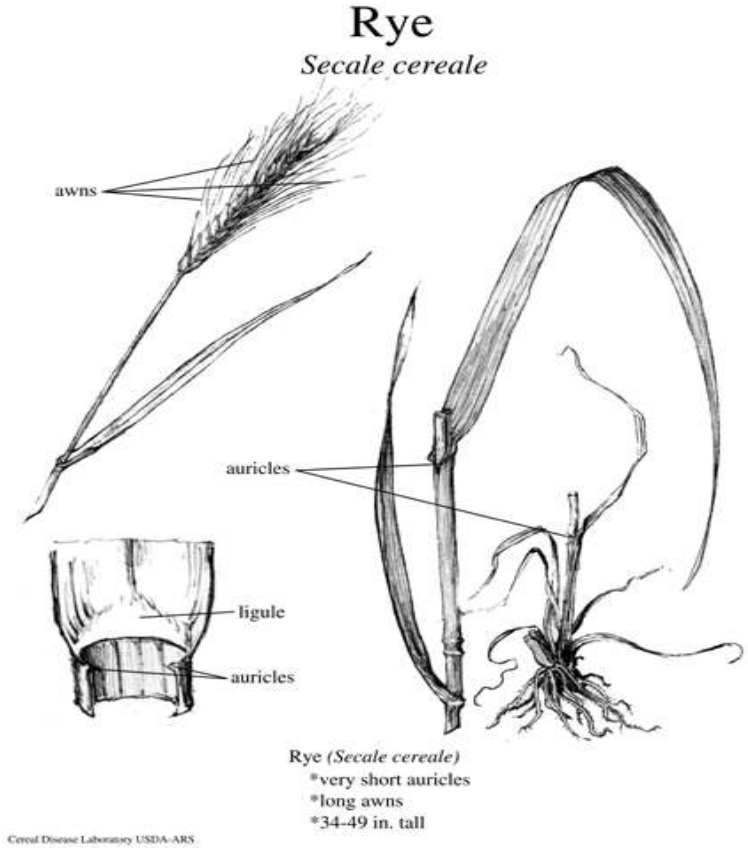
Rye cell wall β-glucosidase: subcloning, expression and purification of recombinant protein from *E.coli*

Nicolas Rochereau
François Rabelais University, Tours, France

January 2007 – June 2007
Södertörn University College, Huddinge, Sweden



Examiner: Assoc Prof Magnus Johansson
Supervisors: MSc Therese Gradin, Assoc Prof Gabriele Delp and Prof Lisbeth Jonsson, Dept.of Life Sciences, Södertörn University College.

Abstract

Several plant defense systems consist of enzymes that act on glucosides and produce a toxic compound. In the intact plant tissue the substrate and enzyme are kept apart. The system studied here consists of the substrate 2-*O*-β-D-glucopyranosyl-4-dihydroxy-1,4-benzoxazin-3-one and the enzyme glucan 1,3-β-glucosidase in rye. The aim was to determine the properties of a cell wall β-glucosidase.

Two different systems for expression and purification of β-glucosidase fused to a tag were used: a 6xHistidine tag system and a thioredoxin tag system. The sequence of the β-glucosidase had previously been determined so now the gene was subcloned into *E.coli*. A direct PCR on colonies, a test expression, a restriction digestion of plasmids and sequencing was made to analyze the transformation, which all turned out successful. Then the β-glucosidase solubility was determined. Finally a purification of the β-glucosidase from *E.coli* under native conditions and a pNPG assay was carried out.

For the $(His)_6$-tagged protein, the recombinant β-glucosidase tended to end up in the insoluble pelleted fraction which indicated formation of inclusion bodies.

The cell wall 1,3-β-glucosidase was soluble with the thioredoxin system, but the percentage of soluble protein fraction was around 5% only of the total protein. In eluates from a nickel-nitrilotriacetic acid column the presence of recombinant protein was confirmed with Western blot, but contaminating bands were also present. Purified elauted fractions did not exhibit detectable β-glucosidase activity. It was not possible to purify active enzyme. From a BLAST search it was clear that the most similar enzymes all had putative glycosylation sites and lack of glycosylation could be a reason for the protein not to fold properly.

Introduction

Since plants cannot move they have come up with different ways of defending themselves, other than running away. Secondary metabolites are often involved in the chemical defense mechanisms of plants. Plants make a lot of secondary metabolites. Many of these are highly dangerous to various attackers and are often stored in specific vesicles or in the vacuole (Poulton 1990). Hydroxamic acids (Hx), which are the most considerably studied secondary metabolites, play a role in the defence of cereals, in the detoxification of herbicides and in allelopathic effects of the crops (Niemeyer 1988). Several systems consist of enzymes that act on glucosides and produce a toxic compound. The system studied here consists of the substrate DIBOAGlc (2-*O*-β-D-glucopyranosyl-4-dihydroxy-1,4-benzoxazin-3-one) which is an Hx and the enzyme glucan 1,3-β-glucosidase.
The system is present in rye (*Secale cereale*), maize (*Zea mays*) and wheat (*Triticum aestivum*). In the intact plant tissue the substrate and enzyme are kept apart. The β-glucosidase comes together with the substrate (DIBOAGlc) when the tissue is broken, for example at an insect attack. The β-glucosidase acts on the glucosides and breaks the glycosidic bond. Then DIBOA (2,4-dihydroxy-l,4-benzoxazin-3-one) is formed which spontaneously produces the toxic compound BOA (2-benzoxazolinone).

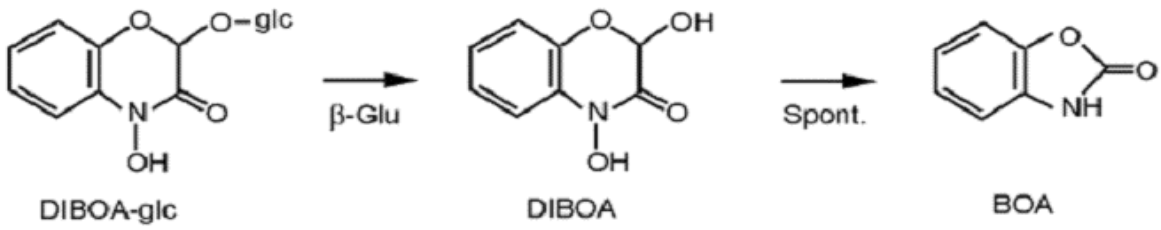


Figure 1. The reaction catalysed by rye β-glucosidase and the subsequent transformation of DIBOA to the more toxic benzoxalinone, BOA.

A lot of things are known about the β-glucosidases. For example β-glucosidases cut glycosidic linkages, they have got an acidic pH optima (pH 5-6), they are ubiquitous in the living world (mammal, plant, insect, fungus, and bacterium) (Esen 1993). β-glucosidases have been implicated in numerous different plant functions, such as growth regulation by the release of active plant hormones (Dietz el al. 1999), lignification and plant defense by the release of plant toxins (Dharmawardhana et al. 1999).

