*Bgl1A variants

The reusability of the enzyme is a most attractive characteristic of immobilized enzyme applications. The β-glucosidase variants immobilized on Eupergit® C without spacer arm (Method 1, Fig. 1) was recycled 3 times (Fig. 5a-b) and N221S/P342L immobilized on Eupergit® C250L with glutaraldehyde spacer arm (Method 2, Fig. 1) was recycled 10 times at the temperatures 80, 85, 90 and 95°C (Fig 5c). The preparations were washed and centrifuged between the trials and no activity could be detected in the wash fractions (data not shown).

In order to investigate the industrial practicability of an immobilized enzyme, the loss of enzyme activity, known as storage stability, is an important parameter to be taken into account. Hence the stability of the immobilized enzymes on cryogels, Eupergit® C and Eupergit® C250L were studied after storage at 4°C. In a combined experiment on storage and reusability, the enzyme immobilized on Eupergit® C250L (Method 2, Fig 1) was stored at 4°C for 6 months and a 10-30% decrease in specific activity was observed both in *wt* and N221S/P342L enzyme (Fig. 6). The decrease in specific activity was further measured by recycling the enzymes for 10 times at the selected temperatures in the 80-95°C interval. As shown in Fig. 6, the immobilized enzyme retained more than 75% (*wt*) and 50% (N221S/P342L) of its original activity after 10 reuses, respectively. The wash fractions were also analysed, but again no activity loss to the supernatant could be observed, using the *p*NPGlc substrate.

After 3 months there was no loss in activity of enzymes immobilized by Method 2 (see Fig. 1) on cryogels but on the other hand the Eupergit® C preparation immobilized by Method 1 had lost approximately 50% of its initial activity (Fig 7). The specific activity of *Tn*Bgl1A immobilized on Eupergit® C250L (immobilized by Method 2, Fig.1) was measured after 6 months storage, and the enzyme still retained more than 75% of its original activity (Fig. 6). This clearly indicated that the enzyme preparations are very stable when immobilized by the glutaraldehyde method.

In another experiment the activity of an immobilized preparation of β-glucosidase variants on Eupergit® C (Method 1, Fig. 1) which had lost approximately 50 % activity after 3 months storage at 4 °C (Fig. 8) was incubated with 1% BSA for 24 h after which activity again was measured. By this BSA treatment, the preparation regained approximately the original activity of the freshly immobilized enzymes (Fig. 7).

3.5. On-line hydrolysis of quercetin glucosides, using *Tn*Bgl1A(N221/P342L) immobilized on Eupergit® C 250L

An enzyme preparation of N221/P342L**,** immobilized on Eupergit® C250L (with glutaraldehyde linker, immobilized in presence of glucose and pretreated with BSA), was used in an on-line hydrolysis experiment. Yellow onion extract mixed with citrate-phosphate buffer was pumped through a cell with the enzyme at 80°C and fractions were collected at different time intervals. In Fig. 9 the concentration of Q-3,4′, Q-4′, Q-3 and quercetin in collected fractions of the on-line hydrolysis can be seen (values at 0 min corresponds to the concentration of the quercetin compounds in the onion extract pumped through the on-line system). The results of the on-line hydrolysis of the onion extract is that all Q-4′ is hydrolyzed to quercetin and that Q-3,4′ is firstly hydrolyzed to Q-3 and secondly to quercetin. Hydrolysis of Q-3 is incomplete. In summary, the immobilized enzyme appears successful for the deglycosylation step of quercetin glucosides in hot yellow onion water extract, but the method have to be further optimized if the goal is to achieve complete hydrolysis of the quercetin glucosides.

4. Discussion

Previous results in our laboratory have shown that the thermostable β-glucosidase A from *T. neapolitana*, *Tn*Bgl1A, as well as a mutated variant thereof, can be used as a catalyst for the conversion of quercetin glucosides to quercetin and glucose in yellow onion extracts obtained from pressurized hot water extractions [3, 8]. Use of the free thermostable enzyme in solution connected to the high temperature extraction methodology allowed catalysis without extensive cooling, but use of free enzyme can be expensive as it necessitates batch wise additions. Immobilization was thus an interesting alternative allowing recovery and reuse of the biocatalyst. Due to the high temperature conditions in the extraction process, covalent immobilization methods were chosen, and two different approaches explored: (a) direct enzyme binding to the matrices via the oxirane groups, and (b) enzyme binding to the support via an ethylene diamine/glutaraldehyde spacer.

Three epoxy-activated acrylic polymer matrices with different texture properties were used in the study and included particulate beads of Eupergit® C and Eupergit® C250L (differing in pore size and oxirane number) and a monolithic cryogel. The suitability of the support material for application in the high temperature method was confirmed, as the support materials were stable

and can be used in future on-line flow system where temperatures up to 150°C (the current upper limit for protein stability [22]) are applicable. Coupling to epoxy- and glutaraldehyde groups on the supports for immobilization to amino and hydroxyl groups on the enzyme surface, exposed to the aqueous medium [12] was also successful, and judged to be due to the Lys, Thr and Ser residues evenly distributed on the enzyme surface as shown by inspection of the 3D-structure model of *Tn*Bgl1A [23]. In principle, a number of different groups *e.g*. amino, hydroxyl, thiol and phenolic can react with epoxy- and aldehyde groups [24] but the amino-group of Lys and hydroxyl of Thr/Ser are most favored [12] for immobilization.

All three types of support materials (Eupergit® C; Eupergit® C250L and cryogels) used in this study are examples of epoxy activated acrylic supports, but have some differences. Eupergit® C and Eupergit® C250L are micro-porous, epoxy-activated acrylic beads commercially available with a diameter of 100-250 μm and a pore size of 10-100 nm [10], while the cryogels are produced through gelation at subzero temperature with properties such as elastic and sponge-like morphology and with a pore size of 1-100 μm, giving a macroporous gel reported to be highly flow permeable [25, 26]. Ligands can be coupled directly or via spacer arms to a reactive group, *e.g.* epoxy-groups, to both Eupergit® C and cryogels [14]. In all cases, the specific activity (U/mg) of the immobilized *Tn*Bgl1A-variants did not decrease in the temperature interval 80-95°C, which is an improvement as the soluble counterpart was only stable up to 90°C [8]. Some differences in enzyme activity were however observed between the supports (even though all had the same reactive groups) which are likely due to their different nature [27] and will be discussed further.

Higher specific activity was observed on Eupergit® C250L than on Eupergit® C, despite having the same composition. The reasons for the lower specific activity of Eupergit® C could be, either its smaller pore diameter which may cause diffusion limitation [13], or the higher content of the oxirane groups of Eupergit® C that could result in the formation of multiple covalent bonds between the support and enzyme, which might disturb the enzyme structure and cause some loss in activity [21].

The low activity on the cryogel compared to the Eupergit® supports could be explained if it is assumed that a substantially large portion of enzymes immobilized on the cryogel were entrapped in the intrinsic section of the carrier, limiting access of the substrate to the active site due to diffusion limitation. This explanation is supported by the fact that the specific activity was

higher when the reactive groups in the support were secluded from its surface through short spacer arms. Thus, the introduction of spacer arms helps in avoiding steric hindrance effects as it guarantees an additional distance between the support and the enzyme and places the enzyme in a better position for catalytic action after immobilization [28, 29]. The ideal situation for enzymes immobilized on cryogel is hence to pass the substrate through the monolith (e.g. combined with an on-line flow system), and this way the substrate could access to the active site of the enzyme entrapped in the intrinsic section of the carrier and offer a higher activity [26].

One problem with enzymes directly immobilized onto the epoxy groups of both Eupergit® C and Eupergit® C250L (Method 1, Fig. 1) was a more rapid loss of enzymatic activity, as previously described by [29], so the technique that led to higher storage stability was considered superior and further experiments were carried out using enzyme immobilized via the glutaradehyde linker (Method 2, Fig. 1). The enzyme immobilized by this method retained more than 75% specific activity even after six months (on Eupergit® C250L) while no loss in activity was observed on cryogels.

To maintain high and stable activity of immobilized enzymes (which are most important factors to be considered) two methods, using additives, have been used. In the first method glucose was added to the enzyme solution before immobilization, which according to [20] changes the microenvironment and protects the active site of the enzyme during immobilization, leading to increased specific activity. Glucose has however previously been shown to activate *Tn*Bgl1A [23, 30], and the slight improvement seen in specific activity (1.3× higher) in our experiments, is in the same range as the previously reported activating effect (1.1× higher specific activity). Hence, even though specific activity increased we cannot conclude that glucose has any significant role in protecting the active site during immobilization. In another method, the microenvironment of the enzyme molecules has been changed by addition of BSA, which can easily react with active functionalities, such as epoxy rings or aldehyde groups of the support without the use of additional activating chemistry [31]. Addition of BSA can significantly increase the specific activity and stability, and this effect was observed when stored preparations of immobilized enzyme on Eupergit® C were incubated with BSA. This has been explained as a hydrophilic environment created into the proximity of the enzyme enhancing the hydrolytic activity and stability, and has previously been observed in consecutive modifications of β-glucosidase [20] and pencillin V acylase [32] immobilized on Eupergit® C. It should however be

noted that with our thermostable enzyme, the most promising activities were found when immobilized enzyme was stored long term in the absence of BSA, and conditioned with the BSA immediately prior to use. In contrast to BSA, blocking of residual groups with 2-mercaptoethanol had no effect on enzyme activity, strongly suggesting that it is the introduction of the hydrophilic environment that is favorable for enzyme stabilization [32]. Upon immobilization, the β-glucosidase could be used for many cycles of hydrolysis without any significant loss in activity. Thus, high stability (95 °C) and good reusability makes the immobilized β-glucosidase very attractive for on-line hydrolysis of for example quercetin-glucosides from yellow onion extracts.

Conclusion

In conclusion the present work demonstrates immobilization of active thermostable β-glucosidase variants by covalent binding to Eupergit® C, Eupergit® C250L and cryogel. This is the first report of covalent immobilization of β-glucosidase variants from *T. neapolitana* on acrylic supports, and shows that immobilization of the enzymes allow recycling, improved the thermostability and resulted in a preparation with good operational stability allowing storage (up to 6 months thus far tested). This makes the enzyme a potential candidate for high temperature processing, successfully used for the hydrolysis of quercetin-glucosides (with potential for other by-product modifications) in hot water extractions, and can be a great candidate for various industrial applications.

Acknowledgements

Oskar Werner is thanked for help with SEM analysis and the Department of Medical Cell Biology, Uppsala University for use of SEM instrument, Prof. Patric Jannasch and Carlos Rodriguez Arza (Centre for Analysis and Synthesis, Dept Chemistry, Lund University) are thanked for assistance on TGA analysis and Dr. Fatima Plieva (Biotechnol, Dept Chemistry, Lund University) for help with the cryogel work.

The authors wish to thank Formas (2006-1346 (Sustain-X-Enz), and 2009-1527 (SuReTech)), STINT (YR2009-7015), Oscar & Lili Lamm´s Foundation and "Kungl Fysiografiska Sällskapet i Lund" for funding, and the Higher Education Commission (HEC) of Pakistan for support to Samiullah Khan.

References

[1] S. Panke, M. Held, M. Wubbolts, Curr Opin Biotechnol, 15 (2004) 272-279.

[2] S.B. Rubin-Pitel, H.M. Zhao, Comb Chem High Throughput Screen, 9 (2006) 247-257.

[3] C. Turner, P. Turner, G. Jacobson, K. Almgren, M. Waldeback, P. Sjoberg, E.N. Karlsson, K.E. Markides, Green Chem, 8 (2006) 949-959.

[4] S. Fernando, S. Adhikari, C. Chandrapal, N. Murali, Energy fuels, 20 (2006) 1727-1737.

[5] G. Paes, M.J. O'Donohue, J Biotechnol, 125 (2006) 338-350.

[6] B. Kamm, M. Kamm, App Microbiol Biotechnol, 64 (2004) 137-145.

[7] S. De, R.N. Chakraborty, S. Ghosh, A. Sengupta, S. Das, J Exp Clin Cancer Re, 23 (2004) 251-258.

[8] S. Lindahl, A. Ekman, S. Khan, C. Wennerberg, P. Börjesson, P.J.R. Sjöberg, E.N. Karlsson, C. Turner, Green Chem, 12 (2010) 159-168.

[9] J. Torres-Bacete, M. Arroyo, R. Torres-Guzmán, I.d.l. Mata, M.P. Castillón, C. Acebal, J Chem Tech Biotechnol, 76 (2001) 525-528.

[10] T. Boller, C. Meier, S. Menzler, Org Process Res Dev, 6 (2002) 509-519.

[11] C. Mateo, J.M. Palomo, G. Fernandez-Lorente, J.M. Guisan, R. Fernandez-Lafuente, Enzyme Microb Technol, 40 (2007) 1451-1463.

[12] F. Rusmini, Z.Y. Zhong, J. Feijen, Biomacromolecules, 8 (2007) 1775-1789.

[13] P.-Y. Wang, S.-W. Tsai, T.-L. Chen, J Chem Tech Biotechnol, 83 (2008) 1518-1525.

[14] R. Mallik, D.S. Hage, J Sep Sci, 29 (2006) 1686-1704.

[15] A. Kumar, F.M. Plieva, I.Y. Galaev, B. Mattiasson, J Immunol Methods, 283 (2003) 185-194.

[16] S.O. Ramchuran, O. Holst, E.N. Karlsson, J. Biosci. Bioeng., 99 (2005) 477-484.

[17] P. Turner, O. Holst, E.N. Karlsson, Protein Expression and Purification, 39 (2005) 54-60.
[18] M.T. Martin, F.J. Plou, M. Alcalde, A. Ballesteros, J Mol Catal B-Enzym, 21 (2003) 299-308.

[19] S.V. Rao, K.W. Anderson, L.G. Bachas, Biotechnol Bioeng, 65 (1999) 389-396.

[20] M. Tu, X. Zhang, A. Kurabi, N. Gilkes, W. Mabee, J. Saddler, Biotechnol Lett, 28 (2006) 151-156.

[21] Y.J. Cho, O.J. Park, H.J. Shin, Enzyme Microb Technol, 39 (2006) 108-113.

[22] T. Tanaka, M. Sawano, K. Ogasahara, Y. Sakaguchi, B. Bagautdinov, E. Katoh, C. Kuroishi, A. Shinkai, S. Yokoyama, K. Yutani, FEBS Lett, 580 (2006) 4224-4230.

[23] S. Khan, T. Pozzo, M. Megyeri, S. Lindahl, A. Sundin, C. Turner, E.N. Karlsson, BMC Biochem, 12 (2011) 11.

[24] I. Migneault, C. Dartiguenave, M.J. Bertrand, K.C. Waldron, Biotechniques, 37 (2004) 790-802.

[25] G. Massolini, E. Calleri, J Sep Sci, 28 (2005) 7-21.

[26] D. Josic, A. Buchacher, A. Jungbauer, J Chromatogr B-Analyt Technol Biomed Life Sci, 752 (2001) 191-205.

[27] C. Mateo, R. Torres, G. Fernandez-Lorente, C. Ortiz, M. Fuentes, A. Hidalgo, F. Lopez-Gallego, O. Abian, J.M. Palomo, L. Betancor, B.C.C. Pessela, J.M. Guisan, R. Fernandez-Lafuente, Biomacromolecules, 4 (2003) 772-777.

[28] Q. Luo, H. Zou, Q. Zhang, X. Xiao, J. Ni, Biotechnol Bioeng, 80 (2002) 481-489.

[29] Z. Knezevic, N. Milosavic, D. Bezbradica, Z. Jakovljevic, R. Prodanovic, Biochem Eng. J, 30 (2006) 269-278.

[30] D.A. Yernool, J.K. McCarthy, D.E. Eveleigh, J.D. Bok, J Bacteriol, 182 (2000) 5172-5179.

[31] L. Cao, Curr Opin Biotechnol, 9 (2005) 217-226.

[32] J. Torres-Bacete, M. Arroyo, R. Torres-Guzman, I. de la Mata, M.P. Castillon, C. Acebal, J Chem Technol Biotechnol, 76 (2001) 525-528.

Figure Legends

Fig. 1. Schematic overview of the two main methods used for the covalent coupling of *Tn*Bgl1A (*wt* and N221S/P342L) to the support. Method 1 represent coupling via oxirane groups and Method 2 shows coupling onto the glutaraldehyde groups.

Fig. 2. Scanning electron microscope images of (a) non-heat treated cryogel, (b) heat treated cryogel (1 h at 95°C), (c) non-heat treated Eupergit® C and (d) heat treated Eupergit® C (1 h at 95°C).

Fig. 3. Thermogravimetric analysis of cryogel and Eupergit® C. The change of weight of the sample with temperature is given. The profile shows: (a) weight loss due to loss of water and (b) weight loss due to decomposing of the polymer.

Fig. 4. Surface view of the 3D structure model of *Tn*Bgl1A, constructed using β-glucosidase BglA from *Thermotoga maritima (Tm*GH1 (PDB code 2WC4)) as template. From left to right: (a) front view (with the active site pocket), (b) side view, and (c) view of the side opposing the active site of *Tn*Bgl1A. Lysine residues are labelled blue, and serine or threonine residues are shown in orange.

Fig. 5. Reusability of immobilized β-glucosidase variants (immobilized directly to the epoxy group (a-b) or with spacer arm (c)) on Eupergit® C and Eupergit® C250L was monitored in the temperature range 80-95°C without a noticeable loss of specific activity (U/mg). (a) Immobilized *Tn*Bgl1A (*wt*) on Eupergit® C recycled three times (b) immobilized (N221S/P342L) on Eupergit® C recycled three times (c) immobilized (N221S/P342L) on Eupergit® C250L recycled ten times.

Fig. 6. Reusability and storage stability of the β-glucosidase variants on Eupergit® C250L, immobilized via the glutaraldehyde spacer arm, after storage at 4°C for six months, followed by recycling ten times at 80-95°C. (a) specific activity (U/mg) of *Tn*Bgl1A (*wt*) and (b) specific activity (U/mg) of (N221S/P342L)

Fig. 7. Effect of the addition of 1% bovine serum albumin (BSA) on the specific activity (U/mg) of β-glucosidase variants immobilized on Eupergit® C via the epoxy group. (a) Specific activity (U/mg) of *Tn*Bgl1A (*wt*) at 80-95°C after three months storage (4°C) in the absence of BSA and with 1% BSA. (b) Specific activity (U/mg) of (N221S/P342L) at 80-95°C after three months storage (4°C) in the absence of BSA and with 1% BSA.

Fig. 8. Comparison of storage stability of β-glucosidase variants immobilized directly to the epoxy groups of Eupergit® C and cryogel. (a) Specific activity (U/mg) of freshly immobilized *Tn*Bgl1A (*wt*) and after three months storage at 4°C. (b) Specific activity (U/mg) of (N221S/P342L) before three months and after three months.

Fig. 9. Concentration of quercetin-3,4′-diglucoside (Q-3,4′), quercetin-3-glucoside (Q-3), quercetin-4′-glucoside (Q-4′) and quercetin in fractions collected from on-line hydrolysis of onion extract. Values at 0 minutes correspond to the concentration of the compounds in the extract before being pumped through the cell with the immobilized enzyme variant N221S/P342L.

Table 1
Specific activity of immobilized *Tn*Bgl1A variants at different temperatures using different supports. The preparation with the highest specific activity (for the respective variant) is shown in bold. V_{max} for the free enzyme at 80°C was 540 U/mg (*wt*), and 1890 U/mg (N221S/P342L).

