Expression of a thermophilic beta-glucosidase from *Thermus sp. Z-1* in *E.coli* and *Pichia pastoris*

BACHELORARBEIT 2

In order to obtain the academic degree
„Bachelor of Science in Engineering“

Course of study:
„Engineering, Environmental & Biotechnology**“**
Management Center Innsbruck

Assessor:
Dr. Christoph Griesbeck

Author:
Stefan Farnik
0710351025

Blocking notice

I have applied for the blocking of this bachelor thesis which has been accepted by the members of the examining board of the Mangement Center Innsbruck.

Place, Day / Month / Year _____

Declaration in lieu of oath

I hereby declare, under oath, that this bachelor thesis has been my independent work and has not been aided with any prohibited means. I declare, to the best of my knowledge and belief, that all passages taken from published and unpublished sources or documents have been reproduced whether as original, slightly changed or in thought, have been mentioned as such at the corresponding places of the thesis, by citation, where the extent of the original quotes is indicated.

The paper has not been submitted for evaluation to another examination authority or has been published in this form or another.

Place, Day / Month / Year _____

Acknowledgements

I thank the Austrian Marshall Plan Foundation for being one of their holders of the Marshall Plan scholarship. Without these grants it would have been almost impossible to reside nine months at the University of California, Berkeley. My thanks also go to supervisors at UC Berkeley and members of the Energy Biosciences Institute. They helped me in gaining lab work experience and were and still are creators of my personal enrichment. During my stay in Berkeley I've always been in contact with my university MCI Innsbruck. They also form a big part of achieving the completion of my bachelor thesis. Thank you very much for accompanying me on this way. I send my final thanks to all my supporters including family, friends and colleagues.

Inhaltsverzeichnis

Abstract

A thermostable cellulase mixture is useful for high temperature, contamination-minimized hydrolysis of plant material. Endo-, exoglucanase and beta-glucosidase are coordinated in one degradation tank to create the important substrate *glucose* for organism that produce ethanol/butanol. The glucanases shorten the cellulose chains into cellobiose molecules while the beta-glucosidase does the final cut at the ß-1,4 bond linking the two glucose units.

This project lays focus on a thermostable beta-glucosidase from *Thermus* sp. Z-1, which was found in the soil of the Yukanada hot springs in Japan. It endures temperatures up to 90°C over a pH range from pH 4.0 to 7.0. The gene was synthesized at an extern facility and delivered on a His-Tag containing plasmid. Two popular host systems for the heterologous expression of the enzyme were chosen, *E.coli* and *Pichia pastoris*.

In detail, the *E.coli* hosts were OrigamiB and BL21. The *P. pastoris* strains were the wildtype X33 and the modified strain pichiaPink IV. The plasmids for each host were completed until an error-free cloned emerged out of all colonies. The expression hosts were then transformed with the appropriate plasmid containing the gene of interest. The main part of the project comprehends of the expression analysis in each host. During the research the expression methods had been modified to increase the yield and solubility of the protein. The enzyme activity was tested after each change. SDS-PAGE and Western Blot analysis were implemented to get a visual evidence of the enzyme.

The BL21 system was upgraded with a chaperonin expressing host that was expected to increase the solubility and so the overall activity of the enzyme. In another experiment the aggregated proteins known as inclusion bodies were tried to be refolded.

The purification was run on nickel-beads that have a high affinity to his-tagged proteins. Nature and denaturing conditions were applied to purify the beta-glucosidase.

To conclude, the thermostable glucosidase from *Thermus* sp. Z-1 couldn't be expressed in substantial amounts. The yield was too low and the his-tagged enzyme was difficult to purify.

Kurzfassung

Eine hitzestabile Mischung aus Zellulasen ist sehr nützlich bei der Hydrolyse von Pflanzenmaterial unter hohen Temperaturen und niedriger Kontaminationsrate. Endo-, Exoglucanasen und die ß-glucosidase werden in einem Zersetzungstank vereint, um an die wichtige Komponente Glucose zu kommen, welches von den Organismen zur Produktion vom Ethanol/Butanol als Nahrung verwendet wird. Die Glucanasen zerkleinern die Zelluloseketten in kurze Zellobiosemoleküle, welche von der ß-glucosidase an der ß-1,4 Verknüpfung in zwei Glucoseinheiten geteilt wird.

Dieses Projekt legt sein Augenmerkmal auf eine hitzestabile ß-glucosidase von Thermus sp. Z-1, welches in der Erde an den heißen Quellen von Yukanada in Japan gefunden wurde. Seine Temperaturstabilität reicht bis 90°C über einen pH Bereich von 4,0 bis 7,0. Das Gen wurde von einer externen Einrichtung synthetisiert und mit His-Tag auf einem Plasmid geliefert. Zwei populäre Wirte wurden für die heterologe Expression gewählt: *E.coli* und *Pichia pastoris*.

Im Detail wurden OrigamiB und BL21 als *E.coli*, der Wildtyp X33 und der modifizierte Stamm pichiaPink IV als *P. pastoris* verwendet. Die Plasmide aller Wirte wurde vervollständigt bis ein fehlerfreier Klon aus den Kolonien isoliert wurde. Die Expressionswirte wurde mit diesen transformiert und bereitgestellt. Den Hauptteil dieses Projekt nimmt die Analyse der Expression jedes Wirts ein. Während der Forschung wurden die Expressionsmethoden modifiziert, um die Ausbeute und die Löslichkeit der Proteine zu erhöhen. Die enzymatische Aktivität wurde nach jeder Änderung getestet. SDS-PAGE und Western Blot Verfahren stellten eine visuelle Bestätigung der Enzymproduktion dar.

Das BL21-System wurde durch die Expression von Chaperones verbessert, welche die Löslichkeit der ß-glucosidase erhöhen sollten. In einem anderen Experiment wurde versucht, die unlöslichen Proteine, bekannt als Einschlusskörper, neu zu formen.

Die Reinigung wurde auf Nickelperlen durchgeführt, die eine hohe Affinität zu His-verknüpften Proteinen aufweisen. Natürliche und denaturierende Umgebungsbedingungen wurden für die Reinigung der beta-glucosidase angewandt.

Zusammengefasst lässt sich sagen, dass die hitzestabile beta-glucosidase von Thermus sp. Z-1 nicht in ausreichenden Mengen exprimiert wurde. Die Ausbeute war meist zu gering und das His-verknüpfte Enzym ließ sich nur schwer reinigen.

1. Introduction

Renewable resources control the future of our planet. It is from essential value that we know how make it, use it and renew the whole process. Research is being carried out in many fields nowadays. One of those is the degradation of cellulose into single glucose units [1]. Different enzymes work in collaboration to create the food for the organism that will produce ethanol/butanol. The working group comprehends the endo-, exoglucanase and the beta-glucosidase. In this research the focus is laid on a thermophilic beta-glucosidase from the bacteria *Thermus* sp. Z-1.

1.1. The thermophilic beta-glucosidase from *Thermus* sp. Z-1

The beta-glucosidase (β-D-glucoside glucohydrolase, EC 3.2.1.21) catalyzes the hydrolytic cleavage of ß-1,4 bonds linking two glucose units or glucose-substituted molecules. They have been reported in all three members of living organism (Bacteria, Archaea and Eucarya).
The thermophilic beta-glucosidase from Thermus sp. Z-1 has been found by Takase, M. and Horikoshi, K. Enzymatic property studies have been carried out by the same researchers [2] [3]. Due to its high optimum temperature of 70-80°C the enzyme – found under GenBank: AB034947.1 (see attachment pages) – can be used in high temperature fermentation processes. The chance of contaminants is lower the higher the temperature.
This beta-glucosidase from *Thermus* sp. Z-1 was also chosen because it was heterologously expressed in *E.coli* (K-12) in substantial amounts by et. al Akiyama [4].

1.2. Host systems

Two different host systems are mainly used in the recombinant expression of proteins: Bacteria and Eucarya. The selection was narrowed down to *E.coli* and *P. pastoris* two high efficient expression hosts.