This work will focus on a β-glucosidase in rye. At the subcellular level, β-glucosidases were found both in the cell wall and in the plastid (Nikus et al. 2001). The sequence of the plastidic β-glucosidase was determined, the gene was cloned, purified and the biochemichal properties of the enzyme was defined. The native cell wall 1,3-β-glucosidase was partly purified and had an apparent molecular weight of about 68-69 kDa (Nikus and Jonsson 2003). The cell wall β-glucosidase belongs to the family 3, whereas the plastidic β-glucosidase belongs to family 1 (Henrissat and Davies 1997).

The aim of this project is to try to confirm the results concerning the properties of the partly purified cell wall β-glucosidase by cloning and purification of recombinant protein. The recombinant protein will be characterized with regard to its molecular and kinetic properties.

Based on the peptide sequences of the isolated protein primers could be designed and a full-length cDNA clone was isolated using GeneRacer Kit (Invitrogen, Paisley, UK) (F.Michoux, personal communication). The gene, referred to as "CWR" (Cell Wall Recombinant), was cloned into pCR®-4-TOPO®, referred to as "D4". The construct was sequenced and the result was submitted to GenBank (Accession Number AY586531).

In a previous attempt to subclone and then purify the recombinant protein using the IMPACT (Intein Mediated Purification with an Affinity Chitin-binding Tag)-system (New England Biolabs, Beverly MA, USA) the protein ended up in inclusion bodies and was not soluble. The purification failed (Roos 2006).

Materials and methods

Two approaches were made, $(His)_6$ tag (Qiagen, Solna, Sweden) and Thioredoxin tag (Invitrogen). The $(His)_6$ affinity tag facilitates binding to nickel-nitrilotriacetic acid (Ni-NTA). Indeed, $(His)_6$-tagged protein are bound to metal-chelating surfaces: Ni-NTA or cobalt resins. So the protein could be purified with a Ni-column or Co-column. Thioredoxin is able to accumulate up to 40% of the total cellular protein and still remain soluble. The thioredoxin protein has been mutated to contain a metal binding domain. The mutated histidines interact with a native histidine to form a patch. His-Patch thioredoxin proteins can terefore be purified on metal chelating resins.

Subcloning of the CWR gene

CWR-Qifull.R	→ 5′-AAT TAG AGT GGT ACC CTA CTT GTT GGC CTC GGT GGT- 3′
CWR-QiMat.F	→ 5′-GAT CAT CTC AAG TAC AAG GAT CCG AAG CAG CCT ATT- 3′
GF1	→ 5′-CCT TAC GCT TTT GCT CCG TGT- 3′
GR1	→ 5′-GCT GTC AGG ATT CTC TGA GA- 3′
CWRpET102M.F	→ 5′-CAC CGA TCA TCT CAA GTA CAA GGA TC- 3′
CWRpET102.R	→ 5′-CTA CTT GTT GGC CTC GGT GGT GAG- 3′

Table 1 showing all primers used during cloning and sequencing.

The target gene, CWR, was amplified from the sequencing vector, D4, by PCR (98°C 30 sec, (98°C 10 sec, 71°C 30sec, 72°C 1min) x 31, 72°C 5min) using the Phusion Hot Start High-Fidelity DNA Polymerase kit (Finnzymes, Espoo, Finland). The primers used for the $(His)_6$-system were CWR-QiMat.F, which contained the restriction site for restriction enzyme Kpn I to facilitate directional cloning, and CWR-QiFull.R. The primers used for the thioredoxin system were CWRpET102M.F and CWRpET102.R. The PCR reaction contained: Sterile water 12,4 µL, buffer 4 µL, dNTP 0,4 µL, template (D4) 1 µL, primers 1 µL, polymerase 0,2 µL. Both the HF and GC PCR-buffer were tested. The PCR-product for CWR-QiMat.F and CWR-QiFull.R was cleaned on a spin column using GeneClean (Qbiogene, Irvine, United States). The PCR-product for CWRpET102M.F and CWRpET102.R was run on a 1% agarose gel and cleaned using GeneClean (Qbiogene).
The PCR-product for the His_6-system was digested with Kpn I (New England Biolabs, Ipswich, England). The pQE-1 expression vector (Qiagen) was double digested with Kpn I and Pvu II that produces a blunt end. The PCR-product and vector were gel cleaned and purified with GeneClean (Qbiogene).

The restriction digested pQE-1 vector was combined with a 3 molar excess of insert using Quick Ligation Kit (New England Biolabs).
The PCR-product for the thioredoxin system was ligated into the pET102-D/TOPO vector using cloning protocol according to the supplied manual (Invitrogen).

Transformation

For the $(His)_6$ tag system, *E.coli* strain M15 [pREP4] were prepared according to the QIAexpressionist manual (Qiagen).
An aliquot of M15 [pREP4] cells were mixed carefully with 10 µL of ligation mix and kept on ice for 20 min, transferred to a 42°C heating block for 90 s. Psi broth was added to the mix and after incubation for 90 minutes at 37°C. 100 and 200 µL aliquots were plated out on Luria-Bertani (LB)-agar plates containing 25 µg/ml kanamycin and 100 µg/ml ampicillin. The plates were incubated at 37°C overnight. Positive and negative control plates were prepared. The positive control was made with a control plasmid included in the kit, and the negative control was made without plasmid.

For the thioredoxin system, *E.coli* strain BL21 Star™ (DE3) One Shot® were prepared according to supplied manual (Invitrogen). An aliquot of BL21 cells were mixed carefully with 2 µL of ligation mix and kept on ice for 30 min, transferred to a 42°C heating block for 30 s. S.O.C medium was added to the mix and after incubation for 30 minutes at 37°C. 100 and 200 µL aliquots were plated out on LB-agar plates containing 100 µg/ml ampicillin. The plates were incubated at 37°C overnight.

Direct PCR on colonies

Five colonies, a1-a5 for the $(his)_6$-system and 1-5 for the thioredoxin system respectively, were picked from the LB-plate, diluted in 200 µL H_2O and boiled 5 minutes. 5 µL were used in the PCR reaction.

These 2x5 colonies were analyzed by PCR with CWR-Qifull.R (10 µM), and CWR-QiMat (10 µM). For the thioredoxin system GF1 (10 µM) and GR1 (10 µM) were used. PCR reaction (94°C 30sec, (94°C 30sec, 55°C 30sec, 72°C 1min) x 30) was performed using Taq Polymerase (Fermentas, Burlington, Canada). The PCR reaction contained: Sterile water 8,4 µL, 10X PCR buffer KCl (with 2 mM $MgCl_2$) 4 µL, dNTP 0,4 µL, colony (dilute and boil) 5 µL, primers 1 µL.