*Tn*Bgl1A variants	Support	Additive	Reactive group	Specific activity 80°C (U mg^{-1})	Specific activity 85°C (U mg^{-1})	Specific activity 90°C (U mg^{-1})	Specific activity 95°C (U mg^{-1})
N221S/ P342L	EupC	-	Epoxy	57.8±5.5	58.4±2.6	49.5±2.8	48.2±2.8
	EupC	1%glu	Epoxy	77.0±14.7	80.3±3.0	65.8±6.0	67.3±4.9
	EupC	-	Aldehyde	41.0±2.1	47.9±6.4	36.1±0.76	35.5±1.9
	EupC250L	**1%glu, 1%BSA**	**Aldehyde**	**243.7±11.6**	**254.0±12.0**	**242.1±24.6**	**238.9±3.7**
	Cryogel	-	Epoxy	1.5	1.3	1.1	0.6
	Cryogel	-	Aldehyde	12.8±0.6	16.9±1.8	16.6±2.1	19.0±1.8
	Cryogel	1%glu, 1%BSA	Aldehyde	10.2±0.0035	10.9±0.001	11.7±0.0035	10.0±0.009
Wild type	EupC	-	Epoxy	30.7±4.9	33.2±2.3	32.5±5.6	25.7±6.7
	EupC250L	**1%glu, 1%BSA**	**Aldehyde**	**110.4±14.1**	**131.3±26.6**	**93.8±10.6**	**93.0±2.5**
	Cryogel	-	Aldehyde	2.7±0.4	4.0±0.4	3.9±0.0	3.8±0.6

Figure 1

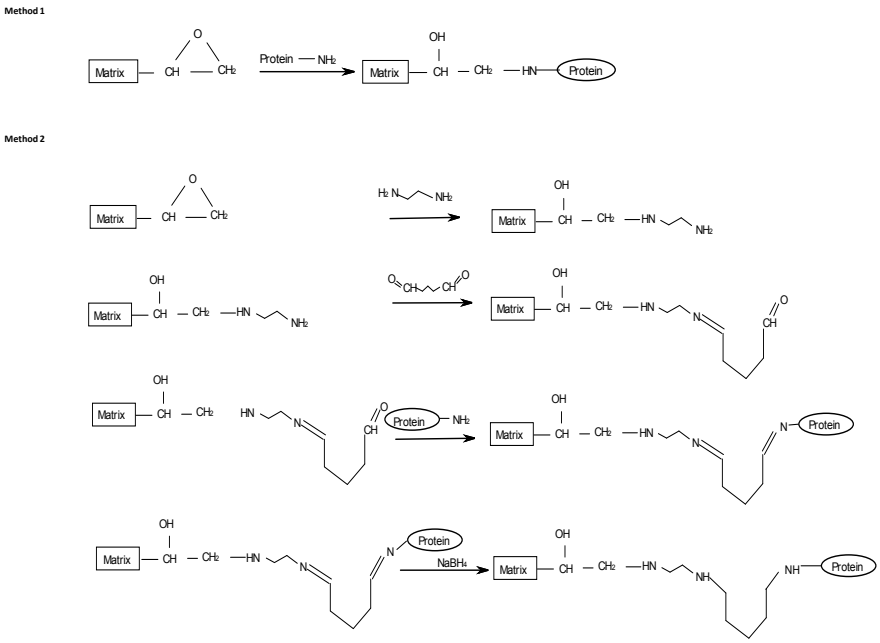
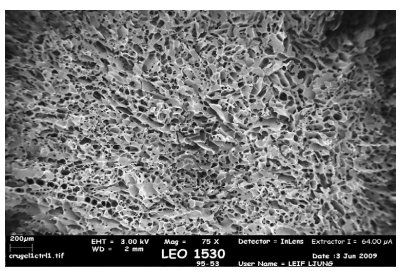
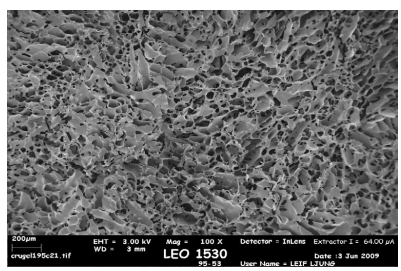


Figure 2

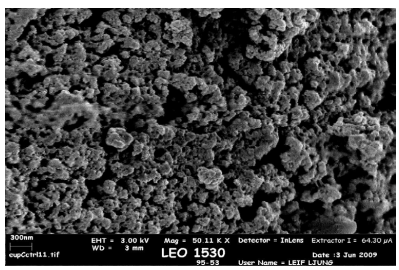
(a)



(b)



(c)



(d)

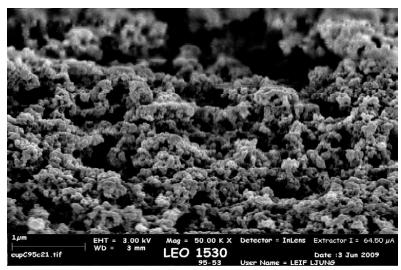


Figure 3

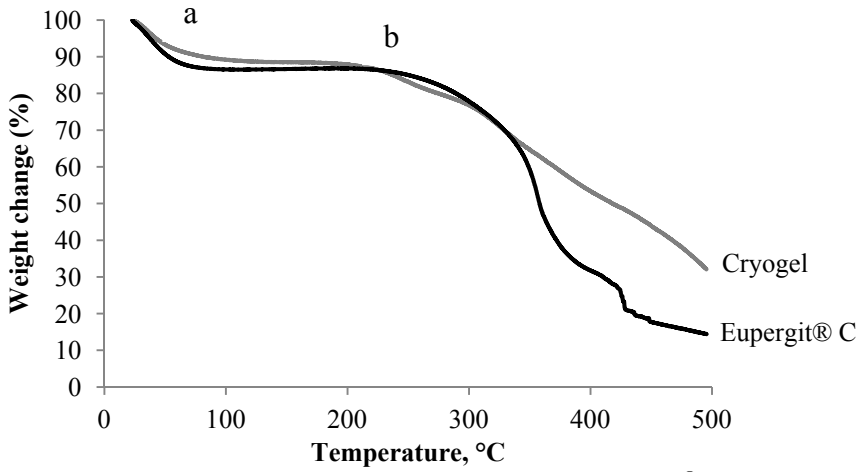


Figure 4

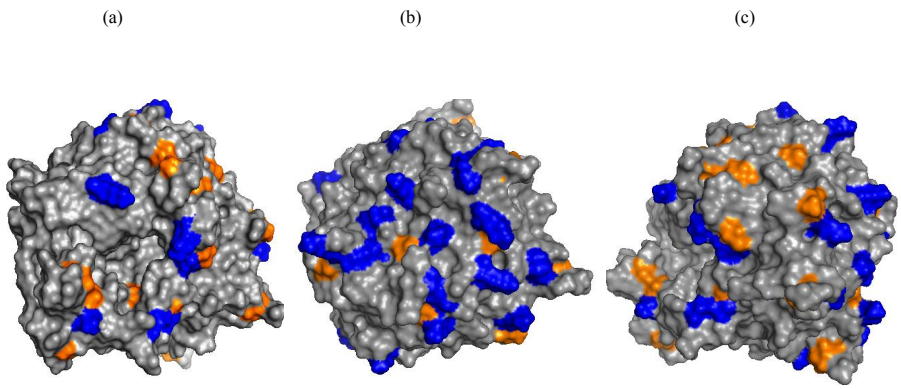


Figure 5

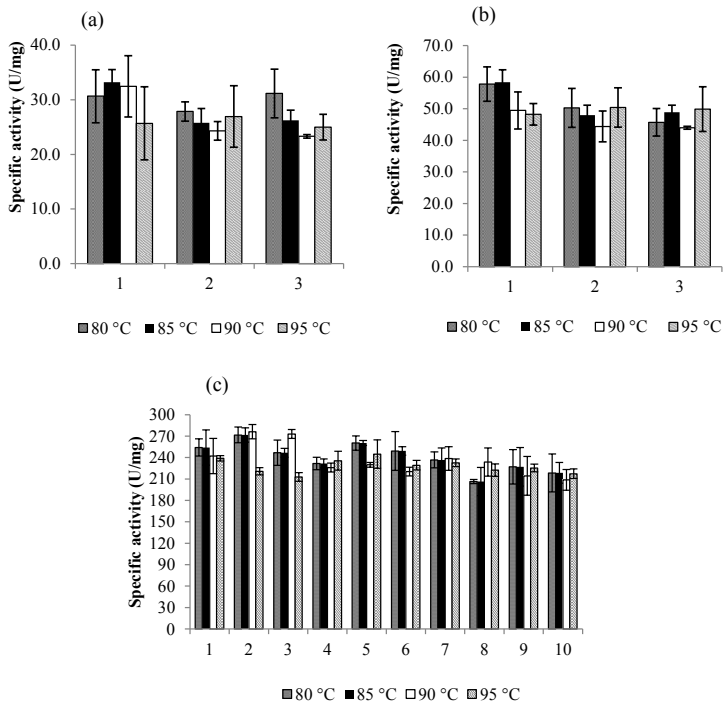


Figure 6

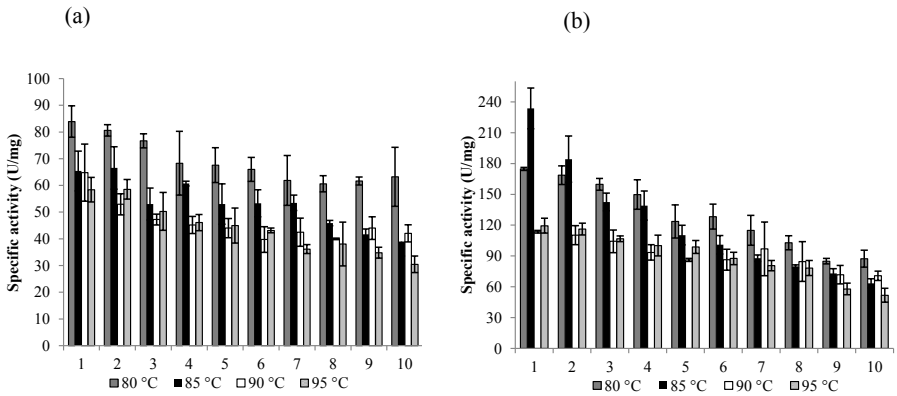


Figure 7

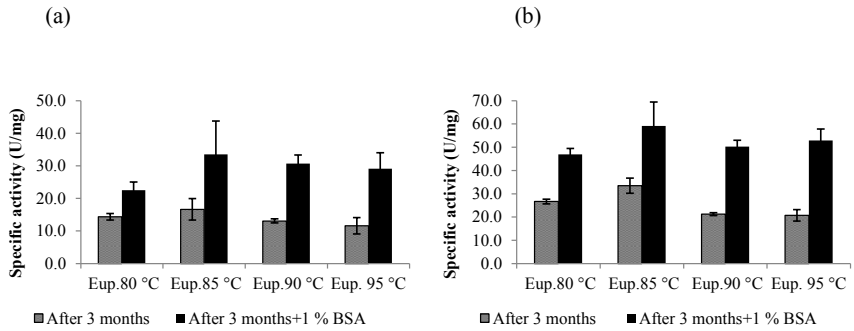


Figure 8

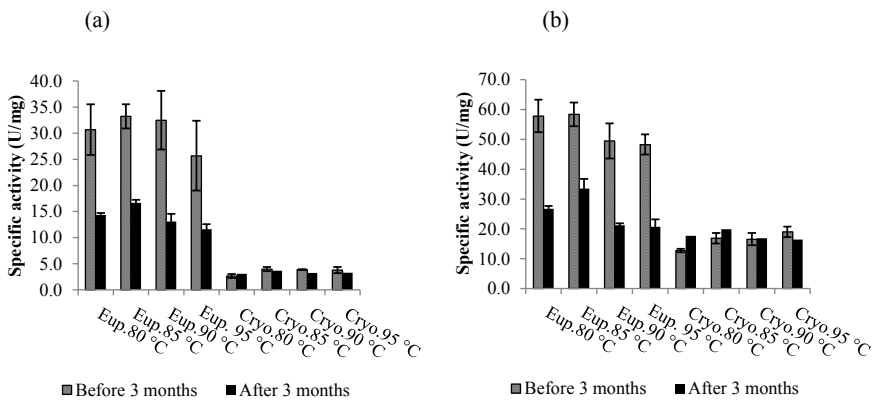
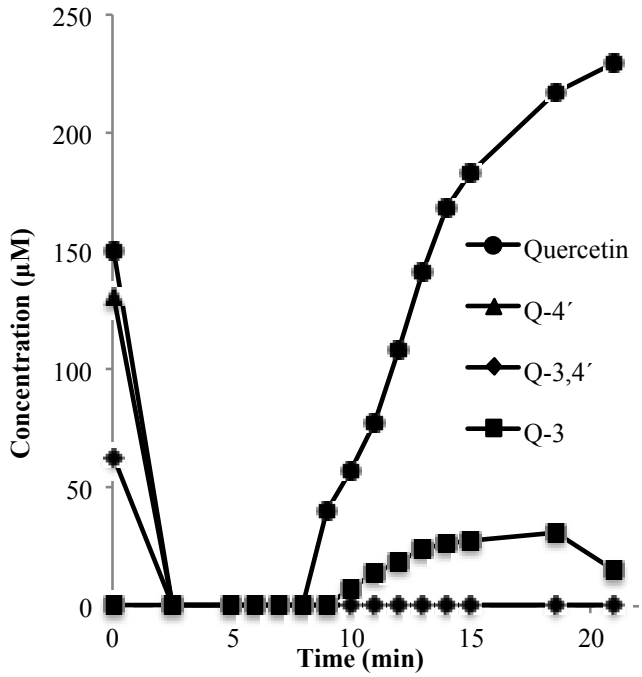


Figure 9


250
200
150
100
50
0
Concentration (µM)
0
5
10
15
20
Time (min)
Quercetin
Q-4´
Q-3,4´
Q-3

Paper IV

An on-line method for pressurised hot water extraction and enzymatic hydrolysis of quercetin glucosides from onions.

Sofia Lindahl[a], Jiayin Liu[a], Samiullah Khan[b], Eva Nordberg-Karlsson[b] and Charlotta Turner[a*]

[a] Lund University, Department of Chemistry, Centre for analysis and synthesis, P.O. Box 124, SE-221 00 Lund, Sweden and [b] Lund University, Biotechnology, Department of Chemistry, P.O. Box 124, SE-221 00 Lund, Sweden.

*Corresponding author: Charlotta.Turner@organic.lu.se, Tel. +46 46 222 81 25

Keywords: On-line, hyphenation, hot water extraction, hydrolysis, flavonoids, quercetin, enzyme, beta-glucosidase, *Thermotoga neapolitana*

Abstract

A home built continuous flow system was used for on-line extraction and hydrolysis of quercetin glucosides from onion using hot water as extraction solvent and a thermostable beta-glucosidase as catalyst for the hydrolysis. *Tn*Bgl1A_N221S/P342L was immobilised on two different support materials, Eupergit® C 250L and cryogel. First, the hydrolysis conditions were optimised regarding temperature (75-95°C), pH (3-6) and volume concentration of ethanol (5-15%) using a standard solution of quercetin-3,4´-diglucoside, with a central composite design. The optimal conditions were 84°C, 5% of ethanol and pH 5.4 (cryogel) respectively pH 5.6 (Eupergit® C 250L). The extraction of quercetin, quercetin-4´-glucoside and quercetin-3,4´-diglucoside from yellow onion was studied using the optimal hydrolysis condition with a flow rate of 1 mL/min or 3 mL/min. The total quercetin extraction yield was 1.7 times higher using the flow rate of 3 mL/min, at a total extraction time of 90 min, compared to a flow rate of 1 mL/min and total extraction time of 240 min. The combined extraction and hydrolysis of yellow-, red- and shallot onion using 5% ethanol and a pH of 5.5 showed higher extraction yield compared to a method using aqueous methanol with 1.2 M HCL. The enzyme stability was tested by a prolonged incubation of the enzyme at 84°C, during which the specific activity using *para*-nitrophenyl-β-D-glucopyranoside (pNPGlc) as substrate was maintained for up to 48 h.

1. Introduction

The forest and agricultural industry produce tons of by-products daily and from these by-products high value compounds can be extracted and isolated before the waste is disposed. Examples of extractable high-value compounds are flavonoids, which are present in fruits and vegetables, such as onions, apples, kale and cabbage [1].

The conventional way to extract flavonoids from plant material is to use solid-liquid extraction using organic solvents, for example methanol, ethanol or ethyl acetate [2]. In recent years, the uses of alternative more environmentally friendly extraction techniques have increased, such as pressurized fluid extraction (PFE), microwave assisted extraction (MAE) and supercritical fluid extraction (SFE) [3]. With these techniques, organic solvents are often still used, but the solvent to sample ratio is smaller and the extraction time is usually shorter, compared to conventional solid-liquid extraction. In PFE, temperatures above the boiling point of the extraction solvent can be used by applying a pressure, which will keep the solvent as a liquid. Advantages of using a hot extraction solvent is that the desorption of flavonoids from the plant material is improved due to the extra energy added, the viscosity and surface tension of the extraction solvent decrease, whereas the diffusivity increases, which improves the wetting of the matrix and the transport of flavonoids out from the matrix.

In PFE, the organic solvent can in some cases be replaced by water, i.e. pressurized hot water extraction (PHWE). PHWE has for example been used to extract quercetin from onion [4-5], quercetin-3-galactoside, kaempferol and isorhamnetin from seabuckthorn leaves [6] and dihydrokaempferol and naringin from aspen knotwood [7]. In this paper, PHWE is combined online with enzyme-catalysed hydrolysis of quercetin glucosides from onion. One disadvantage of using water extraction for quercetin is that the solubility of quercetin in water at ambient conditions is low, 0.002 mg/mL [8]. However, by increasing the temperature, the dielectric constant of water decreases [9] and the solubility of quercetin increases to around 0.7-1.5 mg/mL at 140°C [8]. A disadvantage of using PHWE is that thermolabile compounds may degrade, for example anthocyanins [10], so the choice of temperature has to be selected after the compound of interest.

Quercetin is one of the most expressed flavonoids in plants and it is an antioxidant. Studies, *in vitro*, have shown that quercetin may have positive effects against cancer [11] and neurodegenerative [12] diseases. Quercetin in onion is mainly present as different glucoside forms, of which quercetin-4´-glucoside (Q-4´) and quercetin-3,4´-diglucoside (Q-3,4´) are the two highest abundant ones [13]. In this study, extraction is combined with hydrolysis in order to remove the glucose, which increases the yield of the quercetin aglycone, but also simplifies the qualitative and quantitative analysis. One of the conventional ways to extract and hydrolyse quercetin glucosides from onion is to use aqueous methanol with high concentrations of HCl [14-15]. We have previously shown that the use of PHWE in combination with enzyme-catalysed hydrolysis is a greener alternative compared to the use of aqueous methanol with high concentrations of HCl [16]. Another disadvantage of using the conventional method is that the conditions are quite harsh and quercetin may degrade [14-15]. In this paper, we have developed a continuous flow system for combined PHWE and enzyme-catalysed hydrolysis of quercetin glucosides from onion. Some of the advantages of on-line coupling different techniques are less sample handling and shorter process time.

In this study, a thermostable beta-glucosidase expressed from *Thermotoga neapolitana* has been used enabling the coupling of PHWE with hydrolysis in a continuous flow system without extensive cooling between the extraction and the hydrolysis step. A mutated enzyme (*Tn*Bgl1A_N221S/P324L) was used since earlier results have shown that this has a higher affinity for quercetin-3-glucoside (Q-3) and Q-4´, compared to the wild type enzyme (*Tn*Bgl1A) [16]. *Tn*Bgl1A_N221S/P324L has been immobilised to two different support materials [17], cryogel and Eupergit® C 250L and these two enzymes preparation were used in this paper. By immobilising the enzyme, it can be used in a flow system and it can be re-used. The optimal conditions for hydrolysis are not for sure the same as the extraction, but since the enzyme is more sensitive to different conditions, the hydrolysis conditions were first optimized. A surface response design with three factors; pH, temperature and ethanol volume concentration, and with three levels was used in the hydrolysis optimisation. The optimal conditions were used to study the extraction profiles of quercetin species from yellow onion, but also to investigate the effect of a decrease in pH and changing flow rate. Extraction

and hydrolysis of quercetin glucosides from yellow, red and shallot onion was also compared to a conventional method using aqueous methanol with 1.2M of HCl.

2. Material and Methods

2.1 Materials

Formic acid was purchased from Merck (Darmstadt, Germany). Methanol, quercetin, quercetin 3-β-D-glucoside (Q-3), citric acid monohydrate, disodium hydrogen phosphate and *para*-nitrophenyl-β-D-glucopyranoside *(p*NPG) were purchased from Sigma-Aldrich (Steinheim, Germany). Quercetin 4´-*O*-β-glucopyranoside (Q-4´) and quercetin 3,4´-di-*O*-β-glucopyranoside (Q-3,4´) were purchased from Polyphenols Laboratories AB (Sandnes, Norway). Morin was purchased from Extrasynthese (Lyon, France). Ultrapure water (MilliQ) was used at all times.

2.2 Extraction and hydrolysis system

A home built on-line flow extraction and hydrolysis system was built (Figure 1). To control the flow, an HPLC-pump, HP1050 (Agilent Technologies, Waldbronn, Germany) was used, and to set the temperature for hydrolysis/extraction a GC-oven, HP5890 (Agilent Technologies, Waldbronn, Germany), was used and for time point sampling a fraction collector, Waters fraction collector III (Waters, Milford, US) was used. Tubings in the system were 1/16-inch stainless steel and before the extraction cell a 2-m long coil was placed to ensure equilibration of temperature of the solvent. For extraction of onion, a 10x50 mm Dionex extraction cell with stainless steel inlet and outlet filters was used. Immobilised enzyme was packed in a home built holder of peek (6x12 mm, 0.35 mL) and put in a guard column holder from Agilent Technologies.

2.3 Optimisation of hydrolysis pH, temperature and percentage ethanol

To design the optimisation method for the variables pH, temperature and volume concentration of ethanol, a surface response design with central composite design (CCD) in the Minitab® software was used. Each variable had three levels, the number of experiments was 20, which was divided up into three blocks (one block done within 1 day) and of the 20 experiments, 6 were replicates at the middle level of the parameters. The three levels were for pH; 3, 4.5 and 6, for temperature; 75, 85 and 95°C and for ethanol volume concentration; 5, 10 and 15%. The experiments were performed using the home built on-line continuous flow extraction system described in section *2.2 Extraction and hydrolysis system*. When the oven reached desired temperature, the pump was started, and the gradient used was 0-5 min 100% B, 5-5.5 min 100% A, 5.5-15.5 min 100% A, 15.5-16 min 100% B and 16-21 min 100% B, where solution A was 100 µM Q-3,4´ in 20 mM citrate-phosphate buffer and solution B was 20 mM citrate-phosphate buffer. Amount of immobilised enzyme used was either one piece of cryogel or 0.2 g of Eupergit® C 250L beads, corresponding to 45 µg and 26 µg of enzyme, respectively. The flow rate used was 3mL/min and samples were collected each minute. To the collected samples a fixed amount of morin (internal standard) dissolved in methanol with 5% formic acid was added. An ethanol-H_2O mixture was thereafter added so the final concentration of ethanol was the same for all samples collected (13 v-%). The collected samples were diluted three times with mobile phase (50:50 methanol:H_2O and 0.13 M of formic acid) before HPLC-UV analysis.

2.4 Yellow onion

Two kg of yellow onion, bought at a local supermarket in Sweden, was peeled and the bulb was chopped using an onion chopper to pieces around 5x5 mm and the peel was cut with a scissor. The cut peel and bulb was mixed and divided up in three portions and put in a freezer at -20°C.