1.2.1. *E.coli*

As the selected beta-glucosidase is a prokaryotic gene it is applicable for *E.coli*. They generate high amounts of protein in a fast speed. They are rather easy to handle and therefore make a very simple expression host.
OrigamiB carry the trxB/gor mutations which greatly enhances the disulfide bond formation. Although the z1-ßgl (beta-glucosidase from *Thermus* sp. Z-1) is supposed to just have one disulfide bond stated by an internet protein folding program it increases the chance of a correct folding. It is additionally derived from the lacZY mutant BL21 so it can be induced by IPTG.
The BL21(DE3) contains the T7 polymerase upon IPTG induction and is the basic expression host from Novagen. The Star model is especially designed for the translating toxic genes.

1.2.2. *Pichia pastoris*

Pichia pastoris is a popular yeast used for an exceptional spectrum of heterologous proteins which are mostly eukaryotic [4]. Several beta-glucosidases have been successfully expressed and the yield was substantially high [5] [6].

One big advantage in P. pastoris is the secretion of recombinant proteins with a secretion signal attached. So less purification steps are needed to isolate the target protein. *P. pastoris* is a methylotrophic yeast and therefore has an alcohol oxidase as its promoter which drives the protein expression.

The X33 wildtype strain and the PichiaPink strain IV were chosen for expression. The X33 is a non-manipulated strain and was selected with the antibiotic zeocin. On the contrary, the PichiaPink IV strain uses the complementation of adenine auxotrophy, which makes the colonies turn white when they contain the ADE2 gene. In addition to that, the strain IV has two proteinase knockouts (*prb1* and *pep4)*.

1.3. Chaperones

Chaperones are heat shock proteins that protect proteins under cellular stress and increase the chance of correct folding. The one that was used in this project was called GroEL/ES, two proteins synergistically linked to each other [7]. Et. al Machida has successfully expressed beta-glucosidase in active form in *E.coli* by using chaperones [8].

This chaperonin was being co-expressed within the BL21Star(DE3) host to increase the solubilizing rate of the recombinant enzyme z1-ßgl.

1.4. Goal of the project

The main goal of this project was to express the thermostable beta-glucosidase from *Thermus* sp. Z-1 in *E.coli* and *P. pastoris*, compare the enzymatic activity and determine the values for K_m and k_{kat}. Is the better system located optimizations of that systems are initiated and the yield increased. This enzyme would have been used in the research laboratory for kinetic substrate studies.

2. Materials and Methods

2.1. Materials

Various manufacturers were chosen to order the materials. All companies are settled in the USA, but have delivery points worldwide.
NEB means New England Biosciences and Novagen can be found under EMD Chemicals.

2.1.1. Plasmids

The plasmid with the beta-glucosidase gen on it was gen-optimized for *E.coli* and created by Genscript. It was named z1-ßgl. The OptimumGene™ algorithm was used to improve the gene's parameters like codon usage bias, GC content, restriction sites, repeating sequences and more.
The bacterial expression plasmids were the Champion™ pet303/CT-His (#K6302-03, Invitrogen) and the pet28b/NT-His (#69865-3, Novagen). The eukaryotic expression plasmids were ordered within expression kits: pPICZαC (#K1740-01, Invitrogen) and pPinkα-HC (#A11153, Invitrogen).
The test plasmid to evaluate transformation quality was the pUC19.

2.1.2. Hosts

The bacterial expression systems were the One Shot® BL21(DE3) (#C600003, Invitrogen), BL21Star™ (DE3) (#C602003, Invitrogen), BL21Star™+GROEL/ES (DE3) co-expression host and OrigamiB(DE3) (#70837-4, Novagen).
The eukaryotic expression systems were ordered in the kit-form: One was the EasySelect™ Pichia Expression Kit from Invitrogen (#K1740-01) where the X33 pichia wildtype, phenotype Mut+, was used. The other expression kit was the PichiaPink™ Expression System from Invitrogen (#A11150 and #A11153) where strain 4 (knockout of two major proteases) was mainly used.
For plasmid propagation the One Shot® TOP10 Chemically Competent *E.coli* (#C4040-10, Invitrogen), the One Shot® TOP10 Electrocomp™ *E.coli* (#C4040-52, Invitrogen) and the chemically competent in-house Mach1 cells were transformed.

2.1.3. Solutions and substances

Antibiotic stocks were all 1000X: Carbenicillin (100mg/ml), ampicillin (100mg/ml), zeocin (25mg/ml), kanamycin (60mg/ml) and CAP (chloramphenicol) (34mg/ml).
Further solutions were 1M sopropyl-ß-d-thiogalactopyranoside (IPTG) , 4X SDS-Buffer (in-house), 6X DNA loading dye (#R0611, Fermentas) and 10X BlueJuice™ (#10816-015,

Invitrogen). Complete protease inhibitor cocktail tablets were ordered from Roche (#04693132001).
For the enzymatical assay, 4-Nitrophenyl ß-D-glucopyranoside (N7006, Sigma-Aldrich) was ordered. Protein assays were carried out with the Bio-Rad Protein Assay Dye Reagent Concentrate (#500-0006, Biorad).

DNA ladders were Quick-Load® 100 bp DNA Ladder (#N0467S, NEB), Quick-Load® 1 kb DNA Ladder (#N0468S, NEB) and GeneRuler™ 1kb DNA Ladder Plus, ready-to-use (#SM1334, Fermentas).
Protein ladders were Spectra™ Multicolor Broad Range Protein Ladder (#SM1842, Fermentas), BenchMark™ Pre-Stained Protein Ladder (#10748-010, Invitrogen), BenchMark™ Unstained Protein Ladder (#10747-012, Invitrogen) and 6xHis Protein Ladder (#34705, Quiagen).

2.1.4. Enzymes

All enzymes were ordered from New England Biolabs, :
Restriction enzymes were *Pml*I (#R0532) *Nde*I (#R0111S), *Stu*I (#R0187s), *Fse*I (#R0588), *Nae*I(#R0190), *Xba*I (#R0145), *Xho*I (#R0146), *Mly*I (#R0198) and *Sac*I (#R0156). For dephosphorylation the Antarctic Phosphatase (#M0289S) was ordered.
DNA Polymerases was the thermophilic Deep VentR DNA Polymerase (#M0258)
The ligases were T4 DNA Ligase (#M0202S). and T4 DNA Ligase (conc.) (#M0202T).

2.1.5. Media

For selecting *E.coli* transformants LB-Miller agar plates according to Giuseppe Bertani were used: 0.5% (w/v) yeast extract, 1% (w/v) tryptone, 1% (w/v) NaCl and 1.5% (w/v) agar-agar. 100µg/ml Ampicillin or 60µ/ml carbenicillin (extended lifetime) for pet303 and 50µg/ul kanamycin for pet28b were used. The BL21Star+GROEL/ES needed also 34 µg/ul chloramphenicol. On the contrary, selecting zeocin-resistant colonies was done by using low-salt LB-Miller (5g/l NaCl) and 25µg/ml zeocin.
Expression media for *E.coli* transformants included LB-Miller and YT (1% (w/v) yeast extract, 2% (w/v) tryptone and 1% (w/v) NaCl).
*P. past*oris X33 transformants were selected on YPDS plates (1% (w/v) yeast extract, 2% (w/v) peptone,
2% (w/v) dextrose, 1M sorbitol, and 1.5% (w/v) agar-agar) and YPD (sorbitol was omitted) plates with each 25µg/ml zeocin.
PichiaPink strains were non-specifically grown on YPD plates, groomed in YPDS medium and selected on PAD agar plates (0.4% (w/v) potato starch, 2% (w/v) dextrose, and 1.5% (w/v) agar) without antibiotics due to red/white selection.

Large expression of *P. pichia* transformants was performed in the media according to the *Pichia* expression kit manual.
SOC medium was delivered with the bacterial expression hosts.

2.2. Methods

2.2.1. Construction of the vectors

Everytime the z1-ßgl plasmid from Genscript acted as a template for constructing the vectors. This vector was especially designed for the pet303/CT-His plasmid with *Xba*I upstream and *Xho*I downstream of the codon region of the beta-glucosidase gene.
So the gene for the pet303/CT-His was simply cut out with *Xba*I and *Xho*I respectively, ligated with the digested pet303/CT-His and transformed into chemically competent TOP10 *E.coli* for selection (for restriction, ligation and transformation review 2.2.1 later on and 2.2.2. for selection procedures).
In case of the other three expression plasmids pet28b/NT-His, pPICZαC and pPinkα-HC primers had to be constructed first. Considerations about the start- and stop-codon, his-tag usage, restriction sites, framing-in and melting temperatures were highly essential.