Test expression

LB media (25µg/ml kanamicyn and 100 µg/ml ampicillin) was inoculated with the a1, a2, a3, a5 colonies. At OD_{600} ~0,6 the cultures was split in two and one half was induced with 0,5 mM IPTG. The other one was kept as control. All cultures were incubated for 4 hours at 37°C. 1 mL of both cultures were spun down for 15 minutes at 5000 rpm. The cell-pellet was boiled directly in 200 µL 1xSDS Sample Buffer for 4 minutes. The mix was centrifuged for 15 minutes at 10 000 rpm. Finally 20 µL was mixed with 4xSDS Sample buffer (Tris/HCl pH 6, 252 mM, glycerol 40%, SDS 8%, bromophenol blue 0,04%) and loaded onto a 4-20% Tris-Glycine Duramide gel (Cambrex, Rockland, USA).

Restriction digestion of plasmids

Extraction of pQE-1 with CWR insert was performed using Plasmid midi kit (Qiagen). KpnI and EcoRI (New England Biolabs) were used for the restriction digestion. The expected sizes of fragments were 3381 bp, 1398 bp, and 469 bp.

Sequencing

The plasmids pQE-1-CWRa1 and pET102-D/TOPO-CWRa were sequenced (Cybergene).

Expression

Determination of β-glucosidase protein solubility

LB media (25 µg/ml kanamycin and 100 µg/ml ampicillin) was inoculated with an overnight culture of colony a1 for the $(His)_6$-system and 1 for the thioredoxin system. At OD_{600} of ~0,6 the culture was induced with 1 mM IPTG. A 1 mL sample (uninduced control) was removed before induction, pelleted and resuspended in 50 uL 1xSDS Sample Buffer. The culture was incubated for 4 hours at 37°C. A 1 mL sample (induction control) was removed after induction, pelleted and resuspended in 50 µL 1xSDS Sample Buffer. The culture was centrifuged at 4000xg for 20 minutes, resuspended in 5 mL lysis buffer (NaH_2PO_4 50 mM, NaCl 300 mM, imidazole 10 mM, PMSF 20 µM, pH 8). Cells were lysed by repeated freeze-thawing in liquid nitrogen and cold water. 5 µL respectively of the soluble protein fraction and the insoluble protein fraction was mixed with 1xSDS Sample buffer. All samples were loaded onto a 4-20% Tris-Glycine Duramide gel (Cambrex). The lysate was centrifuged at 10 000xg

at 4°C for 30 minutes. The supernatant was decanted (soluble protein), and the pellet was resuspended in 400 µL 1xSDS Sample Buffer (insoluble protein). Other media, incubation temperatures, lengths of induction and concentrations of IPTG were tested; 2xYeast Tryptone (2xYT) and Terrific Broth (TB) were other media tested, 25°C and 30°C for incubation, 2,5 hours of induction, 0,5 mM of IPTG concentration.

Western blot

A Western blot was performed with samples from the solubility test of the pET102-D/TOPO-CWR1. A PVDF-membrane (Amersham Biosciences, Amersham place, UK) was put in 100% MeOH for 1 minute, Milliq H_2O for 5 minutes and in transfer buffer (96 mM Glycine, 12 mM Tris-Base, 10% MeOH) for 5 minutes. The filter papers and the SDS-PAGE gel were soaked in transfer buffer for a few minutes. Three filter papers, the PVDF-membrane, the gel, and another three filter papers on top were placed in the SemiPhor (Hoefer Scientific Instruments, San Francisco, USA).
The membrane was placed with TBS-T (20 mM Tris-HCl, 500 mM NaCl, pH 7,5, Tween-20, w/v) to rinse any gel residues off for 2x5 minutes. The membrane was then put in blocking solution (25 mL TBS-T + 3% BSA, w/v) for 1 hour at room temperature. The membrane was washed with TBS-T for 2x5 minutes and incubated with primary antibody, Anti-Thio^{TM} Antibody (Invitrogen) in dilution solution (25 mL TBS-T + 1% BSA, w/v) for 1 hour at room temperature. The primary antibody was supposed to bind to the thioredoxin fusion protein. The membrane was washed with TBS-T for 2x5 minutes and incubated with secondary antibody, the affinity purified HRP conjugated Anti-mouse IgG secondary antibody (Sigma, Steinheim, Germany) for 1 hour at room temperature. After final wash steps of TBS-T for 2x5 minutes and TBS for 20 minutes the membrane was incubated in HRP-substrate (5,5 mL 100% MeOH+ 16,5mg 4CN, 21 mL TBS, 40 µL 30% H_2O_2) until color developed.

Large scale expression

1 L of LB media (100 µg/ml ampicillin) was inoculated with an overnight culture of colony 1 for the thioredoxin system. At OD_{600} of ~1,3 the culture was induced with 0,5 mM IPTG. The culture was incubated for 2,5 hours at 37°C, and then the cells were harvested by centrifugation at 3000xg at 4^{o}C for 10 min. The cells were resuspended in 5 mL of lysis buffer. 1 mg/ml of lysozyme was added and incubated on ice for 30 min. After 10 x 6 sec of sonication, DNase I was added and incubated on ice for 1 hour. The lysate was centrifuged at 10000xg for 30 min at 4°C, and the supernatant was saved for the purification.

Purification of the β-glucosidase from *E.coli* under native conditions

1 mL of the 50% Ni-NTA was added with 4 mL cleared lysate and mixed at 4°C for 1 hour. The mixture was added onto a column. First the flow-through was collected, then the column was washed twice with wash buffer (NaH_2PO_4 50 mM, NaCL 300 mM, imidazole 20 mM, pH 8). Samples from the wash steps were also collected. Finally,

the protein was eluted 4 times with 0,5 mL of elution buffer (NaH_2PO_4 50 mM, NaCl 300 mM, imidazole 250 mM, pH 8).

Characterization

pNPG assay

70 µL of substrate (5 mM pNPG (*p*-nitrophenyl glucopyranoside) in 100 mM citrate-phosphate buffer) was mixed with 70 µL of enzyme. The mixture was incubated at 37°C for 15 to 45 min. The reaction was stopped with 70 µL of 0,4 M Na_2CO_3. pNP (*p*-nitrophenol) concentration was determined using a spectrophotometer at 405 nm. A negative control was made with H_2O instead of pNPG. A positive control was made with purified native plastidic β-glucosidase instead of cell wall β-glucosidase. Protein concentration was determined using the Bradford method. 5 µL of sample was mixed with 155 µL dH_2O and 40 µL of Bradford reagent. A standard curve was made with defined concentrations of BSA (Bovine Serum Albumine). Blanks were made without sample and dH2O instead. After a 10 minutes incubation the results were analyzed in a BIO-RAD Model 550 Microplate Reader.

Bioinformatics

The amino acid sequence of the cell wall β-glucosidase was submitted to the BLAST sequence alignment program (http://www.ncbi.nlm.nih.gov/BLAST/). Clustal W at the EBI server (www.ebi.ac.uk/clustalw/) was used to align the sequences and to calculate the sequence alignment scores. In order to investigate putative glycosylation the amino acid sequences of the plastidic and the cell wall β–glucosidases were submitted to NetNGlyc 1.0 Server. (http://www.cbs.dtu.dk/services/NetNGlyc/).