2.5 Static extraction of yellow onion using PHWE

13-15 gram of chopped yellow onion was weighed into 33 mL stainless steel extraction cells, with a glass fiber (GF/B) filter in the bottom. Extraction was

done using a Dionex ASE-200 (Sunnyvalve, CA, USA) using water extraction solvent as described by Turner *et al.* [5]. The pressure was set to 50 bar, initial heating was 5 minutes, oven temperature was 120°C, extraction time was 15 min (5 min x 3 cycles). After extraction, the cells were purged with nitrogen. The collected extracts were pooled and diluted to a total volume of 1120 mL and divided up into 6 aliquots. The extracts were purged with argon and stored at -18°C.

2.6 Hydrolysis with 0 and 5% of ethanol at optimal pH and temperature

The reason, why in this and other experiments an onion extract was used, instead of real onion sample, was to have a constant flow of a certain composition of quercetin glucosides so any difference in the hydrolysis pattern is not due to the extraction kinetics and other sample components.

Yellow onion extract (*2.5 Static extraction of yellow onion using PHWE*) was thawn and 100 mL sample was prepared by mixing 20 mL onion extract, 70 mL of 28.5 mM citrate-phosphate buffer with 100 µM Q-3,4´ (pH 5.4 respectively 5.6), and 5 mL of 99.7% ethanol or 5 mL of H_2O and diluted too 100 mL with H_2O in a 100 mL volumetric flask. The pH was controlled and set to 5.4 or 5.6 with HCl or NaOH. *Tn*Bgl1A_N221S/P342L immobilised to cryogel (one piece, 45 µg of enzyme) or Eupergit® C 250L (0.2 g, 26 µg of enzyme) was loaded into the hydrolysis cell (*2.2 Extraction and hydrolysis system*) and the oven was set to 84°C. When the oven reached 84°C, the pump was started, the flow rate was set to 3 mL/min and the gradient used was 0-5 min 100% B, 5-5.5 min 100% A, 5.5-15.5 min 100% A, 15.5-16 min 100% B and 16-21 min 100% B, where solution A was the 100 mL sample solution above and solution B was 20 mM citrate-phosphate buffer pH 5.4 or 5.6. Fractions were collected for 1 minute each and as a control sample 3 mL of the solution A was used. To all fractions 200 µL of 0.342 mg/mL of morin and 300 µL of aqueous ethanol with 10% formic acid was added, so the concentration of morin, formic acid and ethanol was the same in all collected samples. Formic acid was added to lower the pH below pH 3 so that quercetin would be uncharged. 500 µL of the collected fractions was added to 1000 µL of 50:50 methanol:H_2O and analysed with HPLC-UV.

2.7 Online yellow onion extraction using different ethanol volume concentration, buffer pH and flow rate

Four different extraction experiments have been done, see Table 1 for difference in used flow rate, pH and volume concentration of ethanol.

Table 1

The ion strength of the citrate-phosphate buffer was 20 mM for all 4 experiments and pH was chosen to be 5.5 in experiment 1, 2 and 4 (the middle value of the optimal pH for Eupergit® C 250L and cryogel) and pH 2.6 in experiment 3. Before the extraction started, 2.5 g of onion sample was put in the extraction cell and the oven temperature was set to 84°C. When the oven temperature reached 84°C, the pump was started and extract was collected every minute up to 90 min or 240 min. Experiment 1, 2 and 4 were performed in triplicates and experiment 3 in duplicate.

After extraction at a flow rate of 3 mL/min, 200 µL morin (internal standard) in methanol (c.a. 0.3 mg/mL) and 300 µL mixture of 25% ethanol and 10% formic acid (aqueous solution) was added to the extract with 5% ethanol (Experiment 1 and 3), and 200 µL morin standard in MeOH (approx. 0.3 mg/mL) and 300 µL mixture of 75% ethanol and 10% formic acid (aqueous solution) to the extract with 0% ethanol (Experiment 3). For extract obtained at a flow rate of 1 mL/min (Experiment 4), the volume added of morin, mixture of ethanol and formic acid was proportionally smaller since the extract volume was decided by flow rate. Before HPLC-analysis, 0.5 mL of sample was taken out and diluted with 1 mL of 60:40 MeOH:H_2O (0.13 M formic acid). Only extracts collected each fifth or tenth minute were analysed by HPLC.
Obtained concentration profile was fitted by applying an interpolate curve fit in Kaleida Graph (v 4.03, Synergy Software), and the concentration at each minute was calculated by the function of the fitted curve, using Formula Entry.
Cumulative amount of Q-3,4', Q-4', and quercetin were then obtained using these fitted points.

2.8 Conventional extraction and hydrolysis

The conventional extraction and hydrolysis was done according to Nuttila *et al*. [15], with some modification. 1 g of onion (yellow, red or shallot onion), 4 mg ascorbic acid and 40 mL of methanol/1.2 M HCl (1:1) was added to a glass vial. Extraction and hydrolysis was performed at 80°C for 2h under magnetic stirring in a closed vial. The extracts were cooled to room temperature and 1 mL of each extract was collected and centrifuged. 200 µL of centrifuged extract was mixed with 700 µL of 7:3 H_2O:MeOH and 100 µL of 0.031 mg/mL morin in MeOH and analysed as soon as possible with HPLC-UV. Yellow onion extraction was repeated six times and red- and shallot onion extraction was done in triplicates.

2.9 Online extraction and hydrolysis of yellow, red and shallot onion

In the online extraction and hydrolysis experiments, 20 mM citrate-phosphate buffer, pH 5.5, with 5% of ethanol was used as extraction solvent, the oven temperature was set to 84°C, the flow was 0.75 mL/min and the extraction was done for 120 minutes. When the extraction was combined with hydrolysis the amount of immobilised enzyme used was either one piece of cryogel or 0.2 g of Eupergit® C 250L beads, corresponding to 45 µg and 26 µg of enzyme, respectively, and both these enzyme preparations filled up the reaction vessel of 0.35 mL (6x12 mm). The amount of onion used was 2.5 g for yellow onion, 1.5 g for red onion and 2 gram for shallot onion. After placing the onion in the extraction cell, 1mL of 0.31 mg/mL of morin dissolved in methanol was added and the cell was put under nitrogen stream to evaporate the methanol. The collected extract was diluted to 250 mL with 1:1 ethanol:H_2O. Triplicate extraction and hydrolysis was done for all three onion varieties. Obtained extract was centrifuged and 500 µL of the supernatant was mixed with 450 µL of 7:3 H_2O:methanol and 50 µL of 10% formic acid. The samples were analysed by HPLC-UV.

2.10 HPLC-UV

The HPLC-UV system used was an UltiMate 3000 from Dionex (Thermo Fisher, Germering, Germany). The analytical column used was an Agilent Zorbax SB-C18 (2.1x100 mm, 3.5 µm), with an Agilent Zorbax SB-C8 pre-column (2.1 x 12.5 mm, 5 µm). Depending on the sample, isocratic or gradient elution was used, solvent A: methanol with 0.5% formic acid and solvent B: H_2O, with 0.5% formic acid. For isocratic elution 40 % A was used. In the case of gradient elution either of two following program was applied: (*i*) 40% A 0-35 min, 80% A 35-40 min and 40% A 40-60 min or (*ii*) 40% A 0-27 min, 40-100% A 27-28 min, 100% A 28-43 min, 100-40% A 43-45 min, 40% A 45-60 min. The flow rate of the mobile phase was 0.15 mL/min, the temperature of the column oven was ambient, injection volume was 5 µL and detection was accomplished at 350 nm. Identification and quantification of Q-3,4´, Q-3, Q-4´, quercetin and morin was done using standards and a five point external calibration curve between 0.5-100 µmol/L.

2.11 Stabilisation study of immobilised enzyme

Enzyme activity was estimated at 84°C using *p*NPGlc as substrate in 20 mM citrate phosphate buffer pH 5.6. The reaction measures the release of para-nitrophenol (*p*NP) and was determined spectrophotometrically at 405 nm. A certain quantity of immobilised enzyme beads equivalent to 1.2 µg of *Tn*Bgl1A_N221S/P342L enzyme on Eupergit® C 250L beads in 500 µL citrate phosphate buffer at pH 5.6 (20 mM) was pre-incubated at the specific reaction temperature for 0, 1, 3, 5, 8, 12, 24, 48, 72, 96, 120 and 144 h and in parallel the *p*NPGlc solution (1.46 mM) was pre-heated at the same conditions for 10 min before starting the activity test at certain time interval. 1 mL of the pre-heated *p*NPGlc solution was added to the pre-incubated immobilised enzyme beads. The sample was incubated at 84°C temperature for 1 min and put on ice for 5 min. After 5 min, the samples were removed from the ice and left at room temperature for another 5 min, and then the absorbance was measured spectrophotometrically at 405 nm. The specific activity (U/mg) of the immobilised enzyme was determined and all measurements were performed in triplicate. Enzyme activity was expressed as µmol of *p*NP formed per min at the respective reaction temperature. The decrease in enzymes activity was

monitored with time to determine its stability at the conditions above. Two sample t-test rejecting null hypothesis at 5% significance level was applied on the data to statically verify the significant decrease in stability with time.

3. Results and Discussion

The hydrolysis conditions were first optimised since the enzyme activity and stability strongly depends on both temperature and pH, which also limit the performance of the on-line extraction and hydrolysis process. The extraction in a combined on-line system may thus be slower since the temperature is adjusted to match the activity and stability of the enzyme, as compared to a separate extraction step with optimal temperature for obtaining quercetin species from yellow onion.

3.1 Optimisation of pH, temperature and ethanol percentage for hydrolysis of Q-3,4´

In the optimisation of the hydrolysis, a standard solution of Q-3,4´ was used and three parameters were studied, pH (3-6), temperature (75-95°C) and ethanol volume concentration (5-15%). Ethanol was added to improve the solubility of quercetin, thus minimising losses of the analyte. In Table 2 below the results of the statistical analysis of the surface response design are listed, and the terms that were significant with a 95% confidence interval were pH, ethanol and pH x pH. The R-square values in Table 2 shows that the response surface design was a satisfying match between the model and the experimental data.

Table 2

In the counter plots in Figure 2 it is seen that the hydrolysis works best at low percentage of ethanol, at pH between 5 and 6 and at temperatures around 80-85°C for both of the tested support materials. In Table 3 the optimal pH, temperature and percentage of ethanol from the surface response design can be seen.

Table 3

Figure 2

The temperature interval tested did not have any significant effect (Table 2), which is probably due to that the immobilised enzyme is sufficiently stable at all the temperatures tested [17], so there is no decrease in activity due to thermal degradation. If the hydrolysis had been studied continuously for several hours or at higher or lower temperatures, the temperature might have had a significant effect. The optimal temperature in Table 3 is 84°C for the enzyme, and no difference regarding which support material that was used. These results agree with Khan *et al*. [17] where the specific activity was measured of the immobilised enzyme at 80, 85, 90 and 95°C and no large difference in the activity at the different temperatures were seen. However, the hydrolysis was only studied for a few minutes, which is why a long-term stability test was done in this work at 84°C (section *3.4 Stabilisation study of immobilised enzyme*).

pH on the other hand turned out to have a significant effect and the activity decreased quite rapidly at higher respectively lower tested pHs. The reason why a broader pH interval was selected, despite an already known pH optimum around pH 5 [5], was that pH might have an effect on the extraction step, for instance affecting the solubility of quercetin. When an enzyme is immobilised there might be a change in optimal pH compared to enzyme in solution. The optimal pH value as listed in Table 3 is pH 5.4 (cryogel) respectively 5.6 (Eupergit® C 250L), which is higher compared to pH 5 which was the optimal temperature for enzyme in solution [5]. The difference in optimal pH (Table 3) between the two support materials with 0.2 pH-units might be due to experimental errors, because the RSD values of the six replicates for cryogel and Eupergit® C 250L was 9.6% and 7.5%, respectively, taken at the central point.

The reason why ethanol was included in the optimisation is that the addition of ethanol to the water in the future extraction can increase the extraction efficiency of the quercetin species from onion as well as keep the compounds solubilised when the collected extract has cooled down. Ethanol had a significant

effect, see Table 2, and the optimal volume concentration of ethanol was 5% for the enzyme immobilised on both cryogel and Eupergit® C 250L, Table 3. Ethanol might have a destabilising effect of the enzyme, which might have an apparent effect in the temperature optimum.

Since the optimal ethanol volume concentration was the lowest volume concentration of ethanol included in the design, another experiment was done to see if a lower volume concentration of ethanol, here 0%, would improve the hydrolysis compared to 5%. In this control experiment, a mixture between buffer and onion extract, spiked with Q-3,4´, was used. The conversion efficiency of quercetin glucosides to quercetin between 12 and 14 minutes (in the pump program) was compared to the total amount of quercetin species in equal volume of the spiked onion-buffer solution before hydrolysis, using 0 respectively 5% of ethanol. For enzyme immobilised on cryogel there was a very small difference in conversion efficiency, 71 and 72%, with 0 and 5% ethanol, respectively. In the case of enzyme immobilised on Eupergit® C 250L, the conversion efficiency was 65 and 77%, with 0 and 5% ethanol, respectively. Since the conversion efficiency was similar or higher with 5% ethanol, it was concluded that 5% ethanol was motivated to use. From an extraction point of view, higher volume concentration of ethanol might increase the extraction rate and/or yield of quercetin, but such increase may on the other hand decrease the enzyme activity.

3.2 Extraction profiles of yellow onion using different ethanol volume concentration, pH and flow rate

The optimal hydrolysis condition from the experimental design was 5% ethanol, 84°C and pH 5.4 (cryogel) or 5.6 (Eupergit® C 250L). However, to study possible influence on the extraction efficiency and yield with onion extraction, other experimental conditions were investigated. Figure 3 and Table 1 shows the accumulative amount of extracted Q-4´, Q-3,4´ and quercetin, µmol/g f.w. yellow onion using a flow rate of 3 mL/min at 84°C. It is clear from Table 1 that the addition of ethanol to the buffer facilitates the extraction of all three quercetin species from onion sample. Addition of ethanol may increase the solubility of quercetin species in the extract, thus has a positive effect on the extraction.

Figure 3

Buffer pH can also affect the solubility of quercetin species in water. Quercetin has at least three ionisable hydroxyl groups, of which one has the lowest pKa value of 6.7-7, depending on the measurement method [18-19]. At a pH of 5.5, more quercetin is charged, according to weak-acid equilibrium in aqueous solution, thus is more soluble compared to using a pH of 2.6. However, if high volume concentration of ethanol was used as extraction solvent, a pH as low as 2 was favourable [20]. As shown in Figure 3 the total yield at pH 2.6 is considerably lower than extraction at pH 5.5. In summary, 5% ethanol and pH 5.5 were chosen to be used in the on-line extraction and hydrolysis of quercetin species.

Flow rate is considered to play an important role in extraction. Increasing flow rate will increase the mass transfer of the extracted target compounds. Higher flow rate will maintain a negative concentration gradient, which will drive the extraction forward as determined by the distribution constants of each analyte species. If an extraction is solubility-controlled, i.e. the solubility of the analytes determines the rate of the extraction and desorption kinetics can be neglected, the plots of the extraction yield versus time are straight lines and the slope is proportional to the flow rate [21]. This is exactly what can be seen in Figure 4a in the very beginning of the curves. A solubility-controlled mode of extraction kinetics can be proved by plotting the accumulative amount of quercetin species (Q-3,4', Q-4' and quercetin) against total extraction volume (Figure 4b). Before the extraction volume reached 12 mL, the accumulative yield of quercetin species per volume is the same, regardless of the flow rate used. However, after this point, the curves start to deviate and the extraction rate slowed down (corresponding to after 4 min at a flow rate of 3 mL/min and 12 min at a flow rate of 1 mL/min in Figure 4a), demonstrating that desorption/diffusion effects began to intervene. The very slow and similar extraction rate in the end of the extraction using both flow rates shows that desorption/diffusion effects completely take control of the extraction kinetics, rather than solubility. The total amount of quercetin species extracted by a flow rate of 1 mL/min is

considerably lower than that by 3 mL/min, indicating also the presence of either degradation or adsorption of extracted compounds. Degradation of quercetin has been observed after 15 min in subcritical water extraction of quercetin from onion[4], although conducted at significantly higher temperature (160°C). A study on the degradation of Q-3,4', Q-4' and quercetin standard in a mixture of ethanol, water and formic acid (94:5:1, v/v/v) also showed pronounced degradation of the quercetin species at 110 °C (unpublished data).

Figure 4

The highest total amount of extracted Q-3,4', Q-4' and quercetin (experiment 1, Table 1) during 90 min was 1.4, 2.7 and 3.7 μmol/g (f.w.), respectively. All in all, the best extraction yield in this study was obtained using a buffer pH 5.5, ethanol volume concentration of 5%, and a flow rate of 3 mL/min.

3.3 Combined extraction and hydrolysis, including comparison to conventional extraction

When extraction and hydrolysis of quercetin glucosides from onion was combined in the continuous on-line flow system the flow rate used was 0.75 mL/min, the total extraction/hydrolysis time was 120 min and only *Tn*Bgl1A_N221S/P342L immobilised on Eupergit® C 250L was tested. In Table 4 the quercetin yield from combined extraction and hydrolysis for yellow, red and shallot onion are listed. In Table 4 the results of using the conventional method are also listed.

Table 4

The amount of quercetin quantified in yellow, red and shallot onion using continuous flow on-line extraction and hydrolysis, was compared to using a conventional method, since the quercetin content from one onion to another can vary a lot and therefore it is not trivial to compare the results to other published results.

The conventional method used to extract and hydrolyse quercetin from onion involved the use of aqueous methanol and high concentration of hydrochloric acid. A modified version of the method published by Nuutila *et al*. [15] was employed, i.e. extraction and hydrolysis was done for 120 min at 80°C with 50% methanol and 1.2 M HCl and a few mg of ascorbic acid. The quercetin yield using the on-line extraction and hydrolysis method with enzyme immobilised on Eupergit® C 250L was higher for all three tested onion types, compared to the conventional method (Table 4). The reason why the extraction yield is higher with the new method is that is mild compared to the harsh conditions (1.2 M HCl) used in the conventional method. Hertog *et al*. [14] and Nuutila *et al*. [15] have shown that a longer times than 2 h with 1.2M of HCl leads to degradation of quercetin.

The RSD of the replicates of combined extraction and hydrolysis with red- and shallot onion using the continuous flow system are high, 20-30%. The reason for these large RSD values can for example be that small amount of onion was used in the experiments, 1.5-2.0 g, and it is not easy to weigh in a representative sample, since the amount of quercetin species differ throughout the onion [5, 22]. If larger sample amounts, smaller particle size and/or more replicates had been done, the RSD had probably been lower. Another advantage of using a continuous flow system, compared to a static batch extraction as the one used in the conventional method, is that the extraction solvent is continuously replaced and the solution does not become saturated with quercetin species.

When the extraction was combined with hydrolysis, using immobilised enzyme, a lower flow rate, 0.75 mL/min, had to be used to ensure full hydrolysis. When higher flow rates were used, the hydrolysis was not complete and quercetin glucosides were identified in the extract (data not shown). The extraction and hydrolysis time was only prolonged to 120 min. Table 4 shows that a decrease in the flow rate from 1 mL/min (only extraction) to 0.75 mL/min (extraction + hydrolysis) decreased the total quercetin yield from 4.6 µmol/g f.w. to 3.2 µmol/g f.w.. Extraction without hydrolysis was also tested using the lower flow rate and 120 min, to ensure that there was no adsorption of quercetin to the

support material or any other internal surfaces. The total quercetin yield was then 3.1 μmol/g f.w., i.e. most likely there was no adsorption of quercetin.

The quercetin yield in the combined extraction and hydrolysis of yellow onion using a flow rate of 0.75 mL/min, 3.2 μmol/g f.w., is much lower compared to the total quercetin yield when only the extraction was studied using yellow onion at a flow rate of 3 mL/min, 7.8 μmol/g f.w.. In the continuous flow extraction and hydrolysis setup it was selected to decrease the flow rate and not to increase the amount of *Tn*Bgl1A_N221S/P342L and/or use lower amount of onion. The reason why higher amounts of enzyme could not be used was that the enzyme cell was already fully packed with immobilised *Tn*Bgl1A_N221S/P342L. The amount of yellow onion used was 2.5 g and a lower amount of onion might have resulted in the possibility to use higher flow rate. However, smaller sample size can result in even higher deviation between the replicates. Future experiments will include the use of a longer reaction vessel combined with the use of as high flow rate as necessary to avoid possible degradation of extracted quercetin species and to increase the quercetin yield per time unit.

3.4 Stabilisation study of immobilised enzyme

The unfolding temperature for this enzyme is 100.9°C [23] and the half-life of the enzymes in solution at 85°C is 7 h [16]. Stability of enzymes often increases upon immobilisation [24] and thermostable enzymes can be further stabilised by covalent immobilisation as seen in the case of thermophilic esterases from *Bacillus stearothermophilus* [25]. The stability and reusability of the enzyme has been checked at temperatures up to 95°C, but no loss in activity was observed during the reusability experiments. The reason for this could be that the total exposure time of the immobilised enzyme was around 2 h and also the enzyme was cooled to room temperature and washed after each 11 min [17]. There is no change in the hydrolytic activity at 84°C (optimum temperature for immobilised enzymes) for 5 h when the immobilised form of this enzyme was used in the online flow system (data not shown). To confirm its long-term stability the enzyme was exposed to prolonged heating at the reaction temperature and the activity was monitored for 144 h using *p*NPGlc as substrate. No significant

decrease in activity was observed up to 48 h (Figure 5), which was verified statistically by using two sample t-test with 95% confidence level.

Figure 5

When the extraction was combined with enzymatic hydrolysis in the online flow system, 5% ethanol was used in the extraction procedure but no ethanol was included in the stability studies. In another study it was shown that the half-life of lipase was increased from 85-90 min to 4 h on immobilisation at 70°C [26] while that of catalase was increased from 9 to 29.1 h upon immobilisation on Eupergit® C at 40°C [27]. Compared to several other studies stability of the enzyme studied here is good enough and this enzyme in immobilised form is a suitable candidate for the applications carried out at high temperature. However, to confirm that the enzyme stability is not affected by the addition of ethanol, another stability experiment will be conducted with the enzyme in the presence of ethanol.