2.2.1.1. Primers

Pet28b: Sense primer 5'- *aaaaaa* CATATG atgaccgaaaacgcggaaaatttc and antisense primer 5'- *aaaaaa* CTCGAG tta ctgataccacagcgcgct (*Nde*I & *Xho*I).
pPICZαC: Sense primer 5'- *cacga* CACGTG ag*[1] atgaccgaaaacgcggaaa and antisense primer 5'-*ttgtgc* TCTAGA taccacagcgcgctac (*Pml*I & *Xba*I).
pPinkα-HC: Sense primer 5'- *caagc* GAGTCAGCTA accgaaaacgcggaaa and antisense primer 5'- *gttgat* GGCCGGCC tca **atgatgatgatgatgatg** ctgataccacagcgcgct (*Mly*I & *Fse*I).
The italic letters refer to recognition sequences for the restriction enzymes, the big letters are the restriction sites, the bold letters are coding histidine residues for the artificial his-tag, the underlined letters are a start- or stop-codon and the standard letters are the annealing part of the gene itself. The names of the restriction enzymes are indicated in brackets. The To ensure a correct frame-in with the secretion signal site and the his-tag of the pPICZαC the sense primer contains two additional bases labeled with *[1] and four less bases on the antisense side.
The primers were designed on paper and with the software CLC Genomics Workbench 3. They were ordered from an external facility.

2.2.1.2. PCR

In a next step these primers were used in a PCR (Polymerase Chain Reaction) to generate genes with artificial restriction sides that fit into the appropriate plasmids. The PCR is a common and simple technique to amplify low amounts of DNA to generate a high concentration of the same sample. The primer set the borders and so the length of the resulting DNA piece. The enzyme called *polymerase* adds the missing bases to the single-stranded DNA.
The standard procedure for the Deep Vent Polymerase was used and optimized over time

Components of the Deep Vent Polymerase PCR reaction::

Component	**End Concentration**
Sense primer	0.5µM
Antisense primer	0.5µM
25mM $MgSO_4$	1.5mM
10 mM dNTP	1.5mM
Thermopol Buffer (10X)	1X
DNA template	50-100ng
Deep Vent Polymerase	0.5 units
water	Up to a final volume of 50µl

Table 1 Components of a Deep Vent Polymerase PCR reaction

Cycling conditions adjusted to the use of Deep Vent Polymerase:

Cycle Procedure	**Time**	**Temperature**	**Repeat**
Stage I	04:00	95°C	1x
Stage 2	00:20	95°C	30-35x
	00:30	annealing temp. (variable)	
	02:00	72°C	
Stage3	04:00	72°C	1x

Table 2 Cycling conditions of a PCR reaction

In initialization step 1 the DNA is heated up to the working temperature. Stage 2 contains of the denaturation, the annealing and the extension step. In here the DNA denatures, the primers anneal to the template and the polymerase elongates the DNA. In the final stage 3 all remaining single-stranded DNA are extended. The annealing temperature of a primer pair is simply calculated by substracting 5°C from the melting temperature. The sense and antisense primer should have almost the same melting temperature. If no DNA was made, an annealing temperature screening in 2°C steps is recommended.
All the PCR reactions were done in the Veriti 96 well Thermal Cycler from Applied Biosystems.

2.2.1.3. Agarose gel electrophoresis

The result of the PCR was checked by agarose gel electrophoresis. A mixture out of agarose dissolved in TEB Buffer defines the gradient. The more agarose is used the easier small-sized DNA pieces are separated. To visualize the bands ethidium bromide, a toxic chemical substance, is added to the liquid gel mixture. A loading dye makes the samples compatible for the gel wells and assesses how fast the gel is running.

Plasmids were separated with 0.8% (w/v) gels, all other DNA with 1.2% gels. Loading dye was added to a final concentration of 1X. Depending on the size of the wells a certain amount of sample was loaded on the gel. The voltage was set at 100V.

The finished gel was checked under UV light and if necessary bands were cut out with a clean razor blade. The agarose pieces were purified with the Zymoclean™ Gel DNA Recovery Kit from Zymo Research.

2.2.1.4. Restriction digest

In a restriction digest an enzyme called restriction enzyme recognizes and cuts certain nucleotide sequences also known as restriction sites. Each restriction enzyme produces either sticky ends (with overhang) or blunt ends (no overhang). Three classes subdivide the mass of enzymes into their nature of cleavage. Only type II, which recognizes and cleaves DNA at the same spot, was used in this experiment.

The standard components of a restriction digest with NEB enzymes were as follows:

Component	**End Concentration**
Restriction Enzyme I	depending on the DNA amount
Restriction Enzyme II (if double digest)	
100X BSA	1X
10X NEBuffer (1,2,3,4)	1X
DNA template	defined by user
water	Up to a final volume of 20µl

Table 3 Components of a restriction digest mix

One restriction enzyme was used in a single digestion whereas two were added in a double digestion. The double digest finder from the NEB web page (http://www.neb.com/nebecomm/doubledigestcalculator.asp) was consulted to check the availability of a double digest, the reaction temperature, the buffer and if BSA (bovine serum albumin) was needed or not.

The newly generated z1-ßgl genes from the PCR reaction and their plasmids were cut with the specific enzymes. One unit of enzyme was defined to cut 1µg of DNA in one hour at the reaction temperature. The pair pet28b and z1-ßgl-28b was cut in a double digest with *Nde*I & *Xho*I, pPICZαC and z1-ßgl-pPIC with *Pml*I & *Xba*I. The pPinkα-HC plasmid was cut with *Stu*I

and *Fse*I whereas its gene z1-ßgl-pPink with *Mly*I & *Fse*I. The pet303 and Z1-ßgl-303 were cleaved out of the Genscript plasmid with *Xba*I and *Xho*I.
The cleaved plasmid DNA was incubated with 5 units of Antarctic Phosphatase and 1X Antarctic Phosphatase Reaction Buffer for another 20 minutes at 37°C. That impeded the self-ligation of the plasmid due to removal of the 5' phosphate group. Afterwards the vector and insert were heat inactivated for 20 minutes at 65°C, loaded on an agarose gel and purified with the Zymoclean™ kit.

2.2.1.5. Ligation

During a ligation an enzyme called *ligase* closes the gap between complementary DNA ends. The ligase can't add any more bases. Connecting sticky ends is faster and more efficient than blunt ends because more bases get in touch with each other.

The ligation of the loose DNA pieces was performed in a T4 DNA Ligase reaction. The standard reaction mix of the ligation was as follows:

Component	**End Concentration**
Vector DNA	Ratio 1:2 – 1:6
Insert DNA	
10X Ligation Buffer	1X
T4 DNA Ligase	400U (standard) 2000 U (concentrated)
Water	Up to a final volume of 10µl

Table 4 Components of a ligation mix

This reaction mix was spun down shortly to remove all excess liquids from the tube walls. The process itself was performed at room temperature (20-25°C) for about 10-30 minutes depending which ligase was used. The reaction mix was heat inactivated at 65°C for 10min and immediately used for transformation.

2.2.2. Selecting an error-free clone

During restriction and ligation reactions negative side effects can appear. Some examples are wrong restriction sites, re-ligation of the plasmid and ligation with more than one gene sequence. To ensure a perfect plasmid it has to be sequenced prior to final transformation into expression hosts. First of all the plasmid concentration is way too low to check by sequencing. Plasmid propagation hosts like One Shot™ TOP10 *E.coli* generate more plasmids which are isolated following a purification protocol.

2.2.2.1. Transformation into TOP10 *E.coli* / Mach1 cells

The One Shot™ TOP10 *E.coli* were ordered from outside, whereas the Mach1 cells were produced in-house. That made the Mach1 cells much cheaper. The transformation protocol for the chemically competent cells for both systems is the same. Following the transformation procedure of TOP10 *E.coli* from Invitrogen the transformed cells were spread on selective LB agar plates and incubated overnight at 37°C.