Results

Subcloning of the CWR gene

The CWR gene was amplified from the sequencing vector D4 using PCR. The expected size of the target product was 1818 bp. Two different buffers, GC and HF, included in the PCR-kit (Finnzymes) were tested. Figure 1 shows the products from the amplification, with different dilutions, 1x, 5x and 10x. Both buffers produced products of expected size. The same result was obtained for the thioredoxin CWR (results not shown). Upscale for production of much PCR-product was difficult using the HF buffer, therefore GC buffer was chosen for the small upscale of volume.
The same result was obtained for the “thio”-CWR (results not shown).

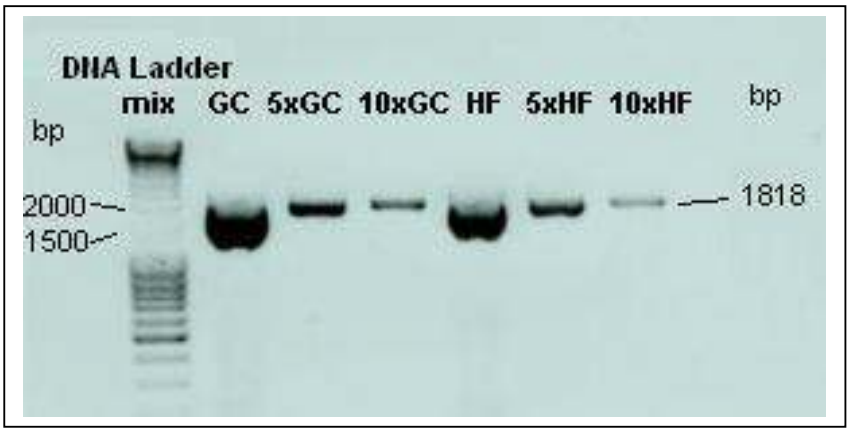


Figure 1: Agarose gel showing PCR-products from the amplification of the CWR-gene in D4 vector. Expected size of target product is 1818 bp. PCR-buffers GC and HF, provided in the kit, were tested. The PCR-products were diluted 5 and 10 times for easier visualization. Ladder: MassRuler^{TM} DNA Ladder mix (Fermentas).

Transformation of competent *E.coli* strains M15[pREP4] and BL21(DE3)

Competent *E.coli* strain M15 [pREP4] was transformed with the ligation mix of pQE-1 vector and PCR-product. Positive control was prepared with the uncut pQE-1 vector, and the negative control was prepared without any plasmid. 200 µL and 50 µL of the transformation mix was accomplished with the ligation mix of pQE-1 vector and PCR-product. The figure 2 shows that for the positive control there was so many colonies that they were unable to be counted. The negative control did not contain any colonies. The plates with the transformed bacteria had colonies. There was a ~5x ratio of colony number between the plates with 200 µL and 50 µL transformation mix. So it was possible to continue with these bacteria. There were similar results for the thioredoxin system. The positive control and the plates with the transformed bacteria had colonies.

	Colonies
Positive control	~10000
Negative control	0
Ligation (200µl)	~4000
Ligation (50µl)	~1000

Figure 2: Table showing bacteria growing on the LB agar containing kanamycin (25 µg/ml) and ampicillin (100 µg/ml). The uncut pQE-1 vector was used as a positive control and the negative control transformation was accomplished without any

plasmid. Two different volumes (200 µL and 50 µL) of competent *E.coli* strain M15 transformed with the ligation mix of pQE-1 vector and PCR-product were plated. Similar results were observed for the thioredoxin system.

Direct PCR on colonies

In order to find out if the CWR insert is present in the pQE-1 plasmid, 5 differents colonies were picked and used as templates in PCR. The expected size of PCR-products is 1818 bp. The same procedure was carried out for the thioredoxin system. The expected size of PCR-products (CWR insert + pET102-D/TOPO plasmid) was 984bp. The figures 3a and 3b show respectively colonies a1, a2, a3, a4, and a5 of pQE-1-CWR, and colonies 1, 2, 3, 4 and 5 of pET102-D/TOPO.
There were bands of expected sizes in colonies a1, a2, a3, a5, and 1, 2, 3, 4, 5.

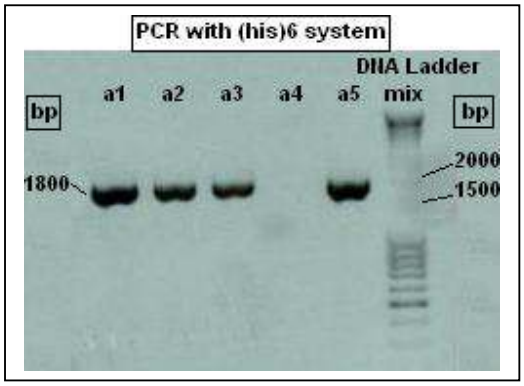


Figure 3a: Agarose gel showing if the insert, CWR, is present in the plasmid, pQE-1. 5 different colonies (a1, a2, a3, a4, and a5) were picked and used as templates in PCR. The expected size of PCR-products is 1818 bp. There is a band of expected size in a1, a2, a3 and a5.
Ladder: MassRuler^{TM} DNA Ladder mix (Fermentas)

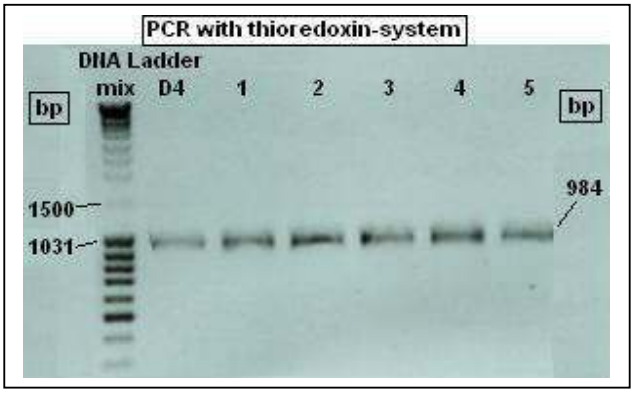


Figure 3b: Agarose gel showing if the insert, CWR, is present in the plasmid, pET102-D/TOPO. 5 differents colonies (1, 2, 3, 4, and 5) were picked and used as templates in PCR. The expected size of PCR-products was 984 bp. There is a band of expected size in colony 1, 2, 3, 4 and 5.
A positiv control (D4) was made with D4 plasmid. Ladder: MassRuler^{TM} DNA Ladder, mix (Fermentas)

Test expression

In order to find out if transformed colonies could produce the β-glucosidase, a test expression was made. Same colonies as shown in figures 3a and 3b were picked. Figure 4 shows the boiled pellet from induced *E.coli* strain M15, transformed with pQE-1 + CWR insert, without IPTG (“c”) or with 0,5 mM IPTG (“i”). The expected size of the protein was ~70 kDa since the size of the β-glucosidase is 68-69 kDa and the

tag is ~1 kDa. There is a strong band for each induced cultured colony, closely to 70 kDa.
For the test expression of pET102-D/TOPO-CWR1 see solubility test, Figure 6. There is a strong band in the induced cultured colony (IC), closely to 80 kDa. The size of the thioredoxin tag is about 13 kDa.