4. Conclusions

The optimal conditions for hydrolysis of quercetin glucosides using *Tn*Bgl1A_N221S/P342L immobilised on cryogel and Eupergit® C 250L using a standard solution of Q-3,4´ were 84°C, pH 5.5 (cryogel) or 5.6 (Eupergit® C 250L) and 5% of ethanol. When the extraction profiles of total quercetin yield from yellow onion were studied, the highest yield was achieved when a flow rate of 3 mL/min was used at a pH of 5.5, temperature of 84°C and 5% of ethanol in the extraction solvent. If a lower pH was used, pH 2.6, the extraction yield decreased compared to at pH 5.5. A decrease in volume concentration of ethanol, from 5% to 0%, also resulted in a lower quercetin extraction yield at pH 5.5. Increasing the flow rate from 1 mL/min to 3 mL/min resulted in a constant yield per solvent volume unit, up to 12 mL, i.e. the extraction was solubility-controlled. However, when desorption/diffusion effects took over and possible degradation of quercetin species began to emerge, a higher flow rate was advantageous. The

online extraction and enzymatic hydrolysis method gave higher quercetin extraction yield compared to using a conventional static extraction with aqueous methanol and 1.2 M of HCl. The developed online extraction and hydrolysis using a water/ethanol mixture and enzyme is slower compared to offline extraction and hydrolysis. By improving the set-up of the online system, the total time can probably be decreased from the 2 h. Improvement can be optimizing the onion enzyme ratio as well as the design of the hydrolysis cell. The dimension of the hydrolysis cell used was 12x6 mm, but by changing the dimension of the cell so that the residence time of the extract in the hydrolysis cell is prolonged a higher flow rate might be used. Temperature stability measurement showed that there was no significant loss in activity up to 48 h at 84°C.

5. Acknowledgment

The authors wish to thank the Swedish Research Council Formas (Sustain-X-Enz 2006-1346, SuReTech 2009-1527), the Swedish Research Council (VR, 2010-333) and Higher Education Commission (HEC) of Pakistan for funding. Rolf Danielsson, Uppsala University, and Jan Åke Jönsson, Lund University, for help with the statistical analysis.

6. References

1. M.G.L. Hertog, P.C.H. Hollman, M.B. Katan, Content of potentially anticarcinogenic flavonoids of 28 vegetables and 9 fruits commonly consumed in the Netherlands. *Journal of Agricultural and Food Chemistry* **1992**, 40 (12), 2379-2383.
2. C.D. Stalikas, Extraction, separation, and detection methods for phenolic acids and flavonoids. *Journal of Separation Science* **2007**, 30 (18), 3268-3295.
3. A. Ekman, M. Campos, S. Lindahl, M. Co, P. Börjesson, E.N. Karlsson, C. Turner, Addition of value in bioresource utilization by sustainable technologies in new biorefinery concepts, 2011, Manuscript
4. M.J. Ko, C.I. Cheigh, S.W. Cho, M.S. Chung, Subcritical water extraction of flavonol quercetin from onion skin. *Journal of Food Engineering* **2011**, 102 (4), 327-333.
5. C. Turner, P. Turner, G. Jacobson, K. Almgren, M. Waldeback, P. Sjöberg, E.N. Karlsson, K.E. Markides, Subcritical water extraction and beta-glucosidase-catalyzed hydrolysis of quercetin glycosides in onion waste. *Green Chemistry* **2006**, 8 (11), 949-959.
6. M.S.Y. Kumar, R. Dutta, D. Prasad, K. Misra, Subcritical water extraction of antioxidant compounds from Seabuckthorn (Hippophae rhamnoides) leaves for the comparative evaluation of antioxidant activity. *Food Chemistry* **2011**, 127 (3), 1309-1316.
7. K. Hartonen, J. Parshintsev, K. Sandberg, E. Bergelin, L. Nisula, M.-L. Riekkola, Isolation of flavonoids from aspen knotwood by pressurized hot water extraction and comparison with other extraction techniques. *Talanta* **2007**, 74 (1), 32-38.
8. K. Srinivas, J.W. King, L.R. Howard, J.K. Monrad, Solubility and solution thermodynamic properties of quercetin and quercetin dihydrate in subcritical water. *Journal of Food Engineering* **2010**, 100 (2), 208-218.
9. S.B. Hawthorne, Y. Yang, D.J. Miller, Extraction of organic pollutants from environmental solids with subcritical and supercritical water. *Analytical Chemistry* **1994**, 66 (18), 2912-2920.
10. E.V. Petersson, J.Y. Liu, P.J.R. Sjoberg, R. Danielsson, C. Turner, Pressurized Hot Water Extraction of anthocyanins from red onion: A study on extraction and degradation rates. *Analytica Chimica Acta* **2010**, 663 (1), 27-32.
11. A. Murakami, H. Ashida, J. Terao, Multitargeted cancer prevention by quercetin. *Cancer Letters* **2008**, 269 (2), 315-325.
12. K. Ono, T. Hamaguchi, H. Naiki, M. Yamada, Anti-amyloidogenic effects of antioxidants: Implications for the prevention and therapeutics of Alzheimer's disease. *Biochimica Et Biophysica Acta-Molecular Basis of Disease* **2006**, 1762 (6), 575-586.
13. R. Slimestad, T. Fossen, I.M. Vagen, Onions: A source of unique dietary flavonoids. *Journal of Agricultural and Food Chemistry* **2007**, 55 (25), 10067-10080.
14. M.G.L. Hertog, P.C.H. Hollman, D.P. Venema, Optimization of a Quantitative HPLC Determination of Potentially Anticarcinogenic Flavonoids in

Vegetables and Fruits. *Journal of Agricultural and Food Chemistry* **1992**, 40 (9), 1591-1598.
15. A.M. Nuutila, K. Kammiovirta, K.M. Oksman-Caldentey, Comparison of methods for the hydrolysis of flavonoids and phenolic acids from onion and spinach for HPLC analysis. *Food Chemistry* **2002**, 76 (4), 519-525.
16. S. Lindahl, A. Ekman, S. Khan, C. Wennerberg, P. Borjesson, P.J.R. Sjoberg, E.N. Karlsson, C. Turner, Exploring the possibility of using a thermostable mutant of beta-glucosidase for rapid hydrolysis of quercetin glucosides in hot water. *Green Chemistry* **2010**, 12 (1), 159-168.
17. S. Khan, S. Lindahl, C. Turner, E. Nordberg Karlsson, Immobilization of thermostable beta-glucosidase variants on acrylic supports for biocatalytic processes in hot water, 2011, Submitted manuscript
18. J.M. Herrero-Martinez, M. Sanmartin, M. Roses, E. Bosch, C. Rafols, Determination of dissociation constants of flavonoids by capillary electrophoresis. *Electrophoresis* **2005**, 26 (10), 1886-1895.
19. S.V. Jovanovic, S. Steenken, M. Tosic, B. Marjanovic, M.G. Simic, Flavonoids as antioxidants. *Journal of the American Chemical Society* **1994**, 116 (11), 4846-4851.
20. E. Kiassos, S. Mylonaki, D.P. Makris, P. Kefalas, Implementation of response surface methodology to optimise extraction of onion (Allium cepa) solid waste phenolics. *Innovative Food Science & Emerging Technologies* **2009**, 10 (2), 246-252.
21. S.B. Hawthorne, A.B. Galy, V.O. Schmitt, D.J. Miller, Effect of SFE flow-rate on extraction rates - classifying sample extraction behavior. *Analytical Chemistry* **1995**, 67 (15), 2723-2732.
22. S. Hirota, T. Shimoda, U. Takahama, Tissue and spatial distribution of flavonol and peroxidase in onion bulbs and stability of flavonol glucosides during boiling of the scales. *Journal of Agricultural and Food Chemistry* **1998**, 46 (9), 3497-3502.
23. S. Khan, T. Pozzo, M. Megyeri, S. Lindahl, A. Sundin, C. Turner, E.N. Karlsson, Aglycone specificity of Thermotoga neapolitana beta-glucosidase 1A modified by mutagenesis, leading to increased catalytic efficiency in quercetin-3-glucoside hydrolysis. *Bmc Biochemistry* **2011**, 12.
24. A.M. Klibanov, Enzyme stabilization by immobilization. *Analytical Biochemistry* **1979**, 93 (1), 1-25.
25. R. Fernandezlafuente, D.A. Cowan, A.N.P. Wood, Hyperstabilization of a thermophilic esterase by multipoint covalent attachment. *Enzyme and Microbial Technology* **1995**, 17 (4), 366-372.
26. N. Nawani, R. Singh, J. Kaur, Immobilization and stability studies of a lipase from thermophilic Bacillus sp: The effect of process parameters on immobilization of enzyme. *Electronic Journal of Biotechnology* **2006**, 9 (5), 559-565.
27. O. Alptekin, S.S. Tukel, D. Yildirim, D. Alagoz, Immobilization of catalase onto Eupergit C and its characterization. *Journal of Molecular Catalysis B-Enzymatic* **2010**, 64 (3-4), 177-183.

Figures and Tables

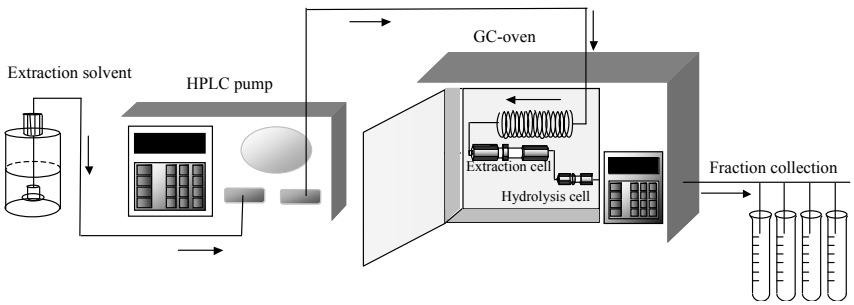
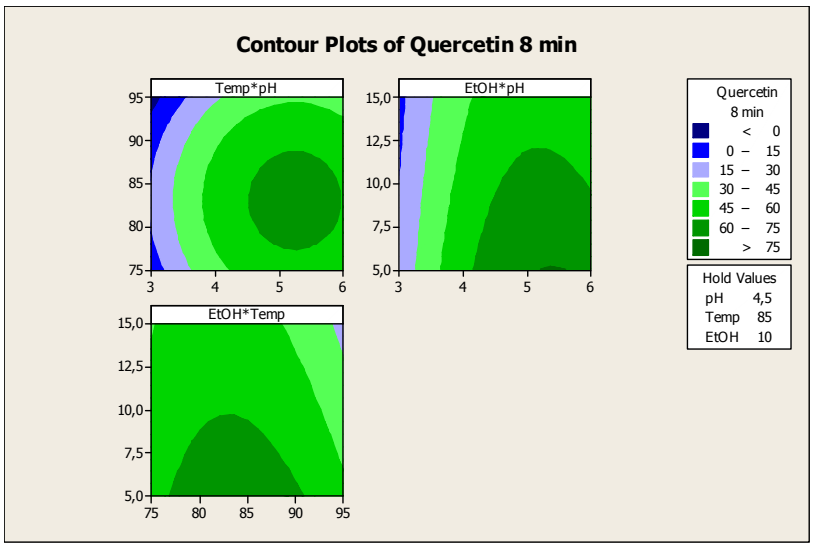


Figure 1. Schematic view of the online continuous flow system

(a)



(b)

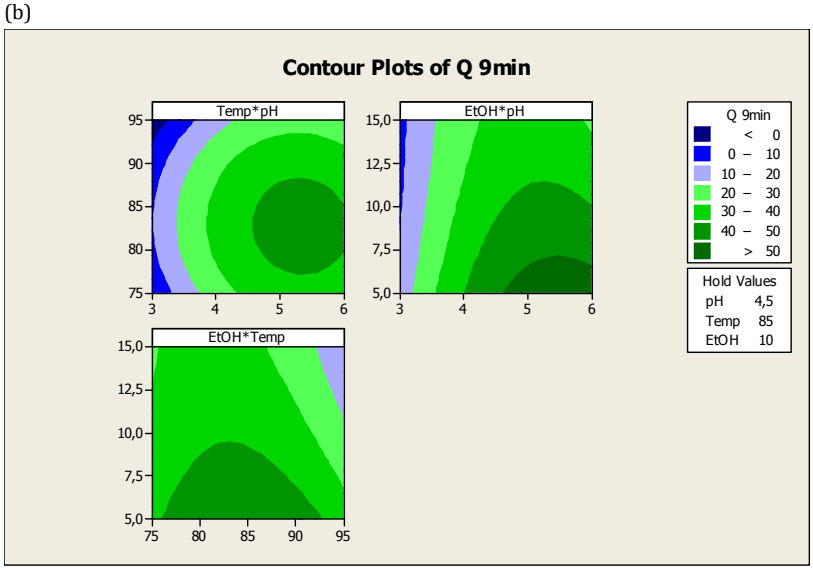


Figure 2. Counter plot of amount detected quercetin in 8 minute or 9 minute fraction if *Tn*Bgl1A_N221S/P342L was immobilised on (a) cryogel and (b) Eupergit® C 250L, respectively.

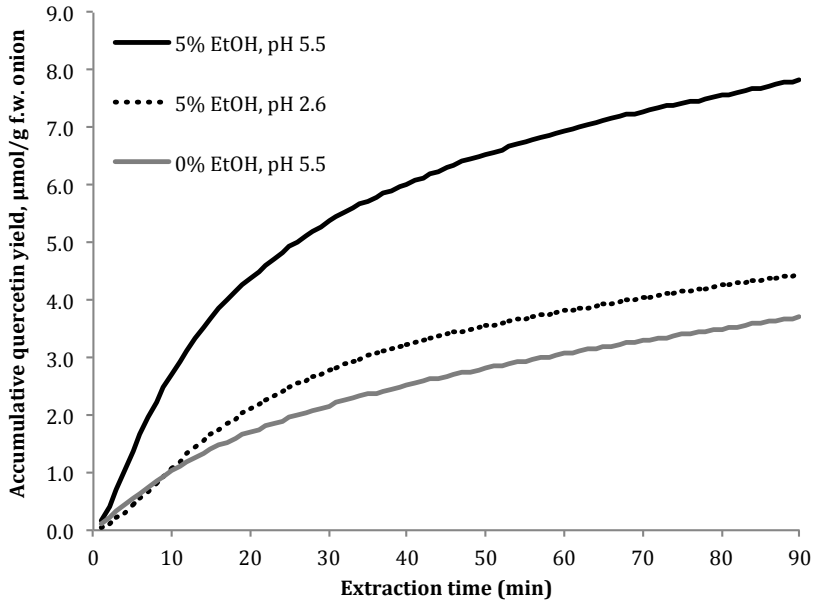
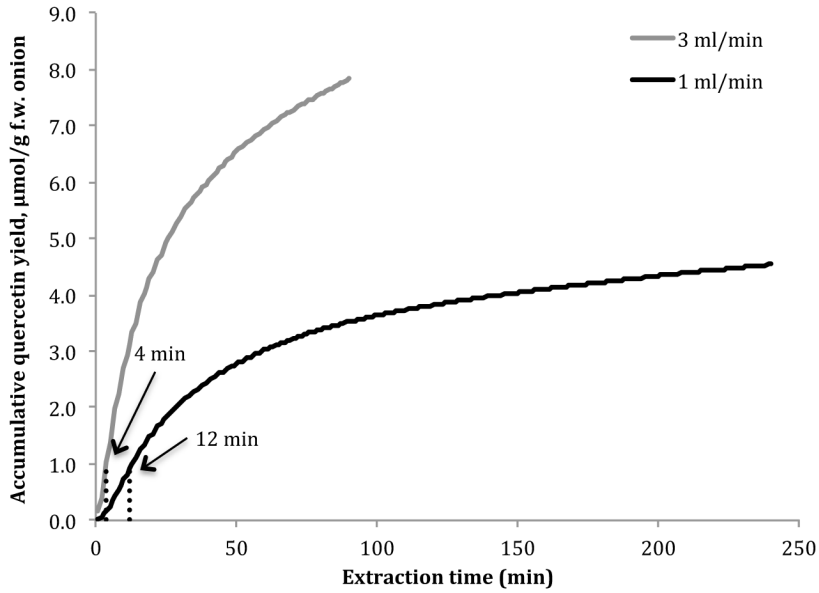


Figure 3. Accumulative quercetin yield of quercetin species (Q-3,4′, Q-4′ and quercetin) as function of extraction time using a flow rate of 3 mL/min at 84°C. f.w. = fresh weight

(a)



(b)

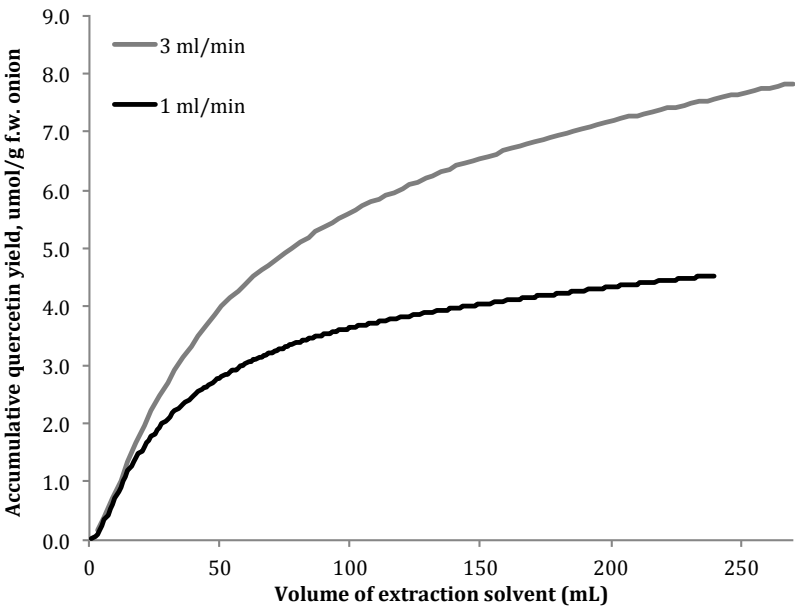


Figure 4. Extraction profiles of total accumulative quercetin species (Q-3,4′, Q-4′ and quercetin) yield as function of (a) time and (b) volume of extraction solvent using a flow rate of 1 mL/min or 3 mL/min. f.w. = fresh weight

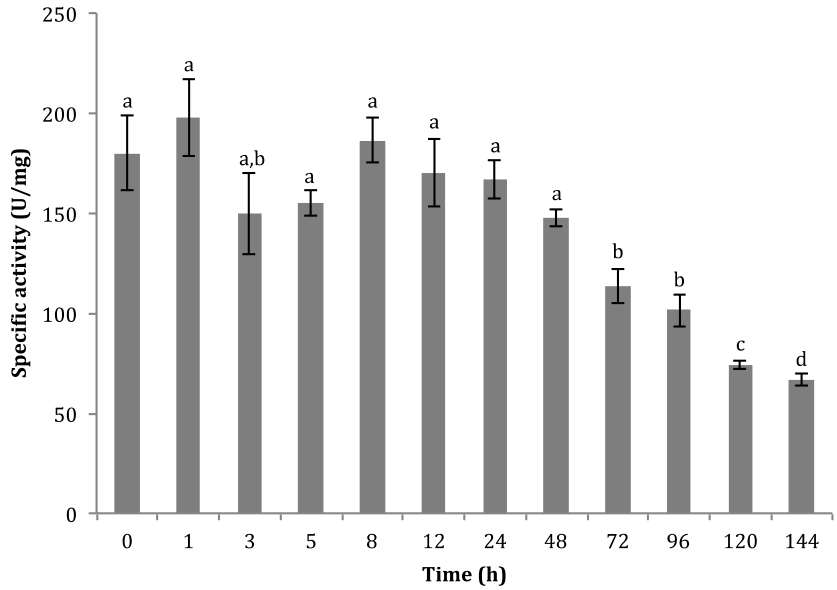


Figure 5. Specific activity of enzyme using *p*NPGlc as substrate after 0-144 h of incubation of *Tn*Bgl1A_N221S/P342L at 84°C. The data is divided into four significant groups (a, b, c and d) on the basis of t.test, within one group the values lie closer with no significance difference.

Table 1. Extraction yield of Q-3,4′, Q-4′ and quercetin, µmol/g f.w. onion from the four different extraction experiments. f.w. = fresh weight

Experiment	Flow (mL/min)	Extraction time (min)	Solvent volume (mL)	pH	Ethanol (%)	Q-3,4′ (µmol/g onion f.w)	Q-4′ (µmol/g onion f.w)	Quercetin (µmol/g onion f.w)	Total quercetin (µmol /g onion f.w)
1	3	90	270	5.5	5	1.4	2.7	3.7	7.8 ± 1.1
2	3	90	270	5.5	0	0.4	1.2	2.1	3.7 ± 1.4
3	3	90	270	2.6	5	0.8	1.3	2.3	4.5 [a]
4	1	240	240	5.5	5	0.9	1.2	2.5	4.6 ± 0.1

[a] n=2, i.e. no standard deviation calculated

Table 2. Result of the surface response design for cryogel and Eupergit® C 250L, using the quercetin concentration in the 8 min fraction as output. Figures in bold indicate that the effect is significant in a 95% confidence interval.