Some of the plasmids were transformed into One Shot® TOP10 Electrocomp™ *E.coli* according to the electroporation protocol. In general, this method offers higher efficiency.

2.2.2.2. Screening and isolating the plasmid

After the colonies had developed on the plates some of them were separately inoculated into 10ml of LB medium with antibiotics at 37°C overnight. The next day the cultures that indicated growth were purified with the QIAprep Spin Miniprep Kit from Quiagen.

The amount of DNA was quantified with the NanoDrop ND-1000 Spectrophotometer from Thermo Scientific (Example Figure 1).

Sample ID	User ID	Date	Time	ng/ul	A260	A280	260/280	260/230
1	Default	6/11/2009	6:00 PM	27.39	0.548	0.287	1.91	1.10
2	Default	6/11/2009	6:01 PM	14.61	0.292	0.114	2.57	0.06
3	Default	6/11/2009	6:02 PM	22.49	0.450	0.249	1.80	0.97
4	Default	6/11/2009	6:05 PM	44.98	0.900	0.482	1.86	1.31
5	Default	6/11/2009	6:06 PM	32.50	0.650	0.482	1.35	0.07

Figure 1 Nanodrop example

If there were too many colonies on one plate, a strict screening of about 20-30 colonies had to be made. Using a 96 deep-well plate, every 1ml of LB with appropriate antibiotics was inoculated with these colonies, grown up for about 5-6 hours at 37°C on a high-speed shaker (300rpm). In a Colony-PCR 1µl of the culture was added instead of the DNA template. Comparing the component table of *2.2.1.2 PCR*, the first denaturing stage was set at 7 instead of 4 minutes to disrupt the cells. T7 or gene-specific primers identified the clones holding the gene which were visualized by an agarose gel. The cultures were re-inoculated in 10ml of LB and the same procedure stated above began.

2.2.2.3. Sequencing the plasmid

The purified plasmid DNA was then prepared for sequencing at an exterior facility called *Elimbio* (www.elimbiopharm.com). At least 500ng of plasmid were needed to ensure a qualitative measurement. T7 primers for the bacterial plasmids, 5′ *AOX1 Pichia* Primer in combination with 3′ *AOX1* Pichia Primer for the pPICZαC and 5′ *AOX1* Pichia Primer + 3´ *CYC1* primer for the pPinkα-HC were used as sequencing primers.

The results of each sequence were reviewed thoroughly. If one sample was absolutely correct the appropriate plasmid was propagated again and stored at -20°C for further experiments.

2.2.3. Transformation of bacterial plasmids into expression hosts

The error-free bacterial plasmids pet303+z1-ßgl and pet28b+z1-ßgl were transformed into their expression hosts following the manufacturer's protocol from Novagen (OrigamiB) and Invitrogen (BL21). Both protocols describe the transformation method of chemically competent cells.
Pet303-z1-ßgl was only transported into OrigamiB whereas pet28b+z1-ßgl into OrigamiB and all three BL21 strains. The solutions were spread on selective LB plates using Carbenicillin for pet303+z1-ßgl, Kanamycin for pet28b+z1-ßgl and Kanamycin with Chloramphenicol for the co-expression host BL21+GROEL/ES with pet28+z1-ßgl.
For an expression background control the wildtype of each strain was grown up at the same time.
The plates were put into a 37°C incubator until the first colonies developed after a few days.

2.2.4. Transformation of eukaryotic plasmids into expression hosts

Only an electroporation reaches an efficiency that is high enough to transform yeast cells. As electrocompetent cells are not commercially available due to their fast decay, all the strains including X33 and pPink had to be made electrocompetent on the day of the transformation.
The slightly different procedures are described in the manufacturer's protocols from Invitrogen: *EasySelect Pichia Expression Kit* (X33) and *PichiaPink Expression System* (pPink).
The settings for the BTX electroporator ECM 630 for *P. pastoris* were 1500V, 200Ω, 25µF, 1 pulse. The cuvette size was 1mm. The transformations were spread on plates after at least 2hours of recreation in a 30°C room.
These plates were placed in a 30°C oven until the first colonies developed.

2.2.5. Expression of z1-ßgl in *E.coli*

The standard procedure for expressing proteins in OrigamiB was set as follows:
10ml LB media with 1X Carbenicillin was inoculated with one colony from the transformation plate. These grew for about one day at 37°C. A volume of 1L LB was inoculated with the overnight culture to an OD of 0.01-0.02. After reaching an OD of 0.5 the culture was induced with 0.5mM isopropyl-ß-d-thiogalactopyranoside (IPTG). The shaker was set at 37°C and 200rpm.
After 7-8 hours expression the cells were harvested by centrifugation. In case of large cultures (>100ml) the pellet was resuspended in 3ml PSB (Phosphate Buffered Saline). The suspension was disrupted with the French Pressure Cell 40k from Thermo twice at 2500psi. The lysate was filled into ultracentrifuge tubes and spun down until the supernatant solution was clear. In case of small-scale expressions (<100ml) the pellet was resuspended in 200-300µl *E.coli* Lysis Buffer (50mM Phosphate pH 7.8, 400mM NaCl, 100mM KCl, 10% (v/v) Glycerol, 10mM imidazole) and sonicated several times in 15sec intervals at 70% amplitude on the Sartorius Labsonic M. It was

necessary to chill down every step to 4°C. The supernatant of the lysate and the pellet were immediately used or stored at -20°C for further analysis.

The standard procedure for expressing proteins in BL21 was set as follows:
BL21(DE3) and BL21star (DE3) colonies were inoculated in each 5ml of LB with 1X Kanamycin, BL21star+GROEL/ES (DE3) into 5ml of LB with 1X Kanamycin and 1X Chloramphenicol. The cells grew overnight at 37°C and were re-inoculated into 50ml LB with antibiotics to an OD of 0.01-0.02. Into the BL21Star+GROEL/ES medium 20% (w/v) L-Arabinose was added to a final concentration of 0.5%. The induction time for BL21 (DE3) and BL21Star (DE3) was set at an OD of 0.6, the one of BL21Star+GROEL/ES (DE3) at 1.0 to develop enough chaperones. 0.5mM IPTG were added to the cells and expressed for 20 hours at 200rpm, 30°C.
The cells were spun down by centrifugation and the pellet was fully resuspended in *E.coli* Lysis Buffer (50mM Phosphate pH 7.8, 400mM NaCl, 100mM KCl, 10% (v/v) Glycerol, 10mM imidazole). The suspension was French pressed twice at 2500psi. The lysate was centrifuged until the supernatant was clear of particles. The supernatant and the pellet were either immediately used or stored at -20°C for further analysis.

2.2.6. Expression of z1-ßgl in *P. pastoris*

The standard procedure for expressing proteins in X33 was set as follows:
The transformants from the plates were inoculated into 5ml of BMGY in baffled flasks at 30°C, 200rpm until the OD was 1.3 – 1.5. After spinning them down at 2500rpm for 5min, the pellets were resuspended in 5ml of BMMY and put back at 200rpm for another 24 hours. Whenever the expression exceeded 24 hours, 40% methanol was added to a final concentration of 0.5%.
When the expression was finished the cells were spun down by centrifugation (2500rpm in 5min for 1-10ml; 10000rpm in 10-20min for >100ml). The supernatant was stored at -20°C and the pellet at -80°C.

The standard procedure for expressing proteins in *Pichia* Pink was set as follows:
The transformants of strain IV as well as the wildtype of strain IV were all inoculated into each 10ml of BMGY and incubated overnight at 30°C. As much cells as needed were spun down to reach an OD of 1.0 before adding 1ml BMMY. After 48 hours at 30°C and 220rpm the expression was stopped by centrifugation (2500rpm in 5min for 1-10ml; 10,000rpm in 10-20min). Both were stored at -20°C for short and -80°C for long terms. Whenever the expression exceeded 24 hours, 40% methanol was added to a final concentration of 0.5%.
The cells were opened by resuspending the pellet in Breaking buffer (50mM sodium phosphate pH 7.4, protease inhibitor cocktail, 1mM EDTA, 5% (v/v) glycerol) and using protocols described in the user manual of the PichiaPink system like glass beads or French press twice at 2500psi. The lysate fractions were separated by centrifugation (30,000rpm for 20min, 4°C) and stored at -20°C for immediate or -80°C for longer terms.