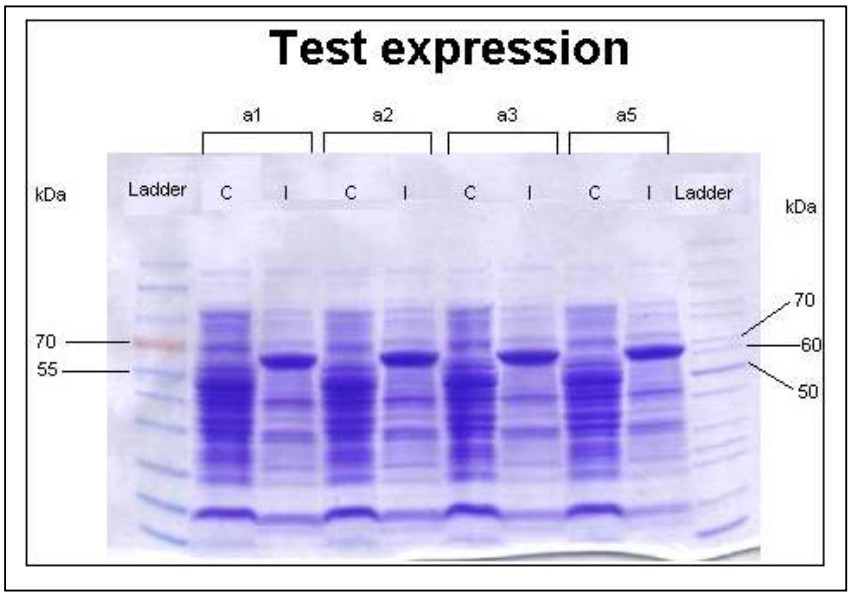


Figure 4: SDS-PAGE showing induced cultures of colonies a1, a2, a3, and a5 for the $(his)_6$-system. This figure shows the boiled pellet from induced *E.coli* strain M15, transformed with pQE-1 + CWR insert, without IPTG ("c") or with 0,5 mM IPTG ("i"). The expected size of the protein is ~70 kDa. There is a clear band at the expected size for all the induced colonies.
Ladder on the left: PageRuler^{TM} Prestained Protein Ladder (New England biolabs)
Ladder on the right: PageRuler^{TM} Protein Ladder (New England biolabs)

Restriction digestion of pQE1-CWRa1

In order to further confirm that the plasmid (pQE-1-CWRa1) contained the insert, a double digestion with EcoRI and Kpn1 was made. This figure shows the plasmid digestion of the colony a1 from the $(His)_6$ system. The expected sizes were 3381 bp, 1398 bp, and 469 bp and there are clear bands at 3,4 kb, 1,4 kb and just below 500 bp.

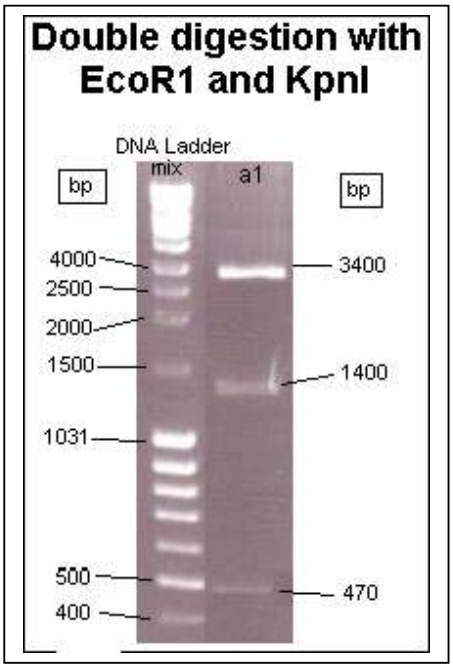


Figure 5: Agarose gel showing if the plasmid has the CWR insert. In this figure, one can see the restriction digestion of plasmids extracted from colony a1 with EcoRI and KpnI. The expected sizes are 3381 bp, 1398 bp, and 469 bp and there are clear bands at 3,4 kb, 1,4 kb and just below 500 bp. Ladder: MassRuler^{TM} DNA Ladder mix (Fermentas)

Sequencing

The results from the sequencing showed that the gene was successfully cloned with no mismatches. The clones pQE-1-CWRa1 and pET102-D/TOPO-CWR1 were chosen to be continued with. However, for the pET102-D/TOPO-CWR1 there was a ~150 bp part of the gene that had not been sequenced.

Expression

Determination of β-glucosidase solubility

The determination of the β-glucosidase solubility is an important step to characterize a protein. So to see if a protein is soluble or not, it could be induced by IPTG, and run on a SDS-PAGE separating the supernatant, containing soluble proteins, and the pellet, containing insoluble proteins.
The solubility of the two recombinant fusion proteins, the $(his)_6$-tagged β-glucosidase and the thioredoxin tagged β-glucosidase, were tested. Colonies a1 and 1 were used respectively. The expected size of the protein was ~70 kDa and ~80 kDa respectively. Figure 6 shows the negative control ("C") which is an uninduced sample and the positive control ("IC") which is the induced sample with 0,5 mM IPTG. The supernatant is the lysate after lysis of the cells and the pellet is the insoluble material after centrifugation. For a1 it was difficult to detect a band in the supernatant, because of

the dilution but it´s clear that there is a big band in the pelleted insoluble material. It means β-glucosidase is not soluble. There was not any success with the other induction times and temperatures. Consequently, the His-tagged system was not a good way to get soluble β-glucosidase.

One the other hand, the supernatant from the colony 1 for the thioredoxin system contained a band at the expected size. The induced control indicates that the amount of induced protein makes up approximately 50% of the total proteins. For the soluble protein fraction the part of induced protein is much smaller, approximately 5%. This culture had been induced at 37°C for 2,5 h. Other induction times and temperatures did not give a higher amount of soluble recombinant protein.