	Cryogel			Eupergit® C 250L		
Term	Coefficient	SE Coefficient	P	Coefficient	SE Coefficient	P
Constant	58,90	4,07		38,63	3,33	
pH	43,70	7,49	**0,0002**	31,12	6,13	**0,0005**
T	-12,59	7,49	0,1237	-10,08	6,13	0,1313
EtOH	-17,09	7,49	**0,0456**	-16,33	6,13	**0,0238**
pH x pH	-43,733	14,23	**0,0120**	-28,14	11,69	**0,0369**
pH x T	0,24	8,37	0,9777	-1,71	6,86	0,8084
pH x EtOH	-8,98	8,37	0,3086	-10,02	6,86	0,1744
T x T	-31,04	14,28	0,0548	-23,56	11,69	0,0716
T x EtOH	-5,65	8,37	0,5149	-7,41	6,86	0,3055
EtOH x EtOH	-0,06	14,28	0,9966	3,08	11,69	0,7978

Cryogel: R-sq. = 89,4% R-sq. (adj.) = 79,8% Standard error of Est = 11,8

Eupergit® C 250L*: R-sq. = 86,7% R-sq. (adj.) = 74,7% Standard error of Est = 9,7*

Table 3. Optimal pH, temperature and vol% ethanol for hydrolysis Q-3,4´ in the continuous flow system.

Support material	pH	Temperature (°C)	Ethanol (vol%)
Eupergit® C 250L	5.6	84	5
Cryogel	5.4	84	5

Table 4. Amount of extracted quercetin from yellow, red and shallot onion using conventional extraction and hydrolysis method (µmol/g f.w.)

Method	**Yellow onion**	**Red onion**	**Shallot onion**
Methanol-HCl	1.4 ± 0.3[a]	2.5 ± 0.5[b]	1.6 ± 0.5[b]
Online extraction and hydrolysis*	3.2 ± 0.2[b]	3.8 ± 1.3[b]	3.2 ± 1.2[b]
pH 5.5, 5% ethanol, 1 mL/min, 240 min	4.6 ± 0.1[b]	-	-
pH 5.5, 5% ethanol, 3 mL/min, 90 min	7.8 ± 1.1[b]	-	-

*On-line method, pH 5.5, 5% ethanol, 0.75 mL/min and 120 min. [a] n=6, [b] n=3

Paper V

Comparison of the active site mutants of *Thermotoga neapolitana* β-glucosidase 1A for hydrolysis of *p*NPGlc and quercetin-3 glucosides

Samiullah Khan[a], Tahir Mahmood[a], Tejas S. Kulkarni[a], Anders Sundin[b], Sofia Lindahl[b], Charlotta Turner[b], Derek T.Logan[c] and Eva Nordberg Karlsson[a*]

[a]*Biotechnology, Dept. of Chemistry, Lund University, P.O. Box 124, SE-221 00 Lund, Sweden.*

[b]*Centre for Analysis and Synthesis, Dept. of Chemistry, Lund University, P.O. Box 124, SE-221 00 Lund, Sweden*

[c]*Biochemistry and Structural Biology, Dept. Of Chemistry, Lund University, P.O. Box 124, SE-221 00 Lund, Sweden*

[*]*To whom correspondence should be addressed*

E-mail*: sami.khan@biotek.lu.se; tahir.mu@hotmail.com;*

Tejas.kulkarni@biochemistry.lu.se; anders.sundin@organic.lu.se;

sofia.lindahl@organic.lu.se; charlotta.turner@organic.lu.se;

Derek.Logan@biochemistry.lu.se

**eva.nordberg_karlsson@biotek.lu.se (corresponding author)*

Keywords: biocatalysis / glycoside hydrolase family 1 / flavonoid glucosides / antioxidant / Mutagenesis / *Thermotoga neapolitana*

Abstract

The enzyme *Tn*Bgl1A from *Thermotoga neapolitana* catalyse hydrolysis of O-linked terminal β-glycosidic bonds at the non-reducing end of glycosides and oligosaccharides. Previous work has shown the enzyme to catalyse hydrolysis of flavonoid glucosides but that glucosylation at the 3-position was less efficiently hydrolysed. Hydrolysis of the flavonoid quercetin-3-glucoside could however be improved by only a single residue change in the aglycone binding site, and to further explore the effect of selected residues on activity, a set of single amino acid changes were introduced close to the aglycone binding region of the active site at positions N220(S/F), N221(S/F), F224(I), F310(L/E) and W322(A). The effect on activity was monitored using the substrates *para*-nitrophenyl-β-D-glucopyranoside (*p*NPGlc) and the flavonoid quercetin-3-glucoside (Q3). All enzyme variants were cloned and overexpressed in *Escherichia coli,* and purified by immobilized metal ion affinity chromatography to a purity of approximately 90%. Flavonoid-glucoside hydrolysis was monitored by developing a HPLC-assay for this purpose. The mutations N220S, N221S and F224I led to a small increase in K_M compared to the wild-type using *p*NPGlc, while for Q3 the N220S and N221S mutations decreased K_M. The turnover of both *p*NPGlc and Q3 was increased most by the N221S mutation. Q3 turnover was also increased by mutating N220S, while N221F and W322A led to a dramatic reduction of *p*NPGlc turnover. The introduced mutations did not significantly affect the thermal stability of the enzyme, which kept an apparent unfolding temperature of 101°C.

Introduction

Glycoside hydrolases (GH) are a widespread group of enzymes hydrolysing the glycosidic bonds between two or more carbohydrates, or between a carbohydrate and a non-carbohydrate moiety [1]. Among the GHs, β-glucosidases catalyse hydrolysis of O-linked terminal β-glycosidic bonds at the non-reducing end of glycosides and oligosaccharides [2, 3]. The GH family 1 contains enzymes found in all domains of life, Which all share the same overall $(\beta/\alpha)_8$ fold, but which have different substrate specificities [2, 3]. In living organisms they play a role in many biological processes, such as the degradation of cellulosic biomass in bacteria and fungi [4, 5], in defence mechanisms in plants [2, 3], in cyanogenesis [6], and in lactase-phlorizin hydrolysis in mammals [2].

Moreover, these enzymes can be used as specific catalysts in many industrial applications, and in this field we have utilized the thermostable β-glucosidase from *Thermotoga neapolitana* (*Tn*Bgl1A) as a catalyst in extraction processes combined with pressurized hot water (and compared to previously used methanol/HCl-methods) to develop sustainable methods to obtain a uniform aglycone form of the antioxidising flavonoid quercetin from onion waste [7-9].

The use of thermostable enzymes in various biotechnological processes is constantly being explored [10], and general advantages include higher reaction rate, reduced risk of contamination, increased substrate and product solubility, stability of enzymes against harsh conditions (such as denaturing agents) and proteolytic enzymes, and reduced cost of extra cooling [10, 11]. Thus far, there are however only a few reports on β-glucosidases from thermophilic bacteria with preferential activity toward flavonoid glucosides [8, 10, 12, 13], but use of the high temperature extraction process demands a stable enzyme, and adds the advantage of better substrate solubility.

Flavonoids are a large class of polyphenolic compounds with antioxidant properties that are found in plants, e.g. anthocyanin in strawberry, raspberry and red plum, flavonones in orange and grape fruits, flavonols in onion, leek, spinach and green cabbage and hydroxyl cinnamate in apple, tomato, pear and peach [14, 15]. Their consumption in human nutrition is associated with some health-enhancing properties, such as prevention of cancer, lowering the risk of vascular diseases, improvement in bone health and reduction in development of degenerative diseases [16-19]. Glycosylation is common in flavonoids and, dependent on their origin, various sugars *e.g.* glucose, xylose, rhamnose and galactose are attached. The β-glycosidic form is most common in nature, and flavonols are mainly found as 3-O-glycosides

but they may also be glycosylated at the 7- and 4′-positions [20]. In yellow onion the dominant flavonoid is quercetin and the predominant glycosylations are glucose bound at the 3 and 4′-positions in the molecule [quercetin-3-glucoside (Q3), quercetin-4′-glucoside (Q4′) and quercetin-3,4′-diglucoside (Q3,4′)] [8]. It is reported that the aglycones are superior to glycosylated forms for some bioactivities, e.g. greater anticarcinogenic effect and higher antioxidative power [21]. Therefore there is interest in converting glycosylated flavonoids into their respective aglycones.

Some GH1 enzymes hydrolyse flavonoid glycosides, dependent on the aglycone moiety, type of sugar and linkage [8, 9, 22]. *Tn*Bgl1A is a good biocatalyst for conversion of quercetin-4′-glucoside to quercetin in yellow onion, but is less efficient in hydrolysis of the glucosylation at the 3-position [8]. A small set of mutations in the active site region showed that it was possible to improve the activity also for Q3-hydrolysis [9], and the double mutant N221S/P342L increased the catalytic efficiency towards Q3 as well as for Q4′ compared to the wild type enzyme [7]. This raised our interest in identifying additional positions of importance for specificity in the active site, and the purpose of this study was to introduce a set of single amino acid changes in the active site region and to analyse the effects on activity using the well-known model substrate *p*NPGlc in parallel with activity analysis on Q3 in order to identify positions affecting one or both substrates. The crystal structure of the enzyme has also recently been determined (Kulkarni T, Logan D, Nordberg Karlsson E, unpublished data), and docking of the substrates was used to monitor putative changes in interactions due to the introduced mutations. The effect of the mutations on thermal stability was also analysed.

Materials and Methods

Chemicals

Methanol, formic acid, citric acid monohydrate, and disodium hydrogen phosphate were purchased from Merck (Darmstadt, Germany). Para-nitrophenol (pNP), quercetin dihydrate, quercetin 3-*β*-D-glucoside (Q3), and *p*NPGlc were purchased from Sigma-Aldrich (Steinheim, Germany). Quercetin 4′-O- *β*-glucopyranoside (Q4′) and quercetin 3, 4′-di-O- *β*-glucopyranoside (Q3,4′) were purchased from Polyphenols Laboratories AB (Sandnes, Norway). Kaempferol-3-O-glucoside (K3), Kaempferol-7-O-glucoside (K7), Isorhamnetin-3-O-glucoside (I3), daidzein, daidzein-7-O-glucoside and morin were purchased from Extrasynthese (Lyon, France). Genistein and genistein-7-O-glucoside were purchased from

Extrasynthese (Lyon, France) and from Sigma-Aldrich (Steinheim, Germany). Ultrapure water (MilliQ) was used at all times.

Mutagenesis

The gene encoding wt *Tn*Bgl1A was cloned in pET22b(+), under control of the T7lac-promoter, as described by Turner et al. [8]. Mutagenesis was performed in order to introduce the following eight designed changes: N220S/F, N221S/F, F224I, F310E/L and W322A, respectively. The QuikChange II site-directed mutagenesis kit (Stratagene) was used to introduce the designed mutations, and the variants were PCR amplified using mutagenic primers (Table 1) according to the instruction of the manufacturer, using the wild-type gene in pET22b+ as template. The mutagenic primers were synthesized at Eurofins (Germany).

Expression and purification

The mutant and wild-type enzymes were produced in 2.5 L batch cultivations at 37°C, pH 7, using the defined medium mAT [23] with 100 µg/mL ampicillin and a dissolved oxygen tension (DOT) above 40 %. Expression was induced at OD_{620nm} = 3, by the addition of 1 mM isopropyl-beta-D-thiogalactopyranoside (IPTG), and continued for 3 h. Production levels were analysed by SDS-PAGE in samples (1 mL) withdrawn hourly after induction. The cells in the samples were pelleted, resuspended in 300 µL 50 mM citrate-phosphate buffer, pH 5.6, ultrasonicated for 2×90 s with a UP400S equipped with a 3 mm titanium probe (Dr. Hielscher, Stahnsdorf, Germany) using a sound intensity of 60 % and a cycle of 0.5 and thereafter centrifuged for 15 min at 13 000 × *g* to separate soluble proteins from insoluble proteins and cell debris.

At harvest, the cells were separated from the cultivation medium by centrifugation (10000 × *g*, 10 min, 4°C), and dissolved in binding buffer (20 mM imidazole, 20 mM Tris-HCl, 0.75 M NaCl, pH 7.5). The ice-chilled cell suspension was lysed by sonication for 5×3 min using a 14 mm titanium probe, with a sound intensity of 60 % and a cycle of 0.5 (UP400S, Dr. Hielscher), centrifuged (30 min, 39000 × *g*, 4°C), heat treated (70°C, 30 min) and again centrifuged (30 min, 39000 × *g*, 4°C). The supernatant was passed through a 0.45 µm Minisart high-flow filter (Sartorius, Göttingen, Germany) and purified on an ÄKTA prime system (Amersham Biosciences, Uppsala, Sweden) by immobilized metal ion affinity chromatography using copper as a ligand as described elsewhere [24]. The fractions containing the purified protein were pooled and dialyzed against 20 mM citrate phosphate

buffer, pH 5.6, overnight using a Spectra/Por dialysis membrane with a 3500 Da molecular weight cut-off ((MWCO)(Spectrum laboratories, Rancho Dominguez, CA, USA). The dialysed protein fractions were stored at 4°C until use.

Protein analysis

The purity of each enzyme variant was analysed by SDS-PAGE according to Laemmli [25]. Expression levels were also analysed by SDS-PAGE after separating insoluble and soluble proteins (described above).

Total protein concentration was estimated at 562 nm by the BCA method according to the manufacturer's instructions (Sigma, Steinheim, Germany) using bovine serum albumin (Sigma-Aldrich) as standard.

Differential scanning calorimetry (DSC)

DSC analysis was made on a MicroCal differential scanning calorimeter (VP-DSC, MicroCal, Northampton, MA, USA) with the cell volume of 0.5072 mL. The samples (in 20 mM citrate phosphate buffer, pH 5.6) were concentrated to 1 mg/mL using Vivaspin (Sartorius AG, Goettingen, Germany) centrifuge tubes with a MWCO of 30,000 Da and were degassed before the scans. The samples were scanned at a rate of 1°C/min in the temperature range of 25-110°C.

Enzyme activity on *p*NPGlc

Enzyme activity, and kinetic parameters (K_M and k_{cat}) were determined at 80°C, pH 5.6 using *p*NPGlc as substrate in 20 mM citrate phosphate buffer, on a Shimadzu UV-1650 Visible spectrophotometer (Shimadzu, Duisburg, Germany). A volume of 980 µL of *p*NPGlc (in a concentration range from 0.09125 to 1 mM) was preheated for 10 min, where after 20 µL of the enzyme solution (12 µg mL^{-1}, 4.56 pmol) was added. Absorbance at 405 nm was measured and plotted by the Shimadzu UV probe 2.01 software as a function of time during 1 min. The extinction coefficient of *p*NP (*para*-nitrophenol) under these experimental conditions was determined as $\varepsilon_{80°C,\ 405nm} = 2.4639 \times 10^3$ mL $mmol^{-1}cm^{-1}$. The kinetic parameters were determined by applying the Wilkinson non-linear regression method using Enzpack (Biosoft, UK).

Activity screening on flavonoid glycosides

Enzyme activities of *wt* and all mutants of *Tn*Bgl1A were screened using fifteen different flavonoid glycosides at 80°C. 150 nmol samples of substrate and 100 nm of Morin (internal standard) dissolved in methanol (in triplicate) were evaporated and 1.5 mL of 20 mM citrate phosphate buffer, pH 5.6, were added. The vials were heated at 80°C for 15 min. A 100 µL fraction was collected and added to 900 µL of mobile phase composed of methanol/water (45:55) and 0.13 M formic acid. The reaction was started by adding 20 pmol of enzyme and 10 min after addition of enzyme, 100 µL fractions were collected and added to 900 µL of mobile phase. Samples were analyzed by HPLC with UV detection (HPLC-UV), as described below.

Enzyme activity of selected mutants on flavonoid 3-glucosides

Enzyme activities of *wt* and the mutants of *Tn*Bgl1A showing the highest activity in the screening above were further screened for three different flavonoid-glucosides glucosylated at position 3 (I3, K3 and Q3). The enzyme assay was done as above, except that the reaction was started by adding 5 pmol of enzyme instead of 20 pmol and samples were taken at 0 and 3 min.

Enzyme kinetics on Q3

To determine the kinetics of the hydrolysis reactions, 50–250 nmol samples of Q3 dissolved in methanol were evaporated and re-dissolved as described above. Fractions were collected at 0, 1, 3, and 5 min after addition of enzyme. The enzyme amount added to Q3 was 5 pmol. Samples were analysed by the HPLC-UV-method below. The enzyme activity using Q3 is expressed in Units (U). One U corresponds to the amount of enzyme required to release 1 µmol quercetin per minute under the described conditions. Obtained data was analysed using Enzpack (Biosoft, UK), using a time-range allowing fit to the Michaelis–Menten curve. Kinetic constants were calculated using the Wilkinson non-linear regression method, supplied in the program.

HPLC analysis

HPLC-UV analysis was performed using the chromatographic system UltiMate 3000 from Dionex (Germering, Germany). An Agilent Zorbax SB-C18 pre-column (2.1x12.5 mm,µm) and column (100×2.1 mm, 3.5 µm) was used for isocratic separation with a methanol:water

(45:55) and 0.13 M formic acid mobile phase at a flow rate of 0.15 mL min^{-1}. The injection volume was 10 μL and detection was accomplished at 350 nm. Quantification of aglycone and glycosides was performed using a five-point calibration curve of standards at concentrations between 0.5 and 30 μM. Each vial taken to analysis had a total volume of 1mL.

Docking simulations

Molecular modeling was performed with the OPLS-2005 forcefield and the GB/SA solvation method for water, using Schrödinger 2010 software suite [26]. Docking simulations of *p*NPGlc and Q3 were carried out with a model based on x-ray diffraction data of the of apo *Tn*Bgl1A (Kulkarni, T, et al, unpublished manuscript). The Protein Preparation Wizard was used to add hydrogen atoms, assign charges, optimize hydrogen bond networks, determine the rotamer of terminal amides, the tautomeric form of histidine, and to analyze the quality of the crystal structure.

The crystal structure of *Tn*Bgl1A was aligned with a previously constructed homology model of *Tn*Bgl1A with respect to the secondary structure [9]. The RMSD (Root-mean-square deviation) (over all α-carbons) with the template was 0.5 Å.

β-D-cellotetraose was transferred from its docked position in the homology model to the crystal structure of *Tn*Bgl1A. The resulting complex was energy minimized to obtain a model of β-D-cellotetraose docked in the active site of the crystal structure of *Tn*Bgl1A. The final position of the cellotetraose was found to be in agreement with the proposed mechanism of GH1 β-glucosidases.

For each of the ligands, *p*NPGlc and Q3, a conformational search was performed, and low energy conformers were placed in the active site of of the *Tn*Bgl1A such that the β-D-glucose part of the ligand was superimposed on the β-D-glucose at the non-reducing end of the docked β-D-cellotetraose.The resulting *p*NPGlc and Q3 complexes were energy minimized. Figures of the docking simulations were prepared by using PyMol [27].

Mutant protein structures were constructed by changing the respective amino acids in the model of the energy minimized native protein complexes and the resulting complexes were further subjected to energy minimizations.

Results and Discussion

Mutagenesis and expression

Several enzymes of known structure are available in the GH1 family, and some of this information was utilized to select residues for mutagenesis at positions close to sugar binding site +1/+2 (or aglycone binding site), that could putatively influence aglycone specificity. Superimposition of the crystal based model of *Tn*Bgl1A [9] with the known structures of other enzymes, (thermostable glucosidase from *Pyrococcus horikoshii* [28], glucosidase from *Paenibacillus polymyxa* [29] and human cytosolic beta-glucosidase (hCBG) [30] was combined with results from a mutational study of hCBG [22, 30]. By this approach five residues were chosen for mutagenesis (N220, N221, F224, F310, and W322). Residue N220 was chosen based on results of mutagenesis published for hCBG [30], in which the corresponding position is occupied by a phenylalanine (F225) proposed to be crucial for aglycone binding and recognition. Mutagenesis of this residue to serine resulted in almost complete loss of activity in hCBG. Based on these observations, the corresponding residue (N220) in *Tn*Bgl1A was mutated to both F and S to monitor the effects of these changes on activity. The neighbouring residue N221 was mutated to S in a previous study, [9] and the change N221S showed increased activity on both Q3 and Q4′, which was suggested to occur via a loss in backbone carbonyl interactions that instead resulted in an increased flexibility of the parallel β-sheets. In this work we have included this variant and also introduced an aromatic residue (F) at this position to monitor the effects on the activity. The changes in residues F224, F310 and W322 (F224I, F310L and W322A) were all made with the intention of replacing the existing residue with a smaller hydrophobic residue. Isoleucine at a position corresponding to 224 in *Tn*Bgl1A can be found in a membrane bound β-glucosidase (I212) from the hyperthermophile *Pyrococcus horikoshii* [28]. Likewise, leucine at position 310 (in loop C of the active site entrance) is found in the *P.horikoshii* enzyme, while in the β-glucosidase from *Paenibacillus polymyxa* a glutamic acid is found at the corresponding position (E316) [29], and this change was also introduced. Mutation of W322 into alanine was simply an attempt to make more space in the aglycone binding part of the active site, close to the +1 sugar binding sub-site.

The *wt* and all mutants were produced in *Escherichia coli* Bl21 (DE3). Analysis of the produced protein by SDS-PAGE showed all enzymes to have very similar production patterns. In all cases the soluble fraction was significantly less than half of the total produced

protein (data not shown). Nevertheless, soluble TnBgl1A was purified using a two-step protocol, including a heat treatment (70°C, 30 min) followed by immobilised metal ion affinity chromatography (IMAC) utilizing the C-terminal His-tag, which yielded a purity at or above 90% in all cases (Figure 1).

Thermostability

Thermostability of the *Tn*Bgl1A *wt* and the enzyme variants mutated at positions 220, 221 and 224 was evaluated by differential scanning calorimetry (DSC). A single transition peak was observed in all cases, which resulted in aggregation during unfolding, and was evident in the form of the precipitate. The denaturation was confirmed to be irreversible in all cases when the scan was repeated for the same sample. The apparent unfolding temperatures (T_m) for the mutants (Table 2) were close to the T_m observed for the *wt* enzyme, showing that the single amino acid changes did not result in any major changes of the thermal stability.