2.2.7. SDS-PAGE and Western Blots

For separating proteins of different sizes, sodium dodecyl sulfate polyacrylamide gel electrophoresesis (SDS-PAGE) was carried out as described in Laemmli et.al [9]. To ease the procedure Criterion Precast gels with different percentages from Biorad were used. The electrophoresis was operated on the Biorad system *Criterion Cell.* The samples were boiled for 8min, loaded and all the gels ran at 150V for about 1h 15min. The bands were visualized with Gelcode Blue Stain Reagent.

For detecting proteins via antibodies the Western Blot system ONE-HOUR Advanced Western Kit (Mouse) from Genscript (#L00242) was used. The method on nitrocellulose was first described by H. Towbin et al. [10]. The blotting of the proteins was carried out on the iBlot Dry Blotting System from Invitrogen. The program was set on 7 minutes transfer time.
The procedure was performed according to the L00242 protocol in the technical manual. Some important notes are that: 1) Mixture 1 was prepared by adding 1µg of Anti-His Antibody into 80µl of WB1 solution. 2) The signal was developed with the LumiSensor Super Chemiluminescent HRP Substrate.
The result of the Western Blot was visualized in the LAS-4000 system from Fujifilm.

2.2.8. Enzyme and protein assays

The activity of beta-glucosidase was measured by using a substance called pNPG (para-nitrophenyl ß-D-glucopyranoside) first described by M. Aminlari et al. [11]. This colormetric substance is composed of a glucose unit and para-nitrophenyl unit that turns yellow when pNPG is cleaved at the ß-1,4 link.
One enzyme reaction mix with a total volume of 100µl contained 50mM sodium citrate, 10mM pNPG and 25µl of the enzyme solution. Negative controls like “no enzyme” or “wildtype” were included. According to Takase et al. [2] the optimum temperature was 80°C. After 10min the reaction was stopped with 100µl of 1M sodium carbonate. The absorbance was read at 400nm with the Paradigm Detection Instrument from Beckman Coulter using the absorbance detection cartridge.
This enzyme assay served mainly as a rule to determine activity or no activity. Without measuring the protein amount of each sample they couldn’t be compared to each other.

Protein concentrations were determined with the Bradford method. The protein assay kit was from Bio-Rad with the standard bovine serum albumin (BSA). The optimized procedure was to create six BSA standard curve points from 0µg to 15µg (0, 1.5, 4.5, 7.5, 10.5, 15 µg) of protein. Each of the six amounts was added into 1ml of 1X dye solution. This provided a range of colors from brown (no BSA) to deep blue. Unknown concentrations of protein were added to fresh 1ml of 1X dye solution until the color fell within range of the standard. All samples were incubated

for 10min and the absorbance was measured at 595nm. Two graphs were blotted, one for the standard curve and the other one for the unknown samples.

2.2.9. Purification of the recombinant z1-ßgl

The 6xHis-tagged proteins have a great affinity towards nickel charged beads. Purifications using this technique are very selective and lead to high concentrations in the eluate. Prior to using a purification method the samples had to be filtered to remove all particles in the solution.

The purification of z1-ßgl was performed by applying various Ni^{2+}-affinity-His-binding methods:

a) Method I: Purification via His-binding resin column
The HisTrap FF (Fast Flow) or HP (High Performance) column was used on the ÄKTAexplorer 100 system, both from GE Healthcare. The resin column of 5ml was equilibrated with Binding buffer (20mM sodium phosphate pH 8.0, 500mM NaCl, 20mM imidazole). The sample was diluted with the same amount of binding buffer. After loading the sample on the column, it was washed with Binding buffer until all non-specific proteins were gone. Then the target protein was eluted with Elution buffer (20mM sodium phosphate pH 8.0, 500mM NaCl, 500mM imidazole). A gradient of 0-100% over a certain time period was included. Fractions were collected with an automatic fraction collector.

b) Method II: Purification via Ni-NTA Magnetic Agarose Beads under nature conditions
This method was performed with small amounts of cell lysate in 1.5ml falcon tubes. The separation device was the magnetic rack 12-Tube Magnet (#36912, Quiagen) consisting of a series of 12 powerful magnetic NdFeB disks. Buffers could be exchanged while the beads where hold by these magnets.
200µl of Ni-NTA agarose suspension was equilibrated with 2x 1.5ml water and 2x 1.5ml Breaking buffer (*P. pastoris*) / Lysis buffer (*E.coli*), always discharging the supernatants. The dry beads were added to 1ml of cell lysate. The samples were incubated for an hour at 4°C while turning up and down. Then the samples were spun down at 12.000g for 1min and placed on the magnetic separator for 1min. The supernatant was removed and the beads washed three-times with 200µl of Wash buffer (50mM Na_2HPO_4, 300mM NaCl, 30mM imidazole; 0.05% Tween20, pH 7.4). Elution was performed with 50µl of Elution buffer (50mM Na_2HPO_4, 300mM NaCl, 250mM imidazole; 0.05% Tween20, pH 7.4). All the steps were done at 4°C.

c) Method III: Purification via Ni-NTA Magnetic Agarose Beads under denaturing conditions
If the His-Tag seems to be inaccessible, the protein has to be purified under denaturing conditions. The samples were denatured in 8M Urea for about 3 hours at 37°C. After equilibrating 600ul of Ni-NTA Beads suspension as stated in Method II the lysate was added. The solution was incubated for one hour at 4°C while shaking up and down until applied to an

8ml filter tube. The washing mixtures were prepared as follows: 10 mixtures with Wash buffer and different urea concentrations ranging from 7.24M to 0M with a total volume of 5ml. One after another mixture was applied to the column so that the immobilized His-proteins could fold back into their nature state. Following that step, the beads were washed 3-times with 5ml pure Wash buffer and eluted 3-times with each 100ul Elution buffer.

A size exlusion chromatography was performed by using the Sephacryl S-100 HR column from GE Healthcare. The *E.coli* Lysis Buffer was used as the mobile phase. Buffer and samples were filtered to remove all particles. The loading and equilibrating steps were consulted from the instructions manual from GE Healthcare.

2.2.10. Chemical and physical treatments of cell lysates and supernatants

2.2.10.1. Heat treatment

This method separates heat resistant from heat incompatible proteins.
The sample is heated up to the optimum temperature of the z1-ßgl at 80°C for one hour. The precipitated proteins are separated by high-speed centrifugation for a couple of minutes. To remove any particles in the solution the sample was being filtered.

2.2.10.2. Ammonium Sulfate Precipitation

Ammonium sulfate alters the solubility of the protein.
This method was used to concentrate the proteins in a solution. An excess of ammonium sulfate was added to the sample and mixed until all the protein was precipitated. The solution was centrifuged at high-speed, the supernatant tossed and the pellet resuspended in as much buffer as needed

2.2.11. Refolding of inclusion bodies

Inclusion bodies (IB) are made out of insoluble proteins that remained in the pellet after centrifuging the lysed cells. Due to recombinant overexpression the proteins can aggregate and are in misfolded state. A sufficient method to denature and slowly re-fold inclusion bodies is described here.
The IBs were washed once with 20ml of 4M urea. After centrifuging the sample at 16.000g for 30min the IBs were resuspended in 20ml of 8M urea and denatured for several hours to overnight in a 37°C shaker. The reaction was stopped by centrifuging again at 16.000g for 30min. The supernatant was collected and interacted with 2ml of equilibrated Ni-NTA Agarose Beads for 1 hour on a 4°C shaker. At the end the sample was loaded on 8ml gravity columns

with filter. The refolding mixtures were prepared as described in *2.2.9* (Method III) with a total volume of 10ml. All 10 mixtures were applied consecutively and the column washed three times with pure Wash buffer. The target protein was eluted 3-times with each 2ml Elution buffer. Incubating times of 3-5min were maintained during each elution step.