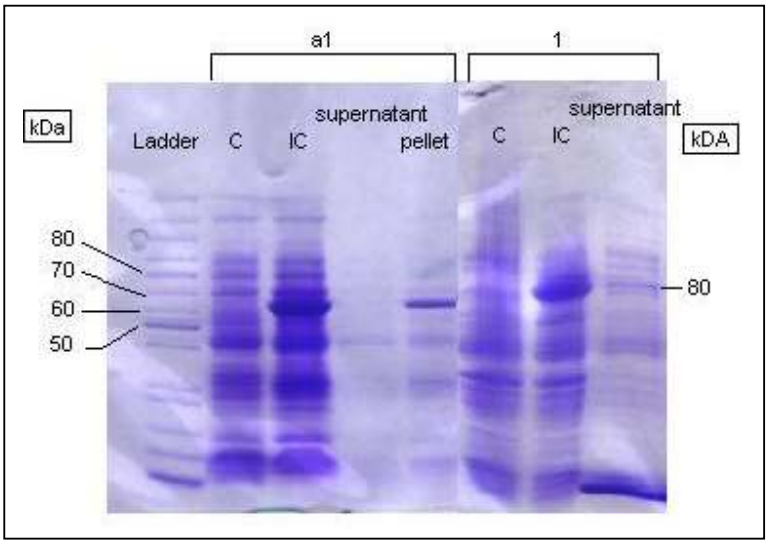


Figure 6: SDS-PAGE showing the solubility of the $(his)_6$–tagged and the thioredoxin tagged β-glucosidase, from a colony a1 and 1 respectively. The expected size of the protein was ~70 and ~80 kDa respectively. The negative control ("C") is noninduced sample and the positive control ("IC") is the induced sample with 0,5 mM IPTG for 4 hours for colony a1 and 2,5 hours for colony 1. The figure shows the supernatant and the pelleted material from induced *E.coli* with 0,5 mM IPTG. Ladder: PageRuler™ Protein Ladder (New England Biolabs).

To confirm the results of a soluble protein fraction for the thio-fusion protein the supernatant fraction from Figure 6 was subjected to Western Blot using anti-thioredoxin anti-bodies. The expected size of the protein was ~80 kDa. The figure 7 shows the supernatant from induced *E.coli* strain BL21, transformed with pET102D/TOPO-CWR1, with 0,5 mM IPTG. The culture was induced for 2,5 h at 37°C. The supernatant from the colony 1 contained a band at the expected size.

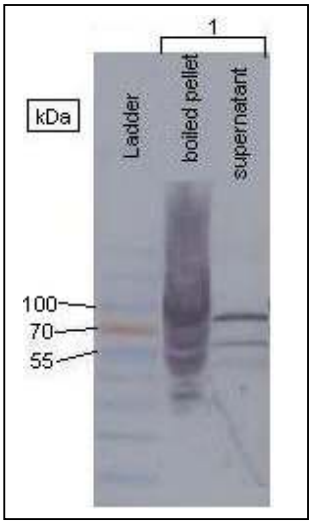


Figure 7: Western blot showing the β-glucosidase solubility. The expected size of the protein is ~80 kDa. The positive control ("boiled pellet") is the induced sample with 0,5 mM IPTG for 2,5 hours. This figure shows the supernatant from induced *E.coli* strain BL21, transformed with pET102D/TOPO CWR insert, with 0,5 mM IPTG.
Ladder: PageRuler^{TM} Prestained Protein Ladder (New England biolabs).

To confirm that it actually was not just a contamination from the pellet in the supernatant another Western Blot was carried out. At the same time, a time course was made to see that it was a gradual induction. The expected size of the protein is ~80 kDa. The figure 8 shows the supernatant and the pellet protein from induced *E.coli* strain BL21, transformed with pET102D/TOPO-CWR1. The uninduced sample ("nonIC") is the negative control and the induced samples ("boiled pellet") are the positive controls. An 80 kDa band is present in the induced sample and in the pellet. Moreover, the same band is present in the supernatant samples. As the time of induction increases, the size of the band in the supernatant increases.

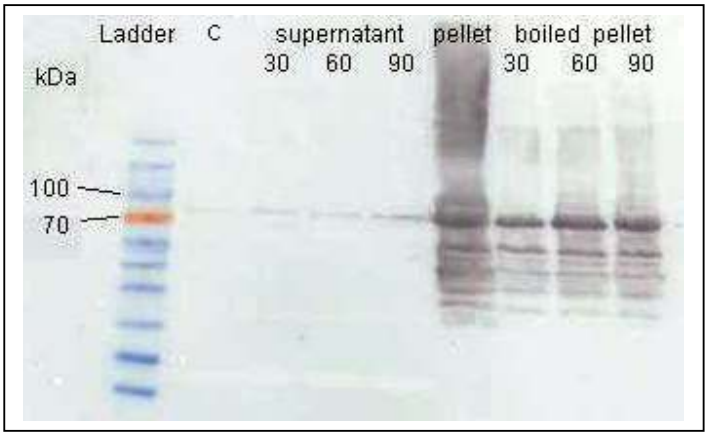


Figure 8: Western blot, with anti-thioredoxin tag antibody, showing the solubility of the β-glucosidase at different times of induction. As the time of induction increases, the size of the band in the supernatant increases.

From left to right: PageRuler™ Prestained Protein Ladder, uninduced control "c", supernatant 30 minutes of induction, supernatant 60 minutes of induction, supernatant 90 minutes of induction, pellet protein from induced *E.coli* strain BL21, boiled pellet 30 minutes of induction, boiled pellet 60 minutes of induction, boiled pellet 90 minutes of induction.

Purification of the β-glucosidase

After the large scale expression, the β-glucosidase was purified with a Ni-column. The expected size was 80 kDa. The figures 9a and 9b show pellet protein from induced *E.coli* strain BL21, transformed with pET102D/TOPO-CWR1, the flow-through with all proteins, the two washing fractions, and the 4 elutions (E1, E2, E3, and E4). A lot of bands were present in eluate 1 and 2. There was one band in these samples at 80 kDa. The 4xE2 sample is an eluate from the column loaded with four times more sample than the other E2. There was an 80 kDa band with a quite same thickness in the 4xE2 sample.

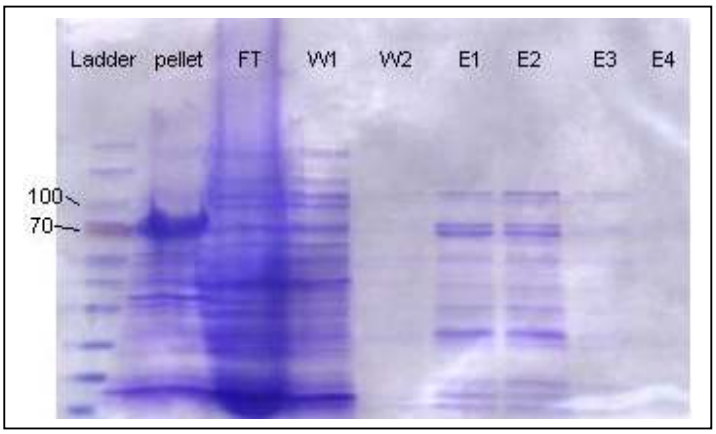


Figure 9a: SDS PAGE showing the β-glucosidase purification with Ni-column. From left to right: PageRuler™ Prestained Protein Ladder, pelleted protein from induced *E.coli* strain BL21, Flow-through (FT), twice wash (W1 & W2), 4 elutions (E1, E2, E3 & E4).