***p*NPGlc hydrolysis and putative substrate enzyme interactions**

The kinetic parameters for the hydrolysis reaction were monitored using the model substrate *p*NPGlc at pH 5.6 and 80°C (Table 3). The mutations N221F and W322A resulted in a significant reduction in the turnover number (k_{cat}) while their K_M values were almost the same as for the *wt* (Table 3). The side-chain of the residue at position 221 makes weak Van der Waals interaction with N291 [9]. Introduction of the bulkier phenylalanine instead of asparagine results in the small local changes in the range of 0.4-0.7Å. The significance of these small local changes is unclear but the indirect effects might have increased the activation energy of the reaction leading to the drastically dropping down of the turnover as compared to the reaction with the wild type enzyme.

Tryptophan-322 is an important amino acid in the aglycone binding site, which makes strong van der Waals interactions with the aglycone part of the natural substrates (Fig. 2). The mutation W322A causes the loss of these van der Waals interactions and creates an extra space in the local area, which may allow the aglycone part of the substrate to find alternate binding modes. This is evident from the Figure 2 where the aglycone part of Q3 makes strong van der Waals interactions with W322 and a major decrease in the percent hydrolysis of Q3 is also observed with the W322A variant (Table 5) due to the loss of these interactions. In case of *p*NPGlc, the observed results are also well in agreement with the above argument as K_M is unaffected, but the turnover of the reaction is drastically decreased suggesting the loss of

interactions affecting the activation energy or alternatively difficulties in removing the products from the active site.

F310L/E are surface mutations, located at the entrance of the active site, and is positioned in one (loop C) of the four loops (A-D) defined as responsible for the overall shape of the active site entrance [30]. F310E introduces negatively charged glutamic acid instead of the hydrophobic character of leucine and phenylalanine. However, neither change at this position, resulted in any major changes in kinetic parameters, and the residue does not directly interact with the substrate. The turnover number for mutant F310E was close to *wt* while F310L was lower than the *wt* but the small decrease in K_M counteracted the decrease in turnover and the overall catalytic efficiency (k_{cat}/K_M) was slightly increased for both mutants compared to the *wt.*

The effect of F224I on the lowering of the catalytic efficiency is not clearly evident from the docking studies. The F224I variant has higher turnover number, but as the increased K_M counteracted the increase in turnover, a decrease in catalytic efficiency was observed. In contrast, the three mutants N220S, N220F and N221S showed marked elevation in the catalytic activity of the enzyme as compared to the *wt*. In these three variants (N220S/F and N221S) there is a slight increase in K_M which is accompanied by an increase in the turnover number, resulting in increased catalytic efficiency compared to the *wt*. Previous work showed that increased K_M of the mutant N221S/P342L (0.66 mM) for *p*NPGlc [9] also resulted in decreased K_M when it was tested in the deglucosylation reaction of the quercetin glucosides (Q3 and Q4´) [7]. Therefore these variations in K_M and (k_{cat}) raised an interest for trying these variants in deglycosylation reactions of flavonoids and isoflavonoids with special focus in Q3, as glucose at the 3-position generally seems more difficult to hydrolyse using GH1 enzymes.

Development of a HPLC method for flavonoid analysis

A HPLC method was initially developed for the determination of enzyme activity with regard to selected flavonoids. The necessity of using HPLC rather than a more simple spectrophotometric assay depends on the fact that both substrates (flavanoid glycosides) and expected products (flavonoid aglycones) absorb light at the same wavelengths. Therefore, it is necessary to include a separation step prior to detection. In order to shorten the total assessment time for each investigated set of enzyme, substrate and hydrolysis condition, especially since a large number of different flavoinoids were investigated in this study, the separation step should be as short as possible. A reversed phase HPLC system was used, with

a mixture of methanol and water as mobile phase containing formic acid to make sure that the flavonoids are uncharged during the separation (Table 4). A wavelength of 350 nm was used for for detection of all substrates and product. The Figure 3 below shows a representative chromatogram showing the partly hydrolysed substrate Q3 as well as its product, Q and morin used as internal standard. Table 4 shows the substrates and correpsonding products investigated in the study, and their retention time using the developed HPLC method.

Screening of flavonoid and isoflavonoid glucoside hydrolysis

The wt *Tn*Bgl1A was previously shown to hydrolyse Q3, but slower than hydrolysis of Q4´ [7-9]. Other GH1 β-glucosidases shown to have activity on flavonoid substrates seem to follow the same trend, frequently completely lacking the ability to hydrolyse glucosylation at the 3-position, e.g. hCBG [22] or β-glucosidase from pig liver [31].The pig enzyme failed to hydrolyse Q3, Q3,4´ and K3 but there was no problem in the hydrolysis of Q4´and 7-glucoside of genistein and daidzein (Fig. 4). This motivated a screening of the hydrolysis of different flavonoid/isoflavonoid glucosides (glucosylated at position 3, 4´and 7) into aglycones using *Tn*Bgl1A and its mutants as catalysts (Table 3). Fixed concentrations of both enzyme and substrate were used, and the hydrolysis data was analysed by HPLC (Table 5). It is clear from the table that the sugars attached at the 7 and 4´ positions were easy to hydrolyse for *wt* and all the mutants and after 10 min all the substrates were converted into product except for mutant N221F, which hydrolysed only 47% of Q4´and 72% of K7. Introduction of phenylalanine at position 220 was also generally slowing down hydrolysis on several substrates (Table 5).

Complete hydrolysis of the flavonoid-3-glucosides was only achieved for the double mutant N221S/P342L. For all other variants, some of the 3-glucoside remained after completion of the 10 min incubation period. It was however only the N220F, N221F and W322A that converted significantly lower amounts of the substrate than the wt enzyme. No hydrolysis at all, was detectable under standard HPLC condition when the glycosylation involved other sugar molecules than glucose.

Screening of flavonoid-3-glucoside hydrolysis

The different behaviour of wt *Tn*Bgl1A and mutated variants in hydrolysis of flavonoid-3-glucosides led us to expand the study for 5 variants, showing equal or improved conversion

as compared to the wt enzyme in the above screening. The percent hydrolysis was measured from HPLC in the same way as above except that the hydrolysis time for the enzyme variants was reduced to 3 min, and the enzyme concentration reduced from 20 to 5 pmol. The data showed that the selectivity of mutants F310E and F310L were same as *wt* but the efficiency of N220S, N221S, N221S/P342L increased for hydrolysis of the 3-glucosidic substrates compared to *wt* (Table 6). A weak trend was also observed, indicating that the I3 was converted faster than K3 which in turn was converted faster than Q3. The same trend was visible in *wt* as well as mutated variants. No difference was clearly visible between the *wt* and the four mutated variants, except that N220S, N221S and N221S/P342L have an activity slightly higher than *wt*. The variant F224I was also converting more substrate than the *wt*, but less than the three variants mutated at position N220 and N221 to S.

Q3 hydrolysis and putative substrate enzyme interactions

The kinetic parameters were determined for the four mutants (N220S, N221S, N221S/P342L and F224I) whose action resulted in the highest conversion of Q3 into quercetin at 80°C and pH 5.6 (Table 3) compared to *wt*. The kinetic data obtained here resulted in a higher V_{max} than previously reported [7]. This is because the previous study was made at 90°C and a later analysis of the CD-spectrum has shown that the alpha-helical parts of the enzyme give a reduced signal at this temperature, indicating partial unfolding (data not shown). This is likely the cause of the decrease in V_{max}, despite a long half-life at this temperature [7], likely caused by reversible denaturation at conditions where complete unfolding is avoided. From the docking studies of the interactions of Q3 with the *wt* and engineered *Tn*Bgl 1A, it is not clear to explain, why the hydrolysis of Q3 is favored and/or unfavored in case of the specific mutations at the different positions as compared to the *wt Tn*Bgl1A (Fig. 2), and direct interactions were only predicted for the N221S mutated enzyme. The largest improvement has been seen for the double mutant N221/P342L due to a decrease in K_M for Q3. This mutant has been reported before and a second spontaneous mutation P342L which is located on the surface of protein opposite to active site must play some indirect function e.g. influencing the conformation, as the effect is lower using the single mutant N221S. P342 does not participate directly in the substrate interaction, but the double mutant shows a minor decrease in K_M and slight increase in k_{cat} increasing its catalytic efficiency compared to the single mutant N221S. The K_M and k_{cat} values for mutant N220S and N221S were in the same range but due to the combined decrease in K_M and increased k_{cat} both have pronounced effect

on the catalytic efficiency. The kinetic parameters for the F224I variant was in Q3 hydrolysis similar to *wt* and no significant change was observed on this substrate.

Conclusions

In this work single amino acid changes were introduced in the aglycone binding part of the catalytic site, and the effect on *p*NPGlc and Q3 hydrolysis was monitored. Changes at position N220 and N221 had large effects on the hydrolysis of both substrates, showing these positions to be of importance for substrate accommodation. In addition the W332 side chain was important, and its removal drastically reduced turnover of *p*NPGlc. With these results another step is taken towards molecular understanding of differences in substrate specificity in glycoside hydrolase family 1.

Acknowledgement

The authors wish to thank Formas (Sustain-X-Enz 2006-1346, SuReTech 2009-1527) for funding and Higher Education Commission (HEC) of Pakistan.

References

1. Cantarel BL, Coutinho PM, Rancurel C, Bernard T, Lombard V, Henrissat B: **The Carbohydrate-Active EnZymes database (CAZy): an expert resource for Glycogenomics**. *Nucleic Acids Res* 2009, **37**:D233-D238.
2. Bhatia Y, Mishra S, Bisaria VS: **Microbial beta-glucosidases: Cloning, properties, and applications**. *Crit Rev Biotechnol* 2002, **22**(4):375-407.
3. Cairns JRK, Esen A: **β-Glucosidases**. *Cell Mol Life Sci* 2010, **67**(20):3389-3405.
4. Bisaria V, Mishra S: **Regulatory aspects of cellulase biosynthesis and secretion.** *Crit Rev Biotechnol* 1989, **9**(2):61-103.
5. Ghosh P, Pamment NB, Martin WRB: **Simultaneous saccharification and fermentation of cellulose - effect of beta-d-glucosidase activity and ethanol inhibition of cellulases**. *Enzyme Microb Technol* 1982, **4**(6):425-430.
6. Lei V, Amoa-Awua WKA, Brimer L: **Degradation of cyanogenic glycosides by Lactobacillus plantarum strains from spontaneous cassava fermentation and other microorganisms**. *Int J Food Microbiol* 1999, **53**(2-3):169-184.
7. Lindahl S, Ekman A, Khan S, Wennerberg C, Borjesson P, Sjoberg PJR, Karlsson EN, Turner C: **Exploring the possibility of using a thermostable mutant of beta-glucosidase for rapid hydrolysis of quercetin glucosides in hot water**. *Green Chem* 2010, **12**(1):159-168.
8. Turner C, Turner P, Jacobson G, Almgren K, Waldeback M, Sjoberg P, Karlsson EN, Markides KE: **Subcritical water extraction and beta-glucosidase-catalyzed hydrolysis of quercetin glycosides in onion waste**. *Green Chem* 2006, **8**(11):949-959.
9. Khan S, Pozzo T, Megyeri M, Lindahl S, Sundin A, Turner C, Karlsson EN: **Aglycone specificity of Thermotoga neapolitana beta-glucosidase 1A modified by**

mutagenesis, leading to increased catalytic efficiency in quercetin-3-glucoside hydrolysis. *BMC Biochem* 2011, **12**:11.
10. Turner P, Mamo G, Karlsson E: **Potential and utilization of thermophiles and thermostable enzymes in biorefining**. *Microbial Cell Factories* 2007, **6**(1):9-9.
11. Vieille C, Zeikus GJ: **Hyperthermophilic enzymes: Sources, uses, and molecular mechanisms for thermostability**. *Microbiol Mol Biol Rev* 2001, **65**(1):1-+.
12. Kuo LC, Cheng WY, Wu RY, Huang CJ, Lee KT: **Hydrolysis of black soybean isoflavone glycosides by Bacillus subtilis natto**. *App Microbiol Biotechnol* 2006, **73**(2):314-320.
13. Song XF, Xue YM, Wang QL, Wu XX: **Comparison of Three Thermostable beta-Glucosidases for Application in the Hydrolysis of Soybean Isoflavone Glycosides**. *J Agric Food Chem* 2011, **59**(5):1954-1961.
14. Boudet AM: **Evolution and current status of research in phenolic compounds**. *Phytochemistry* 2007, **68**(22-24):2722-2735.
15. Proteggente AR, Pannala AS, Paganga G, Van Buren L, Wagner E, Wiseman S, Van De Put F, Dacombe C, Rice-Evans CA: **The antioxidant activity of regularly consumed fruit and vegetables reflects their phenolic and vitamin C composition**. *Free Radic Res* 2002, **36**(2):217-233.
16. Arts ICW, Hollman PCH: **Polyphenols and disease risk in epidemiologic studies**. *Am J Clin Nutr* 2005, **81**(1):317S-325S.
17. Vita JA: **Polyphenols and cardiovascular disease: effects on endothelial and platelet function**. *American Journal of Clinical Nutrition* 2005, **81**(1):292S-297S.
18. Rice Evans CA, Miller NJ, Bolwell PG, Bramley PM, Pridham JB: **The relative antioxidant activities of plant-derived polyphenolic flavonoids.** *Free Radic Res* 1995, **22**(4):375-383.
19. McGhie TK, Walton MC: **The bioavailability and absorption of anthocyanins: Towards a better understanding**. *Mol Nutr Food Res* 2007, **51**(6):702-713.
20. Fossen T, Pedersen AT, Andersen OM: **Flavonoids from red onion (Allium cepa)**. *Phytochemistry* 1998, **47**(2):281-285.
21. Nemeth K, Plumb GW, Berrin JG, Juge N, Jacob R, Naim HY, Williamson G, Swallow DM, Kroon PA: **Deglycosylation by small intestinal epithelial cell beta-glucosidases is a critical step in the absorption and metabolism of dietary flavonoid glycosides in humans**. *Eur J Nutr* 2003, **42**(1):29-42.
22. Berrin JG, Czjzek M, Kroon PA, McLauchlan WR, Puigserver A, Williamson G, Juge N: **Substrate (aglycone) specificity of human cytosolic beta-glucosidase**. *Biochem J* 2003, **373**:41-48.
23. Ramchuran SO, Holst O, Karlsson EN: **Effect of postinduction nutrient feed composition and use of lactose as inducer during production of thermostable xylanase in Escherichia coli glucose-limited fed-batch cultivations**. *J Biosci Bioeng* 2005, **99**(5):477-484.
24. Turner P, Holst O, Karlsson EN: **Optimized expression of soluble cyclomaltodextrinase of thermophilic origin in Escherichia coli by using a soluble fusion-tag and by tuning of inducer concentration**. *Protein Expr Purif* 2005, **39**(1):54-60.
25. Laemmli UK: **Cleavage of Structural Proteins during the Assembly of the Head of Bacteriophage T4**. *Nature* 1970, **227**(5259):680-685.
26. Schrödinger: **LLC**. In. New York, NY; 2009.
27. Schrödinger L: **The PyMOL Molecular Graphics System**. In. Edited by 1.3 V.

28. Akiba T, Nishio M, Matsui I, Harata K: **X-ray structure of a membrane-bound beta-glycosidase from the hyperthermophilic Archaeon Pyrococcus horikoshii**. *Proteins-Structure Function and Bioinformatics* 2004, **57**(2):422-431.
29. Isorna P, Polaina J, Latorre-Garcia L, Canada FJ, Gonzalez B, Sanz-Aparicio J: **Crystal structures of Paenibacillus polymyxa beta-glucosidase B complexes reveal the molecular basis of substrate specificity and give new insights into the catalytic machinery of family I Glycosidases**. *J Mol Biol* 2007, **371**(5):1204-1218.
30. Tribolo S, Berrin JG, Kroon PA, Czjzek M, Juge N: **The Crystal Structure of Human Cytosolic β-Glucosidase Unravels the Substrate Aglycone Specificity of a Family 1 Glycoside Hydrolase**. *J Mol Biol* 2007, **370**(5):964-975.
31. Lambert N, Kroon PA, Faulds CB, Plumb GW, McLauchlan WR, Day AJ, Williamson G: **Purification of cytosolic beta-glucosidase from pig liver and its reactivity towards flavonoid glycosides**. *Biochimica Et Biophysica Acta-Protein Structure and Molecular Enzymology* 1999, **1435**(1-2):110-116.

List of abbreviations

DSC: differential scanning calorimetry

GH1: Glycoside Hydrolase family 1

HPLC: high performance liquid chromatography

K_M: The Michaelis constant

k_{cat}: turnover number

PCR: polymerase chain reaction

*p*NPGlc: para-nitrophenyl-beta-D. glucopyranoside

*p*NP: para-nitrophenol

*Tn*Bgl1A: *Thermotoga neapolitana* beta-glucosidase A from Glycoside hydrolase family 1

Q: quercetin

Q3: quercetin 3-*β*-D-glucoside

Q4′: quercetin-4′-glucoside

Q3,4′: quercetin 3, 4′-di-O- *β*-glucopyranoside

Din: daidzein

D7: daidzein-7-O-glucoside

Gin: ginistein

G7: Ginistein-7-O-glucoside

Gin: ginistein

K: Kaempferol

K7: Kaempferol-7-O-glucoside

K3: Kaempferol-3-O-glucoside

I: Isorhamnetin

I3:Isorhamnetin-3-glucoside

SDS-PAGE: sodium dodecyl sulfate polyacrylamide gel electrophoresis

3D: three dimensional

TABLES

Table 1. Oligonucleotides used for cloning and mutagenesis

Primer	Direction	Mutation	Sequence[a]
1	Forward	N220F	5′- GGAAAGATAGGGATTGTTTTC**TT**CAACGGATACTTCGAACCTGC
2	Reverse	N220F	5′- GCAGGTTCGAAGTATCCGTTG**AA**GAAAACAATCCCTATCTTTCC
3	Forward	N220S	5′- GGAAAGATAGGGATTGTTTTC**TC**CAACGGATACTTCGAACCTGC
4	Reverse	N220S	5′- GCAGGTTCGAAGTATCCGTTG**GA**GAAAACAATCCCTATCTTTCC
5	Forward	N221F	5′ GGAAAGATAGGGATTGTTTTCAAC**TT**CGGATACTTCGAACCTGC
6	Reverse	N221F	5′-GCAGGTTCGAAGTATCCG**AA**GTTGAAAACAATCCCTATCTTTCC
7	Forward	N221S	5′GGAAAGATAGGGATTGTTTTCAAC**TC**CGGATACTTCGAACCTGC
8	Reverse	N221S	5′- GCAGGTTCGAAGTATCCG**GA**GTTGAAAACAATCCCTATCTTTCC
9	Forward	F224I	5′- GGATTGTTTTCAACAACGGATAC**A**TCGAACCTGCAAGTGAG
10	Reverse	F224I	5′- CTCACTTGCAGGTTCGA**T**GTATCCGTTGTTGAAAACAATCC
11	Forward	F310L	5′- GCCAGGGTGTCC**C**TCGTCGAAAGAAACC
12	Reverse	F310L	5′- GGTTTCTTTCGACGA**G**GGACACCCTGGC
13	Forward	F310E	5′- CCGGCCAGGGTGTCC**GAG**GTCGAAAGAAACCTTCCC
14	Reverse	F310E	5′- GGGAAGGTTTCTTTCGAC**CTC**GGACACCCTGGCCGG
15	Forward	W322A	5′- CCAAAACCGCCATGGGAT**TC**GAGATCGTTCCTGAGGG
16	Reverse	W322A	5′- CCCTCAGGAACGATCTC**GA**ATCCCATGGCGGTTTTGG

[a]Induced changes are shown in bold.

Table 2. The melting temperatures of the selected expressed constructs measured by differential scanning calorimetry**.**

Enzyme	**T_m (^{o}C)**	**ΔT_m (^{o}C)**
wild-type*	101.9	
N220S	101.6	-0.3
N221S/P342L*	100.5	-1.4
N221S	100.8	-1.1
F224I	101.1	-0.8
F310L	101.4	-0.8
F3101E	nd	nd
W322A	100.3	-1.6

*Data taken from reference [9]

Table 3. Michaelis-Menten constants. For both *p*NPG and quercetin glucoside hydrolysis measurements were made at pH 5.6, 80°C, (*wt* and mutants).

Enzyme	K_M (mM)	k_{cat} (s^{-1})	k_{cat}/K_M (s^{-1} mM^{-1})
		*p*NPG-hydrolysis	
Wild-type	0. 24 ± 0.04	485 ± 31	2000
N220S	0.31 ± 0.05	757 ± 54	2433
N220F	0.30 ± 0.03	849 ± 40	2857
N221S	0.43 ± 0.11	1170 ± 150	2724
N221F	0.24 ± 0.009	22 ± 0.4	93
F224I	0.64 ± 0.06	803 ± 41	1247
F310E	0.19 ± 0.01	508 ± 10	2714
F310L	0.17 ± 0.01	371 ± 9.1	2144
W322A	0.22 ± 0.01	61 ± 1.2	282
		Q3-hydrolysis	
wt	0.13 ± 0.06	13.6 ± 3.2	101
N220S	0.093 ± 0.01	39.8 ± 2.8	428
N221S	0.092 ± 0.03	40.8±5.7	446
N221S/P342L	0.074 ± 0.01	45.6 ± 3.3	618
F224I	0.15 ± 0.07	14.4 ± 3.9	95

Table 4. Substrate and products used for assessment of enzymatic activity, and their corresponding retention time in HPLC. At UV Visible 350 nm.