3. Results

3.1. Completion of the vector and expression studies in OrigamiB

3.1.1. Completion of the vector pet303-z1-ßgl

The vector pet303-z1-ßgl was completed as described in *2.2.1.4* and *2.2.1.5*. Showing the first step in completing this vector, the z1-ßgl-303 gene out of the Genscript z1-ßgl plasmid and the pet303 are cleaved (Figure 2).

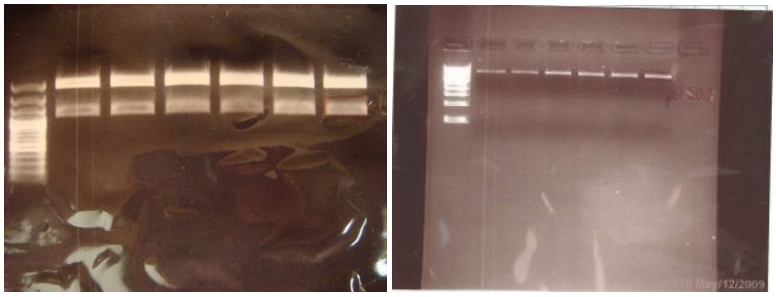


Figure 2 Cleaved z1-ßgl-303 gene and pet303

1µl of the ligation mix was transformed into chemically competent TOP10 *E.coli* according to *2.2.2.1*. Three colonies were picked from the agar plate, re-inoculated in 10ml of LB/Carb, purified and send for sequencing with the T7 primers according to 2.2.2.2/3.

3.1.2. Expression of the z1-ßgl-pet303 in OrigamiB

An error-free clone was transformed into OrigamiB described in 2.2.3 and grown on LB/Carb plates and a plasmid-less wildtype on LB/Kan for background control.

Two transformants and one wildtype were cultivated in each 1L LB with specific antibiotics. Just one transformant and the wildtype were induced with 1mM IPTG at an OD of 0.5. Further steps were carried out according to *2.2.5*.

The first purification was performed as in Method I of *2.2.9*, whereas the Binding buffer held 30mM imidazole and the pH of both buffers was set at 7.4. The connected column type was a HisTrap FF (FastFlow) which has lower binding capacity than HP columns.

The A_{280} of two elutions and the flowthrough was measured with the UV-spectrophotometer (Figure 3) which gives an approximate data about the protein amount in mg/ml.

Sample ID	User ID	Date	Time	mg/ml	260/280	A280 10mm
Elution I	Default	7/1/2009	5:24 PM	0.18	0.95	0.183
Elution II	Default	7/1/2009	5:24 PM	0.88	0.67	0.877
Flowthrough	Default	7/1/2009	5:25 PM	26.99	1.69	26.991

Figure 3 A_{280} of first z1-ßgl-303 purification

Elution 1 was taken from start till the half of the elution ramp while elution 2 held the main part.

A following SDS-PAGE according to *2.2.7* should have visualized the purified proteins on the gel, but no bands were visible.

Another purification of his-tagged z1-ßgl-303 expressed this time in 2L of LB media was undertaken. All the other conditions were kept as before. The chromatogram showed only little to no absorbance of protein in the eluate.

The course of action was changed and a SDS-PAGE with soluble/insoluble proteins of 5ml cultures of three OrigamiB transformants and one wildtype was run. They were expressed the same way as stated before. A 1min sonication was carried out on all cells as described in *2.2.5*. The supernatant was transferred into a new tube and the pellet was resuspended in 300µl of Lysis buffer. Each 50µl sample were taken and prepared for SDS-PAGE according to *2.2.7*.

30µl of the sample was loaded into each well. No overexpression of the enzyme in any transformant nor fraction was detected (Figure 4).

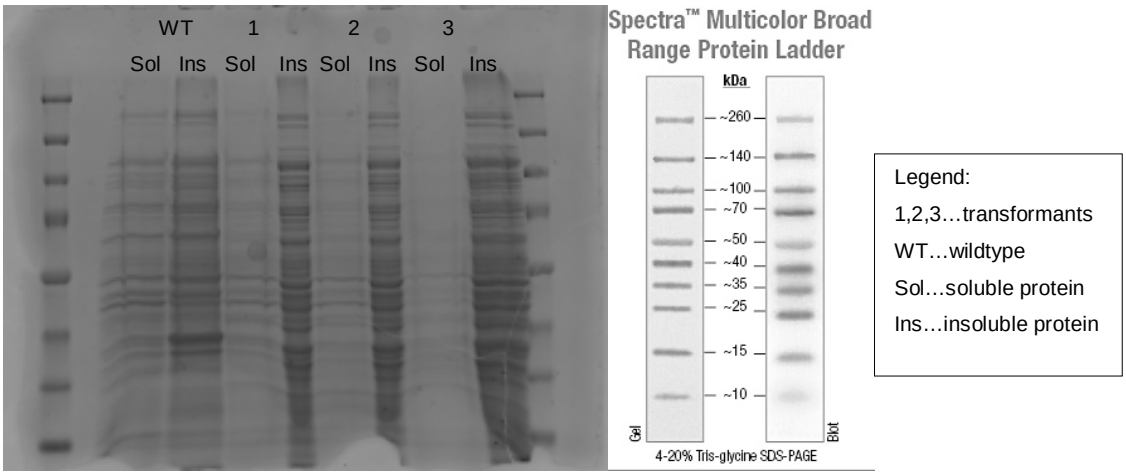


Figure 4 SDS-PAGE 10%, Soluble/Insoluble z1-ßgl-303

3.1.3. Re-screening of an error-free clone and expression in OrigamiB

The sequencing results were checked again thoroughly which revealed a failure in the transformation vector. The digestion and ligation procedures were repeated until an error-free clone emerged by screening 15 transformed TOP10 Electrocomp™ *E.coli* colonies according to *2.2.2.1*. This time, the sequencing data of the ligated vector and the original sequence (GenBank AB034947.1) were aligned in a blast on http://blast.ncbi.nlm.nih.gov/Blast.cgi. They matched to 100% (Figure 5).

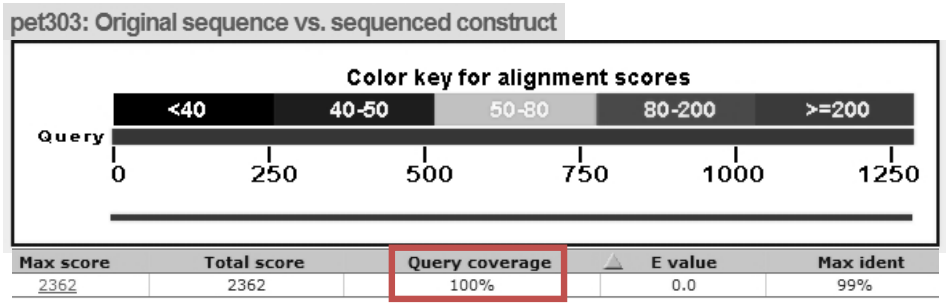
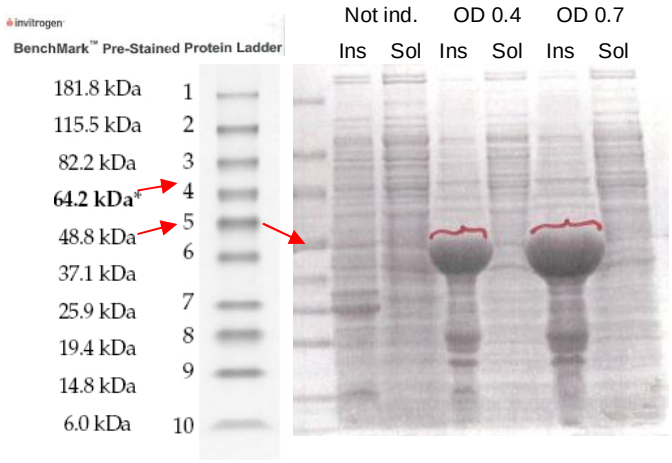


Figure 5 Aligning pet303-z1-ßgl with original sequence

The transformation of the error-free pet303-z1-ßgl occurred as stated before. A new expression was set up with 3x 5ml LB and specific antibiotics at 37°C. Two cultures were induced with 0.2mM IPTG at OD 0.4 and OD 0.7 whereas the third one was not induced. Proteins were expressed overnight at 25°C.