Figure 9b: Western blot, with anti-thioredoxin tag antibody, showing the β-glucosidase purification with Ni-column.
From left to right: PageRuler™ Prestained Protein Ladder, pellet protein from induced *E.coli* strain BL21, Flow-through (FT), twice wash (W1 & W2), 4 elutions (E1, E2, E3 & E4), one elution with 4 times more of supernatant.

pNPG assay

The eluates E1 and E2 from the 250 mL and 1 L culture were subjected to pNPG-assay. There was no detected activity in E1 and E2 samples (results not shown). The activity was also measured in crude lysates from uninduced and induced bacteria. The uninduced sample showed an activity of 3 nmol/(mg, min) and the induced sample showed an activity of 9 nmol/(mg, min). The positive control showed a high activity.

Bioinformatics

In order to know which enzymes that were similar to the cell wall β-glucosidase, the NCBI Blast software was used. This table shows the top hits of the amino acid similarity of the cell wall β-glucosidase. The sequence alignment scores come from alignment with the cell wall β-glucosidase using Clustal W.

Name protein	reference	species	score
β-D-glucan exohydrolase	AAM13694	Wheat	97
β-D-glucan exohydrolase isoenzyme ExoII	AAC49170	Barley	96
Os03g0749300	NP_001051275	Rice	86
Exo-β-D-glucanase	AAF79936	Maize	85
β-D-glucosidase	AAQ17461	Cotton Seeds	77
β-D-glucosidase	CAA07070	Indian Cress	75
β-D-glucan exohydrolase isoenzyme ExoI	AAD23382	Barley	70

Table 2 showing the most similar enzymes to the cell wall β-glucosidase based on amino acid sequences.

In order to investigate putative glycosylation the amino acid sequences of the plastidic and the cell wall β–glucosidases were submitted to NetNGlyc 1.0 Server. This prediction program found that there were four putative glycosylation sites for the cell wall β-glucosidase. For the plastidic the probability of any glycosylation was very low since the protein did not contain a signal peptide, even though it contained sites.

Discussion

There were problems when trying to get a lot of PCR-product for cloning with the HF-buffer. The error rate of Phusion High-Fidelity DNA Polymerase in HF Buffer is lower than that in GC Buffer. Therefore, the HF buffer should be used as a default for high fidelity amplification. However, GC Buffer can improve the performance of Phusion DNA Polymerase. Therefore, GC buffer was chosen for the small upscale of volume.

There was no problem with using the HF-buffer for the thioredoxin primers. The band sizes which were observed on the gels all had the expected sizes. Though, when there is a lot of product loaded onto a gel it tends to give the impression of a lower molecular weight.

After transformation of *E.coli* strains with the two plasmids, the analysis was made with different methods.
PCR on colonies directly (figure 3a and 3b), test expression (figure 4 and 6), the double digestion with EcoR1 and KpnI (figure 5) and the sequencing showed that the transformation had worked. The small part, ~150 bp in the pET102-D/TOPO-CWR1, that was not sequenced will be sequenced if the enzyme shows to be active.
The CWR had been nicely ligated into each vector. Only the colony a4 did not work because sometimes the plasmid closes without the gene. That is why the bacteria had grown on the medium containing antibiotics even though there was no insert present. It is still strange though, because the open plasmid had a sticky and a blunt end so theoretically it would not close. All this experience was an indication also that the preparation of competent *E.coli* was successful.

The previous attempt to express the β-glucosidase using the IMPACT-CN (Intein Mediated Purification with an Affinity Chitin-binding Tag) had not been successful. One possible explanation was that this was maybe due to a too big tag, which interfered or made it difficult for the bacteria to fold the protein properly and therefore produced inclusion bodies. For this reason a construct with a smaller tag, a $(his)_6$-tag was chosen. The figure 6 shows that the β-glucosidase was not soluble when fused to the $(his)_6$-tag. So, it seems that there was not a matter of size.

The solubility of the thioredoxin tagged protein was also tested. Several other proteins that have had problems with solubility had become soluble when fused to thioredoxin (Yasukawa, 1995). Indeed the thioredoxin tag helped the protein become soluble. Nevertheless, this was a very small fraction. The induced control indicates that the amount of induced protein makes up approximately 50% of the total protein. For the soluble protein fraction the part of induced protein is much smaller, approximately 5%.

The fact that so much protein tend to end up in the insoluble pelleted fraction indicates formation of inclusion bodies. This is not a very uncommon result when trying to express a recombinant protein in *E.coli*. Indeed, *E.*coli has a frequent deposition of the expressed protein product into insoluble inclusion bodies (Cabrita and Bottomley 2004). For instance, a maize β-glucosidase was overexpressed in *E.coli*, resulting in accumulation of most of the protein in insoluble inclusion bodies (Jan Zouhar et al. 1999).

The figure 7 confirms this conclusion. Indeed, the supernatant from the colony 1 contained a band at 80 kDa. There are two other bands at ~55 kDa and ~40 kDa. It means the primary antibody, which was used in this Western Blot, was not really specific for the thioredoxin tag. One possibility could also be that there were degradation products form the β-glucosidase fusion protein.

In the figure 8, contrary to the His-tagged system, one 80kDa band is present in the supernatant samples. This figure shows us as the time of induction increases, the size of the band in the supernatant increases. It means as the time of induction increase, the cells produce a protein.

The figure 9a shows there were a lot of bands present in eluates 1 and 2. The figure 9b shows that there were four other bands, though a lot weaker. It means the Ni-column was not so specific to bind the thioredoxin tag. There was a 80 kDa band with quite the same thickness, on judged by the eyes in the 4xE2 sample. It means when the eluate quantity increase, the β-glucosidase quantity after purification did not increase. This was not expected.

The results of the pNPG assay did not show any success. It was probably due to a weak concentration of the β-glucosidase in eluate samples. For some eluates it was a bit unclear whether there was activity or not since it was so low. On top of this, the eluates were not pure so there could be a risk that it were co-purified *E.coli* proteins that were responsible for the activity. One way to solve this could be to make a native gel and make an activity stain directly on it. The results shows a very small difference between uninduced and induced sample (supernatant): 3 nmol/(mg, min) and 9 nmol/(mg, min) respectively. Of course, there was activity, which came from the *E.coli* enzymes but it is too diluted. Another possibility for the undetectable activity was that the protein was not active when it was produced in *E.coli* at all. It was not surprising that the positive control gave such a huge activity since the sample was pure.