Substrate	Retention time (min)	Final Product	Retention time (min)
Quercetin-3-glucoside	4.3	Quercetin	9.8
Quercetin-4´-glucoside	5.1	Quercetin	9.8
Quercetin-3,4´-diglucoside	2.9	Quercetin	9.8
Daidzein-7-O-glucoside	2.7	Daidzein	8.3
Ginistein-7-O-glucoside	3.5	Ginistein	14.4
Kaempferol-7-O-glucoside	4.5	Kaempferol	17.3
Kaempferol-3-O-glucoside	5.8	Kaempferol	17.3
Isorhamnetin-3-glucoside	5.8	Isorhamnetin	19.7

Table 5. Comparison of the percent hydrolysis of *wt* and mutants. C= complete conversion of the substrate at the conditions chosen.

Enzyme	Time	Gin	Din	Q/Q4´	K/K7	Q/Q3,4	Q/Q3	K/K3	I/I3
wild-type	10	C	C	C	C	82±0.5	73±0.8	87±1	93±0.4
N220S	10	C	C	C	C	92±0.5	84±0.4	97±0.2	98±0.1
N220F	10	C	C	C	C	10±2	8±1	35±2	27.±2
N221S	10	C	C	C	C	96±0.4	98±0.7	C	C
N221F	10	C	C	47±1	72±2	15±5	22±3	23±11	53±2
N221S/									
P342L	10	C	C	C	C	C	C	C	C
F2241	10	C	C	C	C	87±0.4	84±0.9	96±0.3	C
F310L	10	C	C	C	C	53±0.5	52±0.6	85±0.2	93±1
F310E	10	C	C	C	C	74±0.8	73±2	89±2	91±0.3
W322A	10	C	C	C	C	17±2	12±0.5	35±2	18±2

Table 6. Comparison of the percent hydrolysis of K3, I3 and Q3. nd = not determined

Enzyme	Time	K	Iso	Q
Wild-type	3	13±3	14±0.5	11±3
N220S	3	36±0.6	36±2	23±1
N221S	3	32±0.8	45±1	27±1
N221S/P342L	3	39±0.6	52±1	30±0.6
F2241	3	19±2	21±3	10±1
F310L	3	14±5	nd	8±2
F310E	3	14±1	15±4	10±1

Figures:

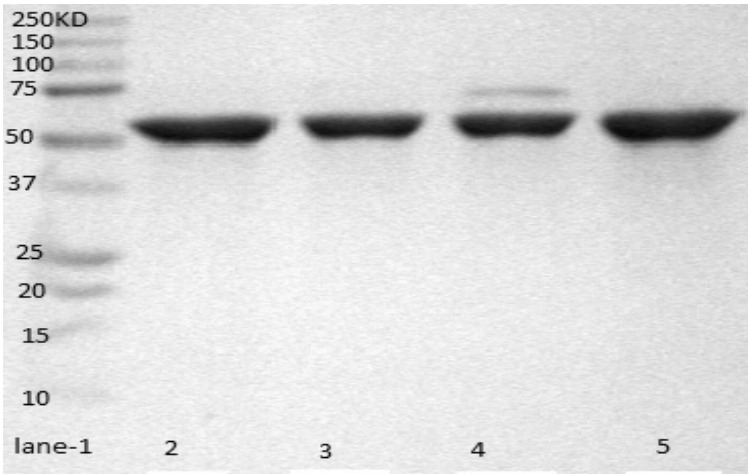


Figure 1. SDS-PAGE analysis of purified enzymes after immobilised metal ion affinity chromatography. lane 1 molecular weight marker, lane 2, *Tn*Bgl1A-F224I, lane 3, *Tn*Bgl1A-N221S, lane 4, *Tn*Bgl1A-N220S, lane 5 *wtTn*Bgl1A

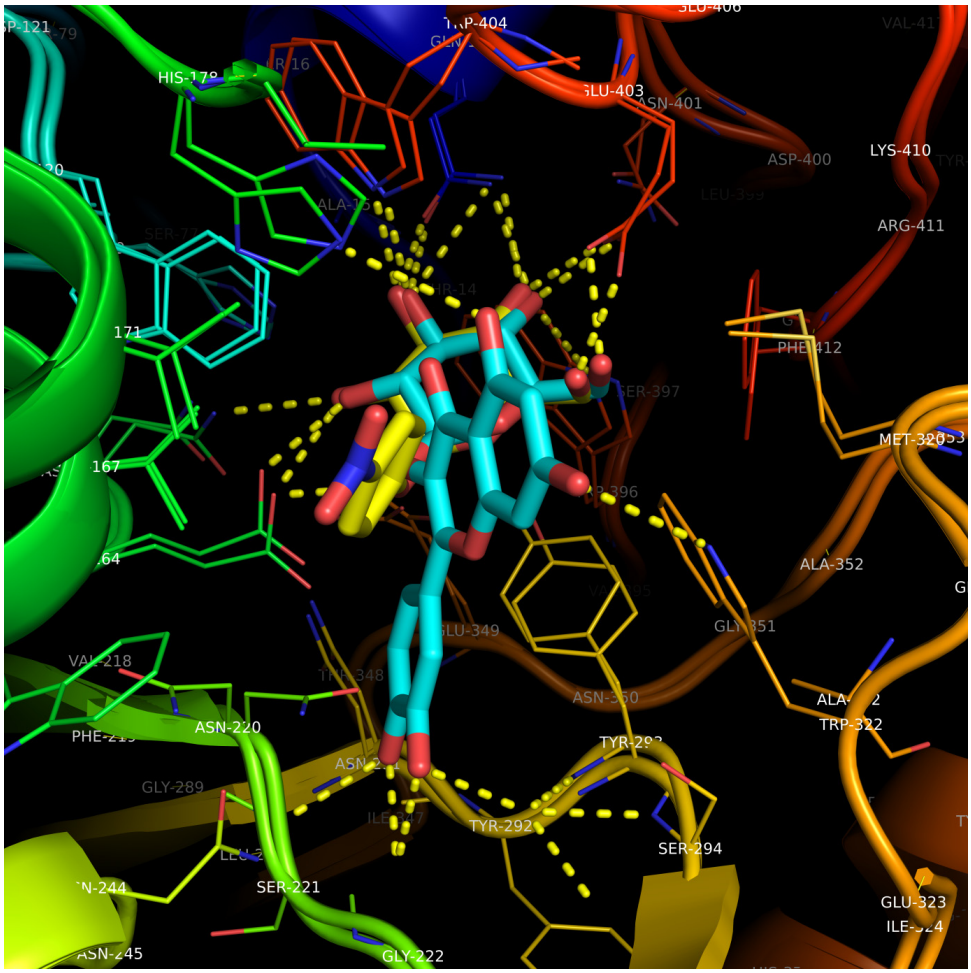


Figure 2. Ligand binding sites showing *p*NPGlc and Q3 bound with the W322A and the wild type variant respectively. The ligand represented in Yellow color (stick representation) is *p*NPGlc and the ligand represented in aquamarine color (also in stick representation) is Q3. Yellow dotted lines represent the hydrogen bond interactions between *p*NPGlc and Q3 and the respective active site amino acid residues.

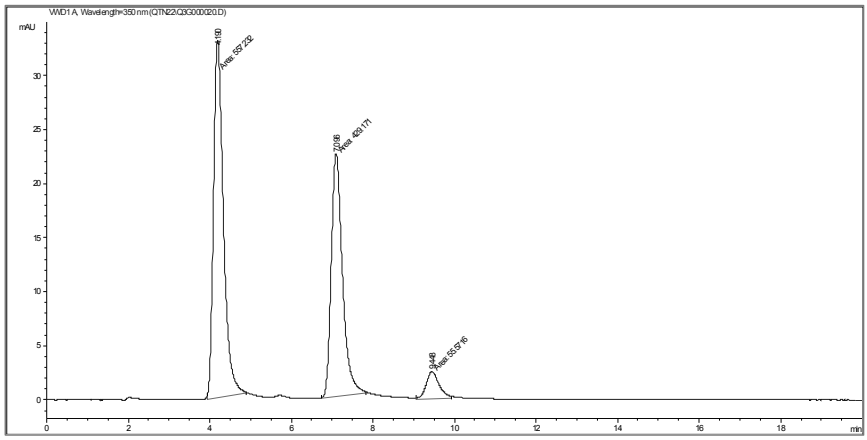


Figure 3. Representative chromatogram of HPLC-UV analysis of Q3 showing Q3,Morin and Q (from left to right).

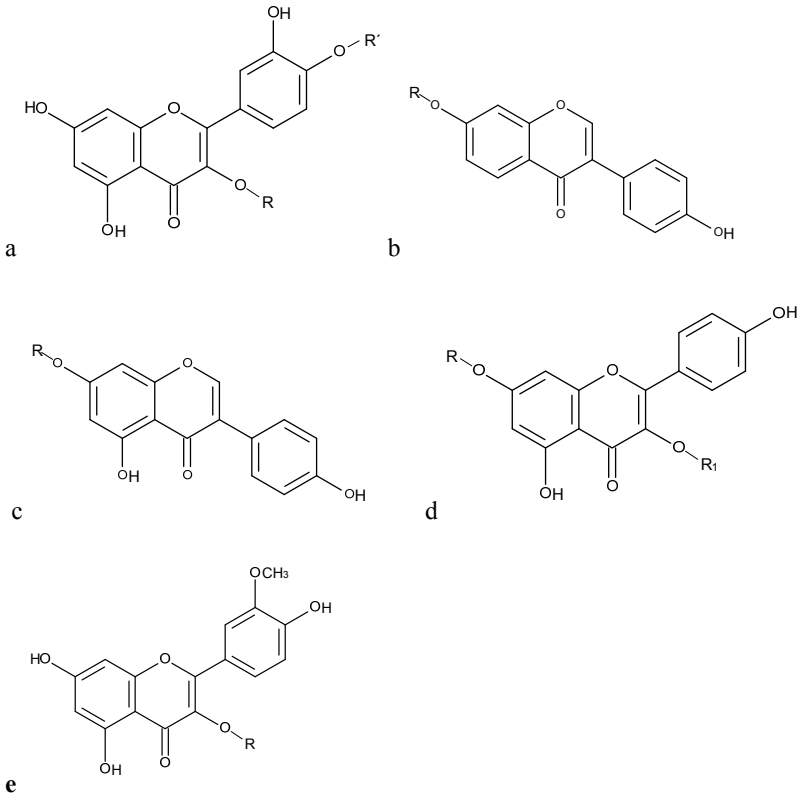


Figure 4. Chemical structure of substrates and product (a) quercetin (R, R´=H), Q3(R=glucose,R´=H), Q4´(R=H,R´=glucose), Q3,4´(R,R´=glucose) (b) Daidzein (R=H), D7(R=glucose) (c) Genistein (R=H), G7 (R=glucose), Kaempferol (R,R_1 =H), K3 (R=H,R_1 =glucose), K7 ((R=glucose,R_1 =H), Isorhamnetin (R=H), I3 (R=glucose)

Paper VI

Increased production of the active soluble form of the thermostable β-glucosidase *Tn*Bgl1A using chaperonin co-expression in *Escherichia coli*.

Samiullah Khan, S M Shariar Shakil, Gashaw Mamo, Eva Nordberg Karlsson
Biotechnology, Lund University, PO Box 124, SE-22100 Lund, Sweden

Abstract

The possibilities to characterize and use a recombinantly produced enzyme are dependent on a successful production strategy. The gene encoding the thermostable beta-glucosidase *Tn*Bgl1A from *Thermotoga neapolitana* was cloned under the control of the T7/lac-promoter in the pET22-vector and expressed in *Escherichia coli* strain Bl21 (DE3). The protein production was directed to the cytoplasm, and yielded a large amount of target protein, but the major part was produced in insoluble inactive form. To investigate possibilities to improve production of the active soluble enzyme in *E. coli*, the gene was synthesised with codon usage optimised for the host and cloned for production in the cytoplasm as well as in the periplasmic space. Cytoplasmic production of the codon-optimized synthetic gene again yielded large amounts of inactive insoluble protein. Secretion to the periplasm (using the pelB leader peptide) did not result in a changed pattern concerning the production of soluble or insoluble protein. Use of different combinations of *E.coli* chaperones in productions directed to the cytoplasm instead led to significant increases in the ratio of the active soluble enzyme, and was a successful strategy to increase the productivity of active thermostable beta-glucosidase in the *E. coli* production system.

Keywords: thermostable, *Thermotoga neapolitana*, GroEL, GroES, DnaK, DnaJ, GrpE

Introduction

The production step is still identified as a bottleneck for characterization and use of many novel enzymes. *Escherichia coli* is a popular choice for the production of foreign proteins for a number of reasons: it can grow at high cell densities, its genetic information is well established and a large number of cloning vectors have been developed to achieve high protein production levels. However, it is relatively common that the over-expressed target protein is accumulating in the cell as biologically inactive aggregates known as inclusion bodies. Sometimes, inclusion-body formation is used as strategy to obtain pure protein, *e.g.* when the protein of interest is toxic to the cell (Rudolph & Lilie, 1996) or is proteolytically unstable (Thomas, et al 1997), but this strategy requires refolding of the target protein. Refolding is often low-yielding and expensive, and hence a variety of different methods have been developed to instead increase the ratio of active soluble protein. These methods include cultivation optimisation (e.g. reducing growth temperature (Vasina & Baneyx, 1997) or inducer tuning (Turner et al, 2005)), genetic engineering (e.g. fusing the gene to a solubilising protein or peptide (Kapust & Waugh 1999; La Vallie et al, 2000; Turner et al, 2005) or co-expression of the target protein with molecular chaperonins, most frequently using GroEL/GroES (Gupta et al, 2009; Ushida et al, 2004).

The cultivation optimization strategies are often relying on a reduced production rate, which in turn results in a higher ratio of active/inactive enzyme, but at the same time reduce the overall productivity, that way reducing misfolded proteins (Vasina & Baneuyx, 1995; Thomas et al 1997). Fusion of the target protein to a solubilising protein [e.g. maltose-binding protein (Kapust & Waugh 1999), thioredoxin (La Vallie et al, 2000), or the Nus protein (Turner et al, 2005)] is often, but not always, resulting in a higher ratio of soluble active target protein, but a common problem is that the hybrid proteins are often expressed at lower levels than the unfused enzymes (Turner et al 2005; Thomas et al, 1997). In addition, molecular chaperones can be co-expressed with the target protein, and utilised to improve solubility of over-expressed proteins (Thomas et al, 1997). Among the folding modulators in *E. coli* cytoplasm, some combinations are available commercially, and are hence attractive alternatives for production optimisation. For instance the cytoplasmic modulators, DnaK-DnaJ-GrpE and GroEL/GroES molecular chaperones are together with a trigger factor (TF, a putative cotranslational folding catalyst) available in different combinations in vectors

compatible with the pET-expression vectors (Nishihara et al, 1998; Nishihara, et al., 2000), which are frequently used in *E. coli* laboratory productions.

In this work, we have evaluated the effect of periplasmic or cytoplasmic location on the production-level of soluble active β-glucosidase1A from *Thermotoga neapolitana,* and the effects of chaperonins on the enzyme produced in the cytoplasm. The target protein from the most successful production was purified using a C-terminal hexa-histidine-tag included in the cloning design, to assure that the downstream processing is not negatively affected by the modifications in the production step.

Materials & Methods

Chemicals

All chemicals were pro-analysi from MERCK Eurolab (Darmstadt, Germany) unless otherwise stated.

Gene synthesis, cloning and transformation

The wild type gene encoding the β-glucosidase Bgl1A was PCR-amplified from genomic *Thermotoga neapolitana* (DSM strain 4359) DNA, and cloned for cytoplasmic expression into the vector pET22b+ (Novagen, Madison, WI), resulting in the plasmid pBGL1A1c, incorporating a C-terminally locatedhexa-histidine-tag, as described by Turner et al, (2006). Primers were forward 5´- TAT TCT TAT CAT ATG AAA AAG TTT CCC GAA GGG TTC and reverse 5´- TAT TCT TAT CTC GAG ATC TGT TAG TCC GTT GTT TTT G. The synthetic gene encoding the same enzyme but with optimised codon usage for *E. coli* was synthesised (Eurofins MWG, GmbH), and cloned into pET22b+ for cytoplasmic expression (pBGL1A2sc) using the same approach as for the wild-type enzyme. In addition the gene was cloned in pET22b+ for expression in the periplasm, in frame with an N-terminally located pelB-leader peptide also in this case incorporating the C-terminally located hexa-histidine-tag (pBGL1A3sp). The primers used for amplification were in this case forward: 5'-ATT GCC ATG GCA AAA TTT CCG GAG GGC TTT C and reverse: 5'- TTG CCT CGA GAT CGG TCA GGC CAT TAT TCT TG. The restriction sites used for cloning in the respective primers are underlined. Gene and vector were ligated by T4 DNA ligase (Invitrogen Life Technologies, Frederick, MD) and transformed by electroporation into *E. coli* Nova Blue (Novagen, Madison, WI). Transformants were screened by PCR and one positive clone from each construct was sequenced (GATC, Germany). The pET-derived plasmids were

transformed into the expression strain *E. coli* BL21 (DE3) (Novagen, Madison, WI). Plasmid inserts were verified by restriction analysis.

The effect of *E.coli* chaperonins on folding was analysed by transformation of pBGL2sc to five different competent BL21 (DE3) cells, each containing a plasmid with different combinations of chaperonins (Table 1).

Expression

Cultivation of all variants was carried out in 100 ml Luria-Bertani medium containing 100 µg/ml ampicillin in baffled 500 ml E-flasks on a rotary shaker at 37°C. The chaperonin coexpression culture (with 20 µg/ml chloroamphenicol) producing the highest amount of active beta-glucosidase was also expressed in defined mAT medium (Ramchuran et al, 2002). Expression was started by adding IPTG (isopropyl-beta-D-thiogalactopyranoside) to a final concentration of 1mM, when the optical density at 620 nm was approximately 0.6. Chaperonin expression in selected cultures was, dependent on the plasmid used (Table 1), at the same time induced by arabinose (2 mg/ml) or tetracycline (5 ng/ml). Samples of 1 ml were taken every hour in the production phase and centrifuged immediately at 6000×*g* at 4°C for 5 min. Pellets and supernatants were stored at –20°C until analyzed.

Enzyme and Protein analysis

The cell pellets were dissolved in 500 µl of binding buffer (20 mM imidazole, 20 mMTris-HCl, 0.75 M NaCl, pH 7.5) and lysed by ultrasonication in a UP400S instrument (Dr. Hielscher GmbH, Stuttgart, Germany) at 60% amplitude, cycle 0.5 for 90 seconds. Soluble protein was separated from insoluble protein by centrifugation at 15000×*g* for 40 minutes at 4ºC. Enzyme activity was measured by a modified β-glucosidase assay using the substrate *p*NPGlc (*para*-nitrophenyl-β-D-glucopyranoside). Absorbance at 405 nm was measured as a function of time during 1 min. The extinction coefficient of *p*NP (*para*-nitrophenol) under these experimental conditions was determined as $\varepsilon_{80°C,\ 405nm} = 2.4639 \times 10^3$ mL $mmol^{-1}cm^{-1}$. 1 U corresponds to the amount of enzyme that will release 1 µmol of reducing sugar equivalents (expressed as glucose) per minute.

The protein content in the purified protein sample was estimated by the BCA-copper method with BSA as a standard (Sigma, Steinheim, Germany).

Soluble and insoluble protein was analyzed by SDS-PAGE according to Laemmli, (1970).

Purification

The cells were harvested by centrifugation (10000 × g, 10 min, 4°C),and dissolved in binding buffer (20 mM imidazole, 20 mMTris-HCl, 0.75 M NaCl, pH 7.5). The cell suspension was lysed by sonication for 5 × 3 min using a 14 mm titanium probe sound intensity of 60% and a cycle of 0.5 (UP400 S, Dr. Hielscher), centrifuged (30 min, 39000 × g, 4°C), heat treated (70°C, 30 min) and again centrifuged. The supernatant was purified on an ÄKTA prime system (Amersham Biosciences, Uppsala, Sweden) by immobilized metal ion affinity chromatography using copper as a ligand as described elsewhere (Turner et al, 2006; Khan et al, 2011)

Results

A synthetic gene encoding TnBgl1A for production in the cytoplasm or periplasmic space.

The thermostable β-glucosidase *Tn*Bgl1A from *T. neapolitana* is a single module glycoside hydrolase with a TIM-barrel fold and with a molecular mass of 53 kDa (Khan et al, 2011). To optimise the production conditions, the gene encoding *Tn*Bgl1A was synthesised with optimal codon usage for *E. coli* production (Fig. 1) and cloned in the pET22+ vector using the same restriction-sites as was used for the wild-type gene that instead was PCR-amplified from genomic DNA of *Thermotoga neapolitana* (Turner et al, 2006). Expression of the native and synthetic genes in the *E. coli* cytoplasm led to in principle identical expression profiles and in both cases the ratio of soluble active enzyme represented only a fraction of the total production which was dominated by insoluble target protein (Fig. 2).

Cloning in frame with the sequence encoding a leader peptide was used to direct the target protein to the periplasmic space. Production in the periplasm, can lead to improved yield of the soluble fraction of the target protein (Retallack et al, 2007), and an added advantage is that the target protein can be released by osmotic shock reducing number of proteins in the extract. Thus, the synthetic gene was also cloned in frame with the 22 amino acid *pelB* leader peptide encoded in the pET22+ vector, and the resulting construct was expressed at similar conditions as the target protein without the leader peptide. Periplasmic production of the beta-glucosidase did however improve neither activity (Table 2) nor production level. The overall production of the target protein in the periplasm was in principle following the pattern observed in the cytoplasmic production (Fig 2).

Coexpression of the synthetic TnBgl1A-gene encoding five variants of molecular chaperonins.