The protein bands were visualized by SDS-PAGE (Figure 6) and the activity of the soluble fraction recorded with a pNPG enzyme assay according to *2.2.8*.



As can be seen for the induced transformants, the beta-glucosidase was expressed almost entirely in the insoluble fraction. No soluble z1-ßgl proteins are visible on the gel.

Figure 6 SDS-PAGE 4-20%, Re-run of 303-z1-ßgl expression

OrigamiB	**not ind**	**OD=0.4**	**OD=0.7**
	0,0524	0,0506	0,0432

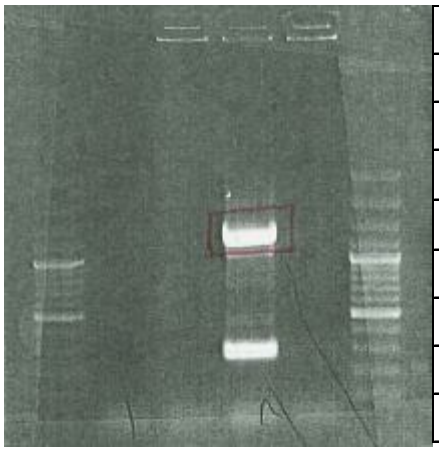
Table 5 Enzyme assay the re-run of 303-z1-ßgl

The values are stated in absorbance units. The enzyme activity of the soluble fraction is almost zero even lower than the one of the non-induced transformant (Table 5).

3.2. Completion of the vector and expression studies in X33

3.2.1. Completion of the vector pPIC-z1-ßgl

The vector pPICZαC-z1-ßgl was completed as described from *2.2.1.1* till *2.2.1.5*. The PCR was successful by annealing the gene-specific primers stated in *2.2.1.1* at a temperature of 57°C (Figure 7).

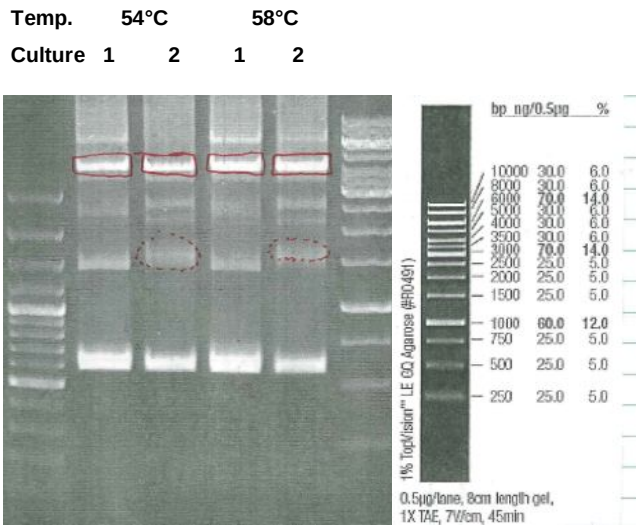


For Prim	1,5
Rev Prim	1,74
$MgSO_4$	3
10 mM dNTP	1,5
Thermopol Buffer (10X)	5
Deep Vent Polym	0,75
DNA	3
Water	33,51
Total	**50**

Figure 7 Agarose gel 1.4%, PCR of pPICZαC z1-ßgl gene

The red circled band was at about 1300bp which is the approximate number of bases of the pPICZαC -z1-ßgl gene with restriction sites. The second lower band was unknown but didn't have any influence on the experiment. The ladder to the left was the Quick-Load 100bp DNA ladder from NEB and was used for comparison.

After completing the vector, the plasmid was sent for sequencing as described in *2.2.2.3*. The results for all miniprepped colonies were bad. A Control-PCR of the miniprepped plasmids with AOX1 primers showed no evidence of the z1-ßgl gene within the MCS (Figure 8).



The red circled band was too far away from the working point (4000bp). The lower band around 600bp could be lead back to a PCR without insert – a gene was polymerized containing the AOX priming sites and the empty MCS.

Figure 8 Control-PCR of z1-ßgl insertion into pPICZαC

All the steps from *2.2.1.1* had to be repeated. The PCR was performed as before at 57°C annealing temperature. The band at 1300bp was cut out, purified as described in *2.2.1.3* and prepared for restriction cleavage. Both the plasmid and the gene were double digested according to *2.2.1.4* at 37°C overnight. No arctic phosphatase was used to dephosphorylate the plasmid. Two ligation mixes were arranged, one in 3:1 the other in 2:1 ratio: The ligations were performed with standard T4 DNA Ligase at 16°C overnight and transformed into chemically competent TOP10 *E.coli* as described in *2.2.2.1*. Each two colonies were picked from 3:1 and 2:1 transformation plate. They were cultured in 2x 25ml low salt LB/Zeocin and purified with the Miniprep method according to *2.2.2.2*.

The sequencing was done by Elimbio with AOX primers where just plasmid A turned out to be correct.

3.2.2. Expression of the z1-ßgl-pPIC in X33

The miniprepped pPICZαC-z1-ßgl plasmid from 3:1 plate was propagated in TOP10 *E.coli* and linearized by an excess of *Sac*I. Transformation of the X33 strain was performed according to *2.2.4*. 250µl and 750µl transformation mix were spread on YPDS/Zeocin plates. The first colonies appeared after 3 days on both plates.

6 colonies of the plates were spread on fresh YPDS/Zeocin plates and inoculated into 5ml of BMGY/Zeocin growth medium. They were cultured at 30°C for 24hours on a 200rpm shaker. The cells were pelleted by centrifugation at 2500rpm for 5min, resuspended in 5ml of BMMY and further expressed at 30°C for 24hours. The expression was stopped by centrifugation at 2500rpm for 5min. Both fractions were stored at -20°C. Samples of the supernatant were loaded on a SDS-PAGE which was run according to *2.2.7*. Because no bands appeared on the gel, the preparation of the samples was reviewed. The samples were extremely diluted prior to SDS-PAGE which lowered the whole protein concentration too much (Figure 9).

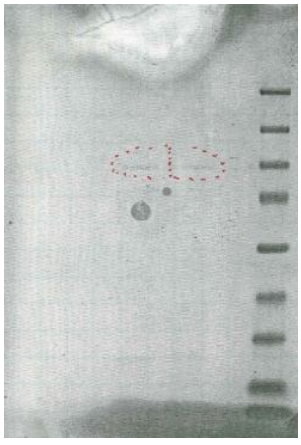


Figure 9 Diluted SDS-PAGE of X33 cultures

Around the red doted area some bands could be detected. This amount of protein is not sufficient to issue a statement.

In a second attempt with concentrated samples there were no bands detectable once again.

3.2.3. Selecting the most suitable medium for expression

Furthermore the best media for the expression was identified. 4 inoculations of the X33 transformants with each 5ml of BMGY were set up. The growth was carried out according to *2.2.6* until the each pellet was resuspended either in BMMY, MMY, BMM or MM medium to an OD of 1.0. The expression proceeded at 30°C over 5 days. Every day one sample was taken from each flask and the culture re-induced methanol to a final concentration of 0.5%.
A small Bradford Assay was performed on the secreted fraction of all media according to *2.2.8*. The amount of protein secreted increased over the days (Figure 10).

Absorbance at 595nm	BMMY	MMY	BMM	MM
Day 3	0,79	0,96	0,52	0,52
Day 5	0,82	1,1	0,55	0,54
Medium	0,6	0,59	0,53	0,52

Amount of Protein	BMMY	MMY	BMM	MM
Day 3	0,03504043	0,24332272	n.d	n.d
Day 5	0,07179613	0,4148493		
Medium	n.d	n.d		

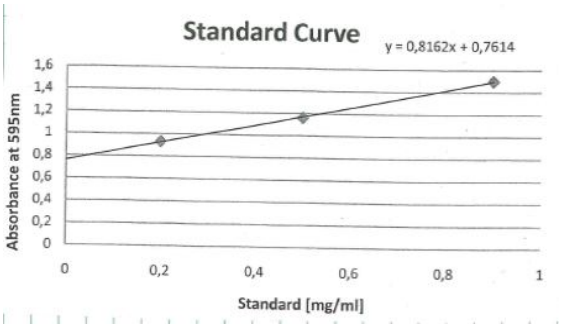


Figure 10 Bradford Assay of the X33 media selecting experiment

The row "Medium" is the standard absorbance of the medium itself and was subtracted from the sample measurements. Using the standard curve equation protein amounts could be converted from absorbance units to mg/ml. The MMY medium had with 0.415 mg/ml the most proteins secreted in 5 days. The media BMM and MM contained only indefinable traces of proteins.