A BLAST search was performed to investigate the most similar proteins to the cell wall β-glucosidase with respect to possible glycosylation (table 2). The partly purified cell wall β-glucosidase glycosylation had been analysed and there was not any proof of glycosylation (Nikus and Jonsson 2003). As Nikus and Jonsson argues, this was very surprising since it is very common for secreted proteins, like the cell wall β-glucosidase, to be glycosylated (Cairns et al. 2000, Varghese et al.1999). The β-D-glucosidase from Indian Cress (table 2) had been sequenced, and the N-glycosylation sites determinated. The primary amino acid sequence contains four putative N-glycosylation sites (Crombie et al. 1998). As well, exo-β-D-glucanase also called ExGase (table 2) which was purified from the cell walls of developing maize (*Zea mays*) shares five putative glycosylation sites with the cell wall β-glucosidase in rye (Kim et al. 2000). The same observation had been observed with the β-D-glucan exohydrolase exo II (table 2). This enzyme had four N-glycosylation sites and there was also some evidence for actual glycosylation (Hrmova et al. 1995).
β-D-Glucan Exohydrolase, Exo I, in barley (table2) which also belongs to family 3 and share the catalytic sites with the cell wall β-glucosidase is glycosylated at three sites according to the three-dimensional structure (Varghese et al. 1999). Two of the corresponding sites are found in the cell wall β-glucosidase. So, if the enzymes that show a big sequence similarity to the cell wall β-glucosidase are glycosylated, there is a big probability that the cell-wall β-glucosidase is also glycosylated.

The N-glycosylation sometimes has a role in the biological activity of the protein. Indeed the N-glycosylation influence the glycoprotein conformation, stability and biological activity for some proteins (Rayon et al. 1998). For the cell wall β–glucosidase there is not enough information from the literature whether the actual glycosylation is important for the activity or not. β-D-Glucan Exohydrolase, Exo I, in barley has three glycosylation sites where one is discussed to have a potential role in guiding of the substrate (Varghese et al 1999). It is not very similar to our β-glucosidase but is similar in regions of the active site and the glycosylation sites.
The importance of glycosylation for the activity may not be so big, it is more likely that it plays a bigger role in folding the protein correctly. This might explain the production of inclusion bodies. In this project, three systems for cloning, expression and purification of the cell wall β-glucosidase have been tried. All of them were based on expression in *E.coli*. The plastidic β-glucosidase was successfully cloned and expressed in *E.coli* (Nikus et al. 2003). The NetNGlyc prediction program suggested that there was no glycosylation since the protein does not contain a signal peptide and therefore it is not very likely that it is exposed to N-glycosylation. *E.coli* cannot make many of the modifications that eukaryotic proteins, like plant proteins, need. Using yeast for expression could be used if protein needs post-translational modifications, like glycosylation.

Three *E.coli* based systems have been tried for cloning, expression and purification of the cell wall β-glucosidase. In this ex-job on one hand *E.coli* produced a weak fraction of soluble β-glucosidase when co-expressed with a thioredoxin tag. One the other hand, *E.coli* probably did not produce a β-glucosidase which was active. Production of a soluble and active β-glucosidase with *E.coli* was maybe not a good way. A good idea is to move into a higher system, like yeast or insect cells, since they have a higher possibility to make posttranslational modifications if needed.

Acknowledgments

This work was made possible by Södertörns Högskola, Flemingsberg, Sweden and I am thankful for being accepted as a student there.
I would like to thank Professor Lisbeth Jonsson for accepting me for this project.
I would like to thank MSc Therese Gradin for her patiently help with the lab work, for the help with the report, for answering my questions about details of lab work and also improve my English skills.

References

Pictures on the front page: http://www.cdl.umn.edu/introduction/crops_images/rye.gif

Asim E (1993) β-glucosidase, biochemistry and Molecular Biology. (Manuscript)

Cabrita LD and Bottomley SP (2004). Protein expression and refolding – a practical guide to getting the most out of inclusion bodies. Biotechnol Annu Rev, 10:31-50.

Cairns K, Champattanachai V, Srisomsap C, Wittmann-Liebold B, Thied B, and Svati J (2000). Sequence and expression of thai rosewood β-glucosidase/β-fucosidase, a family 1 glycosyl hydrolase glycoprotein. J. Biochem. 128:999-1008

Crombie HJ , Chengappa S, Hellyer A, Reid JSG (1998). A xyloglucan oligosaccharide-active, transglycosylating-D-glucosidase from the cotyledons of nasturtium (Tropaeolum majus L) seedlings – purification, properties and characterization of a cDNA clone. The Plant J 15 (1), 27–38.

Dharmawardhana DP, Ellis BE, Carlson JE (1999). cDNA cloning and heterologous expression of coniferin β-glucosidase. Plant Mol Biol 40:365-372

Dietz K-J, Sauter A, Wichert K, Messdaghi, Hartung W (2000). Extracellular β-glucosidase activity in Barley involved in the Hydrolysis of ABA glucose conjugate in leaves. J Exp Bot 51:937-944.

Gupta R, Jung E, Brunak S (2004). Prediction of N-glycosylation sites in human proteins. In preparation

Henrissat B, Davies GJ (1997) Structural and sequence-based classification of glycoside hydrolases. Curr. Op. Struct. Biol. 7:637-644 [PMID: 9345621]. (http://www.cazy.org/)

Hrmova M, Harvey AJ, Wang J, Shirley NJ, Jones GP, Stone BA, Hoj PB, and Fincher GB. (1996). Barley β-D-Glucan exohydrolase with β-D-glucosidase activity. J Biol Chem, 271: 5277-5286

Kim JB, Olek AT, and Carpita NC (2000). Cell Wall and Membrane-Associated Exo-β-d-Glucanases from Developing Maize Seedlings. Plant Physiol. 123(2): 471–486.

Niemeyer H. (1988) Hydroxamic acids (4-hydroxy-1,4-benzoxazin-3-ones), defence chemicals in the gramineae. Phytochemistry, 27: 3358-1988

Nikus J, Daniel G, Jonsson LMV (2001). Subcellular localization of β-glucosidase in rye, maize and wheat seedlings. Physiol Plant 111:466-472

Nikus J, Esen A, Jonsson LMV (2003) Cloning of a plastidic rye (*Secale cereale*) β-glucosidase cDNA and its expression in *E.coli*. (Manuscript)

Nikus J, Jonsson LMV (2003). Cell wall glucan 1,3- β-glucosidase from rye (Secale cereale) active on hydroxamic acid glucosides. (Manuscript)

Poulton JE (1990). Cyanogenesis in Plants. Plant Physiol 94 :401-405

Rayon C, Lerouge P and Faye L (1998). The protein N-glycosylation in plants. J Exp bot, 49:1463-1472

Roos A (2006) Rye cell wall β-glucosidase : purification of recombinant protein from *E.coli* and molecular and kinetic characterization of the enzyme. (Student report)

Varghese JN, Hrmova M, Fincher GB (1999). Three-dimensional structure of a barley β-D-glucan exohydrolase, a family 3 glycosyl hydrolase. Structure 15;7(2):179-90

Yasukawa T, Kanei-ishii S, Maekawa T, Fujimoto J, Yamamoto T, Ishii S (1995). Increase of solubility of foreign protreins in *E.Coli* by copropduction of the bacterial Thioredoxin. J Biol Chem. 270:25328-25331.

Zouhar J, Nanak E and Brzobohaty B (1999). Expression, single-step purification, and matrix assisted refolding of a maize cytokinin glucoside-specific β-glucosidase. Protein Expr Purif 17:153-162.