Co-expression of molecular chaperones has in many cases been shown to have a positive effect on solubilizing difficult target proteins by facilitating proper folding and assembly of the protein (Thomas & Baneyux, 1996; Thomas et al, 1997), although the extent naturally depends on the protein. Five different combinations of genes encoding chaperonins were coexpressed with the synthetic gene encoding the thermostable beta-glucosidase *E.coli* strain BL21 (DE3). All the chaperone encoding plasmids carry an origin of replication which is derived from pACYC (Nishihara, et al. (2000), making them compatible for use in *E.coli* together with pET expression plasmids that carry the ColE1 origin of replication (Table 1). Expression in the respective plasmid was induced when the optical density reached 0.6-0.7. The post induction phase was then allowed to proceed for 4 h. Coexpression with all the five chaperonins (DnaK-DnaJ-GrpE and GroES-GroEL) identified in the *E. coli* folding pathway led to significant improvement in the active soluble fraction of the thermostable beta-glucosidase (Table 2, Fig 3). Selective coexpression with only GroES-GroEL, or with only DnaK-DnaJ-GrpE also led to improvements, but to a lower extent. A similar, relatively low improvement of the active soluble target protein was seen after coexpression with the trigger factor, alone or together with GroEL/ES (Table 2). This shows that cooperation of many different chaperonins in the folding pathway appears to be necessary for a good production yield of soluble active beta-glucosidase using the strong T7/lac promoter in the *E. coli* production system.

Production in complex and defined media, and purification of the produced β-glucosidase by immobilized metal ion affinity chromatography.

The activity screening of the different constructs described above were all made in complex LB-medium. To investigate the influence of the medium composition on the production, the plasmid combination that upon expression showed significantly improved target protein activity (using both the pBGL1A2sc encoding the beta-glucosidase, and the pG-KJE8 plasmid, encoding the DnaK-DnaJ-GrpE, GroEL/ES chaperonins) was also expressed in a defined medium, using glucose as the sole carbon source. Our results showed that use of the defined medium decreased the activity obtained from the cultivation (6.6 U/ml i mAT compared to 9.7 U/ml in LB), combined with a visible decrease of the soluble fraction in SDS-PAGE (Fig. 4). The production levels were however still favourable compared to other

coexpression combinations, showing that the defined medium can be utilized if this is desirable from a cultivation control perspective [*e.g.* allowing use of glucose limited fed-batch technology (Ramchuran et al, 2002; Ramchuran et al, 2005)].

Purification of the β-glucosidase from the chaperonin coexpression culture was made as a two step procedure using heat treatment, followed by immobilised metal ion chromatography step utilizing the C-terminal hexa-histidine tag included in the cloning design (Table 3). This resulted in a purity of approximately 80-90 % (data not shown) which was comparable to the purity obtained in a similar procedure using extracts from productions where the β-glucosidase was produced without the aid of the chaperonins. This shows that the presence of the chaperonins do not interfere in the down-stream processing steps, securing the possibilities to purify the target protein with comparable methodology.

Discussion

In this work we show that the production of active soluble thermostable β-glucosidase is significantly enhanced by the coexpression of the *E.coli* chaperones. The two molecular chaperonin teams (DnaJ-DnaK-GrpE and GroEL/GroES) are known to suppress off pathway aggregation reactions and to promote folding via ATP-coordinated cycles where folding intermediates are bound and released (Thomas et al, 1997).

Among the five chaperones encoding plasmids (divided into the teams DnaK-DnaJ-GrpE and GroEL/ES) (Nishihara et al, 1998), the construct holding both teams of chaperones appeared to be the most efficient for keeping the target protein in solution. This is in accordance with the cooperative role of the two chaperone teams (Gragerov et al, 1992). In a first step, shielding of hydrophobic stretches is made by DnaK-DnaJ in the newly synthesised chain, contributing to obtain correct folding by preventing interactions with folding intermediates and cellular components. The GrpE binding to DnaK then catalyse release of the bound protein. Coexpression of DnaK-DnaJ-GrpE team did indeed lead to increased activity of TnBgl1A, but only to a moderate extent (approximately 1.3 × increase). The more frequently utilised chaperonin team GroEL/ES was slightly more efficient in aiding production of the active beta-glucosidase (1.7 × increase). In principle, the effect using either of the two chaperonin teams was in the same range. The GroEL/ES chaperonins are believed to facilitate productive folding of proteins that have reached an intermediate folding state (Thomas et al, 1997). Use of the GroEL/ES chaperones in coexpressions of the target protein is well established, and has in many cases been shown to increase the yield of correctly folded target

proteins (reviewed by Thomas et al, 1997). However, the effect is dependent on the target protein, and many cases are also known where the chaperones are not aiding production of the active form. For *Tn*Bgl1A the coexpression of GroEL/ES did almost double the activity of the enzyme. Addition of all five chaperonins however led to a 6× increase in the activity, showing the expected cooperative effect when all chaperones where present during the production.

Addition of the trigger factor to the GroEL/ES team did not further improve the activity of the target enzyme, despite the fact that the trigger factor (TF) alone led to an increase in activity approximately corresponding to that shown using one of the two chaperone teams described above. This factor is suggested to play a role folding as it can be found associated to the nascent polypeptide, and also interact with the GroEL chaperone (Kandror et al, 1997; Nishihara et al, 2000).

In this work we have utilised fixed concentrations of the inducers for the chaperone productions, thus additional improvements of the production levels may be achieved by further optimizations in these levels, as indicated by Nishihara and co-workers (2000). We also see that the activity levels are higher in small batch cultivations using complex medium, so also in the choice of medium is affecting the overall production of the active beta-glucosidase. In all, we have shown that coexpression of *E.coli* chaperonins is significantly improving the yield of folded beta-glucosidase, significantly increasing the productivity of the protein, facilitating further use and characterization of this thermostable enzyme.

Acknowledgements

Financial support from the EU FP7 program AMYLOMICS and the Swedish research council Formas (229-2009-1527) funding the SuReTech research collaboration is greatly appreciated.

References

Gupta P., Gholsalkar, A., Mishra, S., Chaudhuri, T.K. Enhancement of over expression and chaperone assisted yield of folded recombinant aconitase in *Escherichia coli* in bioreactor cultures. *J Biosci Bioeng*, 107, 102-107 (2009)

Gragerov, A., Nudler, E., Komissarova, N., Gaitanaris, G. A., Gottesman, M. E., Nikiforov, V. Cooperation of GroEL/GroES and DnaK/DnaJ heat shock proteins in preventing protein misfolding in *Escherichia coli*. *Proc. Natl. Acad. Sci. USA,* 89, 10341–10344 (1992)

Kandror, O., Sherman, M., Moersholl, R., Goldberg, A.L. Trigger factor associates with GroEL in vivo and promotes its binding to certain polypeptides. *J. Biol. Chem,* 272, 1730-1734 (1997)

Kapust, R. B., Waugh, D. S. *Escherichia coli* maltose-binding protein is uncommonly effective at promoting the solubility of polypeptides to which it is fused, *Protein Sci,* 8, 1668-1674 (1999)

Khan, S., Pozzo, T., Megyeri, M., Lindahl, S., Sundin A., Turner C., Nordberg Karlsson, E. Aglycone specificity of *Thermotoganeapolitana* β-glucosidase 1A modified by mutagenesis, leading to increased catalytic efficiency in quercetin-3-glucoside hydrolysis. *BMC Biochem*, 12:11 (2011)

Laemmli, U.K. Cleavage of structural proteins during assembly of the head of bacteriophage T4, *Nature,* 227, 680-685 (1970)

LaVallie, E. R., Lu, Z., Diblasio-Smith, E. A., Collins-Racie, L. A., and McCoy, J. M. Thioredoxin as a fusion partner for production of soluble recombinant proteins in *Escherichia coli*. *Methods Enzymol,* 326, 322-340 (2000)

Nishihara, K., Kanemori, M., Kitagawa, M., Yanagi, H., Yura T. Chaperone Coexpression Plasmids: Differential and Synergistic Roles of DnaK-DnaJ-GroE and GroEL-GroES in Assisting Folding of an Allergen of Japanese CedarPollen, Cryj2, in *Eschericia coli*. *Appl. Environ. Microbiol,* 64, 1694-1699 (1998)

Nishihara, K., Kanemori, M., Yanagi, H., Yura, T. Overexpression of Trigger Factor Prevents Aggregation of Recombinant Proteins in Escherichia coli. *Appl. Environ. Microbiol*, 66, 884-889 (2000)

Nordberg Karlsson, E., Holst, O. Tocaj, A. Efficient production of thermostable xylanases from *Rhodothermus marinus* in *Escherichia coli* fed-batch cultures. *J. Biosci. Bioeng,* 87, 598-606 (1999)

Ramchuran, S., Nordberg Karlsson E., Velut, S., de Maree, L., Hagander, P., Holst, O. Production of heterologous thermostable glycoside hydrolases and the presence of host-cell proteases in substrate limited fed-batch cultures of *Escherichia coli* BL21(DE3) *Appl. Microbiol. Biotechnol,* 60, 408-416 (2002)

Ramchuran, S. O., Holst, O., Nordberg Karlsson, E. Effect of postinduction nutrient feed composition and use of lactose as inducer during production of thermostable xylanase in *Escherichia coli* glucose-limited fed-batch cultivations *J. Biosci. Bioeng*, 99, 477-484 (2005)

Retallack, D.M., Schneider, J.C., Mitchell, J., Chew, L., Liu, H. Transport of heterologous proteins to the periplasmic space of *Pseudomonas fluorescens* using a variety of native signal sequences. *Biotechnol. Letters,* 29, 1483-1491 (2007)

Rudolph, R. and Lilie, H.In vitro folding of inclusion body proteins. *FASEB J* 10, 49-56 (1996)

Thomas, J. G., and Baneyx, F. Protein misfolding and inclusion body formation in recombinant *Escherichia coli* cells overexpressing heat-shock proteins, *J. Biol. Chem,* 271, 11141-11147 (1996)

Thomas, J. G., Ayling, A., Baneyx, F. Molecular Chaperones, Folding Catalysts, and the Recovery of Active Recombinant Proteins from *E.coli*. *Appl. Biochem. Biotech*, 66, 197-238 (1997)

Turner, C., Turner, P., Jacobson, G., Waldebäck, M., Sjöberg, P., Nordberg Karlsson, E., Markides, K. Subcritical water extraction and β-glucosidase catalyzed hydrolysis of quercetin in onion waste. *Green Chem,* 8, 949-959 (2006)

Turner, P., Holst, O., Nordberg Karlsson, E. Optimized expression of soluble cyclomaltodextrinase of thermophilic origin in *E.coli* by using a soluble fusion-tag and by tuning of inducer concentration *Prot. Express. Purif,* 39, 54-60 (2005)

Ushida E., Kagawa, N., Sakaki, T., Uruchino, N., Sawada, N., Kamakura, M., Ohta, M., Kato, S., Inouye, K. Purification and characterization of mouse CYP27B1 overproduced by an *Escherichia coli* system coexpressing molecular molecular chaperonins GroEL/ES. *Biochem. Biophys. Res. Commun,* 323, 505-511 (2004)

Vasina, J.A., Baneyx, F. Expression of aggregation-prone recombinant proteins at low temperatures: a comparative study of the *Escherichia coli cspA* and *tac* promoter systems, *Prot. Express.Purif.,* 9, 211-218 (1997)

Figures:

Figure 1. The coding sequence of the synthetic gene (red), optimised for codon usage in *Escherichia coli* is shown together with the original gene (PCR-amplified from genomic DNA of *Thermotoga neapolitana*). Red boxes indicate codons that are identical in the original and synthetic gene. The complete gene is has 444 codons, of which 183 are unchanged in the synthetic gene. The majority of the changes (229 codons) only concern changes in the third position, while 11 codons (at 32 positions) encoding the three amino acids arginine, serine and leucine are changed to alternative codons. [S: TCG (to AGC), TCC (to AGC), TCT (to AGC), TCA (to AGT) and in one case AGC to TCG, R: AGA (to CGG, CGC, CGT), AGG (to CGC, CGT), L: CTC (to TTG), CTG (to TTA), CTT (to TTA, TTG) and also TTG (to CTG, CTC).]

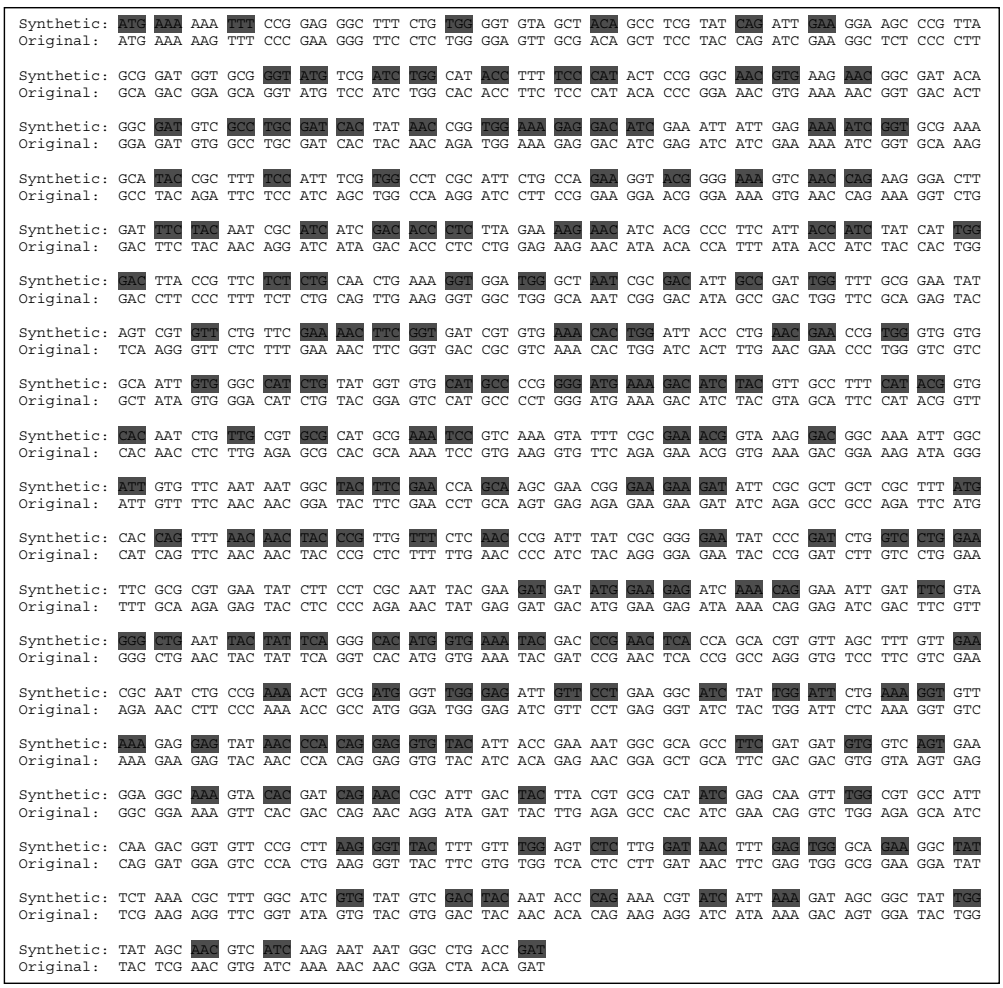


Figure 2. Production of *Tn*Bgl1A. Expression of the synthetic gene [lane 1 soluble (S), lane 2 insoluble (I) protein fraction] and original gene [lane 4 (S) and 5 (I)] in the cytoplasm of *E. coli* strain Bl21(DE3). Expression of the synthetic gene cloned in frame with a sequence encoding the pelB-leader peptide for production in the periplasmic space [lane 7 (S), lane 8 (I)] resulted in a similar production pattern. Lane 3 and 6 shows the molecular weight marker. The beta-glucosidase has a molecular weight of 53 kDa (indicated by an arrow in the figure). M= molecular weight marker. The molecular weights of the protein standard are indicated to the right.

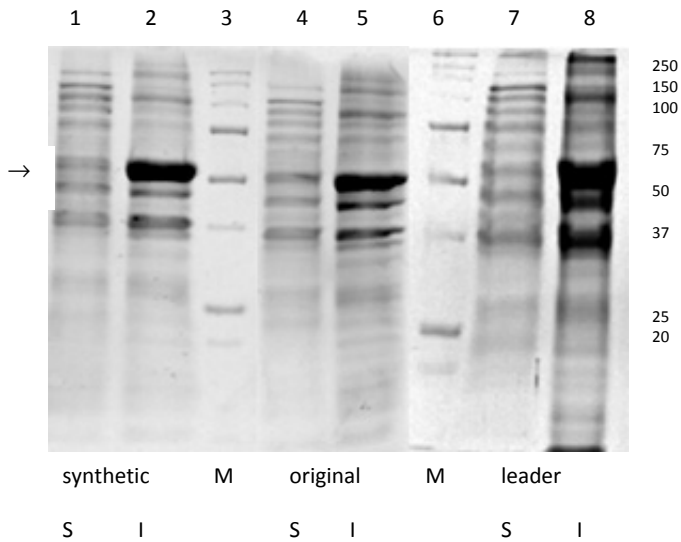
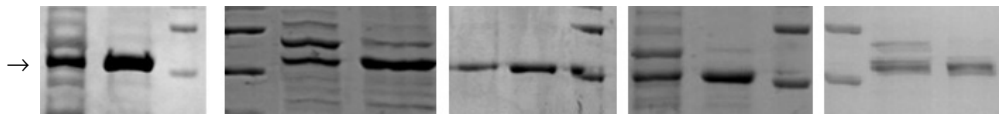


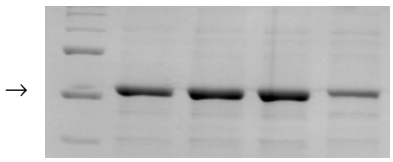
Figure 3. Soluble (S) and insoluble (I) *Tn*Bgl1A, 4 h after induction in cultures coexpressing different chaperonin combinations. The position of *Tn*Bgl1A (53 kDa) is indicated by an arrow on the left side. The standard protein bands (M) corresponding to molecular weight 75 and 50 kDa are shown for each cultivation. The 60kDa chaperonin GroEL is visible in the soluble protein fraction in cultures with plasmid pG-KJE8, pGro7 and pG-Tf2. In pTf16 the 56 kDa trigger factor is visible.



S I M M S I S I M S I M M S I

pG-KJE8 pGro7 pKJE7 pG-Tf2 pTf16

Figure 4. Comparison of production levels using complex (LB) and defined (mAT) growth medium in the culture coexpressing the *Tn*Bgl1A encoding gene (from pBGL1A2cs) and the 2 chaperonin teams (from pG-KJE8). The lanes from left to right: M (molecular weight marker), SLB (soluble fraction in LB-medium), ILB (insoluble fraction in LB), ImAT (insoluble fraction in mAT-medium), SmAT (soluble fraction in mAT-medium). The position of TnBgl1A is indicated by an arrow.



Tables:

Table 1. Chaperone and beta-glucosidase encoding plasmids utilized in the production trials. The chaperonin encoding plasmids are commercially available from TaKaRaBioSci, inc.

Plasmid	**Encoded Chaperones/Enzyme**	**Promoter**	**Inducer**	**Selection Marker**	**Origin of replication**	**Reference**
pG-KJE8	DnaK-Dnaj-GrpE GroES-GroEL	*araB Pzt-1*	L-Arabinose Tetracycline	Cm	pACYC	Nishihara et al, 1998
pGro7	GroES-GroEL	*araB*	L-Arabinose	Cm	pACYC	Nishihara et al, 1998
pKJE7	DnaK-Dnaj-GrpE	*araB*	L-Arabinose	Cm	pACYC	Nishihara et al, 1998
pG-Tf2	GroES-GroEL-tig	*Pzt-1*	Tetracycline	Cm	pACYC	Nishihara et al, 2000
pTf16	tig	*araB*	L-Arabinose	Cm	pACYC	Nishihara et al, 2000
pBGL1A1c	TnBgl1A	*T7/lac*	IPTG	Amp	ColE1	Turner et al, 2006
pBGL1A2sc	TnBgl1A (from synthetic gene)	*T7/lac*	IPTG	Amp	ColE1	this work
pBGL1A3sp	TnBgl1A (from synthetic gene)	*T7/lac*	IPTG	Amp	ColE1	this work

Table 2. Activity levels for the different constructs in the heat treated extract. (heat treatment 70°C, 30 min). The cdw was determined using a standard curve (cdw vs. OD at 620 nm) obtained from several batch cultivations without induction (Nordberg Karlsson et al, 1999) and 1 OD-unit was equal to 0.44 g cdw /L. (n.a.= not applicable). The plasmid encoding the synthetic gene used in the chaperone experiments is shown in bold.

Plasmids in Bl21(DE3)	**Activity (U/ml)**	**Final cdw (g/L)**	**Activity/cell dry weight U/g**	**Ratio**
pG-KJE8, pBGL1A2sc	9.7	1.9	5.1	5.7
pGro7, pBGL1A2sc	2.7	1.8	1.5	1.7
pKJE7, pBGL1A2sc	2.1	1.8	1.2	1.3
pG-Tf2, pBGL1A2sc	2.2	1.5	1.5	1.7
pTf16, pBGL1A2sc	2.3	1.4	1.6	1.8
pBGL1A2sc	**1.6**	**1.7**	**0.9**	**0.9**
pBGL1A1c	1.5	2.6	0.6	n.a.
pBGL1A3sp	1.8	1.9	0.9	n.a.

Table 3. Purification table of TnBgl1A after a production using coexpressed DnaK-DnaJ-GrpE and GroEL/ES.

Step	Volume (ml)	Total protein (mg)	Activity (U)	Recovery (%)	Specific activity (U/mg)	Purification (fold)
Cell extract	86	70	800	100	11	1
IMAC	0.5	0.4	200	25	500	45