A Western Blot according to *2.2.7* was performed later on. No z1-ßgl sources were detected (Figure 11). The bands visible show the protein ladder.

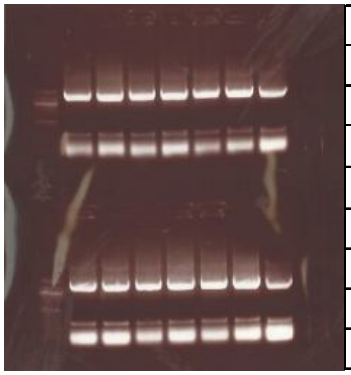
Figure 11 Western Blot of X33 samples of different media

In the meanwhile the PichiaPink system with the pPinkα-HC plasmid was introduced. The work on the X33 was put aside.

3.3. Completion of the vector and expression studies in BL21

3.3.1. Completion of the vector pet28b-z1-ßgl

The generation of the z1-ßgl-28b gene was easily achieved in the first attempt by annealing the primers stated in *2.2.1.1* at 52.7°C in a PCR reaction according to *2.2.1.2*. After running an agarose gel all bands perfectly arranged around 1300bp (Figure 12).



pet28b for	1,37
pet28b rev	1,48
MgSO4	3
10 mM dNTP	1,5
Thermopol Buffer (10X)	5
Deep Vent Polym	0,75
DNA	3
Water	33,9
Total	**50**

Figure 12 PCR of the z1-ßgl-28b gene

The bands were purified with the Zymoclean Kit and prepared for a double digest as described in *2.2.1.4*.

The double digest finder from NEB suggested to use NEBuffer4 +BSA at 37°C (Figure 13).

Enzyme	Cat#	Temp	Supplied NEBuffer	Supplements		% Activity in NEBuffer			
				BSA	SAM	1	2	3	4
NdeI	R0111	37°C	NEBuffer 4	No	No	75	100	75	100
XhoI	R0146	37°C	NEBuffer 4	Yes	No	75	100	100	100

Double Digest Recommendation(s) for NdeI + XhoI:

- Digest in NEBuffer 4 + BSA at 37°C.

Note: The above recommendation is based on the experimental results. Please check Suggested NEBuffers for Double Digestion.

Figure 13 Double Digest Finder for pet28b and z1-ßgl-28b gene

The DNA pieces were loaded on two agarose gels after a 24 hour reaction (Figure 14). The plasmid negative control (supercoiled DNA) of the left gel in the last column runs farther than the cleaved plasmid, which on the other hand gives back the exact base length (5368bp +/- cleaved MCS).

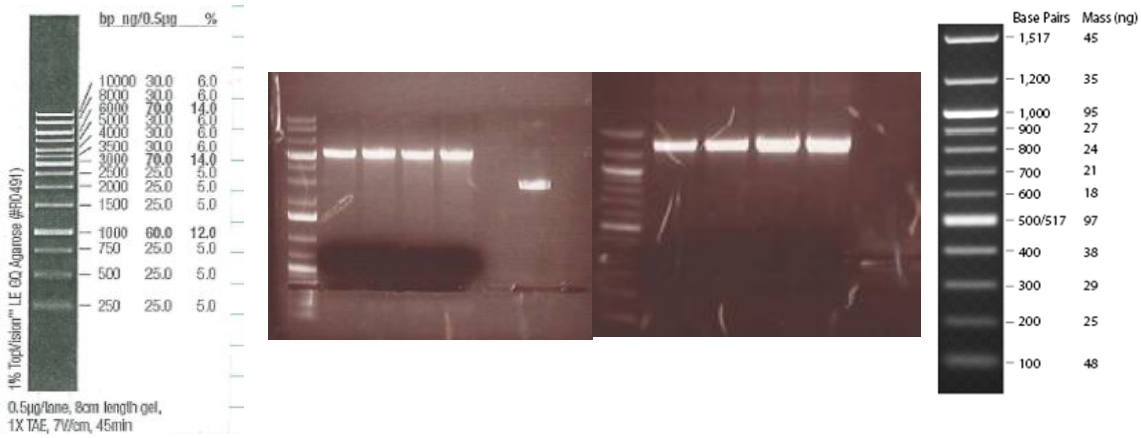


Figure 14 Restriction of plasmid and gene of pet28b

The modality of calculating the ratio of insert to vector was slightly changed. Instead of using the weight mass, the molecular mass of the DNA was involved (Table 6). To improve the ligation credibility five ligation mixes were performed: Three normal with different ratios, one without GOI (Gene of Interest) and one without ligase.

Ratio	**3**	
Mass_vector	20	fmol
Mass_insert	60	fmol

pet28b + z1-ßgll	**Ligation**	Conc [fmol/ul]
Vector	2,26	8,84
Insert	0,51	118,35
Ligation Buffer 10X	1	10
T4 Ligase	2	2000 U/ul
Water	4,23	-
Total	**10**	

Table 6 Ligation of pet28b and z1-ßgl-28b

Each time the ratio was increased by 1 the amount of the insert increased by 20fmol.
The ligation was performed according to *2.2.1.5*.

The completed vectors were transformed into electrocompetent TOP10 *E.coli*, screened and sequenced as described in *2.2.2*.
Running an agarose gel after the colony-PCR of 15 electroporated colonies some of them exposed themselves to contain the GOI. Except colonies 2, 3, 14 and 15 every other colony held the z1-ßgl-28b gene (Figure 15).

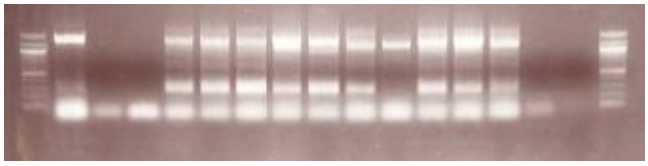


Figure 15 Screening of pet28b colonies for GOI

Four of five candidates which were sent to the DNA sequencing facility had the GOI inside. The aligning on both sides of the pet28b-z1-ßgl- gene were perfect (Figure 16). The empty part in the middle is not of interest as long as both restriction ends are perfectly framed in.

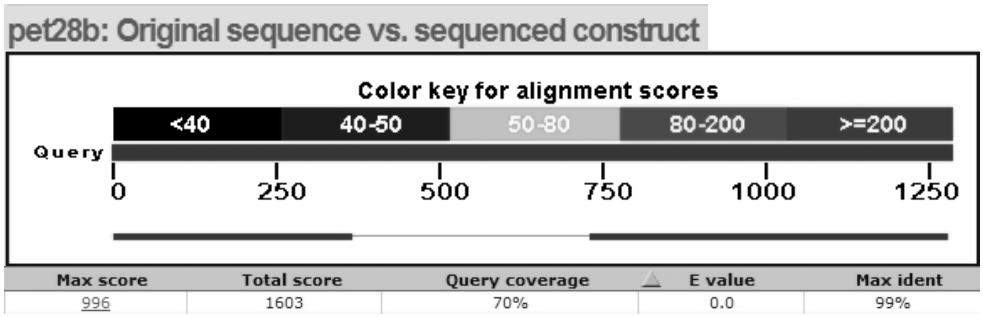


Figure 16 Aligning pet28b-z1-ßgl with original sequence

3.3.2. Expression of z1-ßgl-pet28b in BL21 strains

The correct plasmids were propagated via Miniprep and transformed into all three BL21 strains afterwards. The transformation was done according to *2.2.5* by adding each 8ng/µl of empty pet28b plasmid and completed vector pet28b-z1-ßgl into BL21 (DE3), BL21Star (DE3) and BL21Star(DE3)+GROEL/ES.

After a few days the first colonies developed on LB plates (Figure 17).

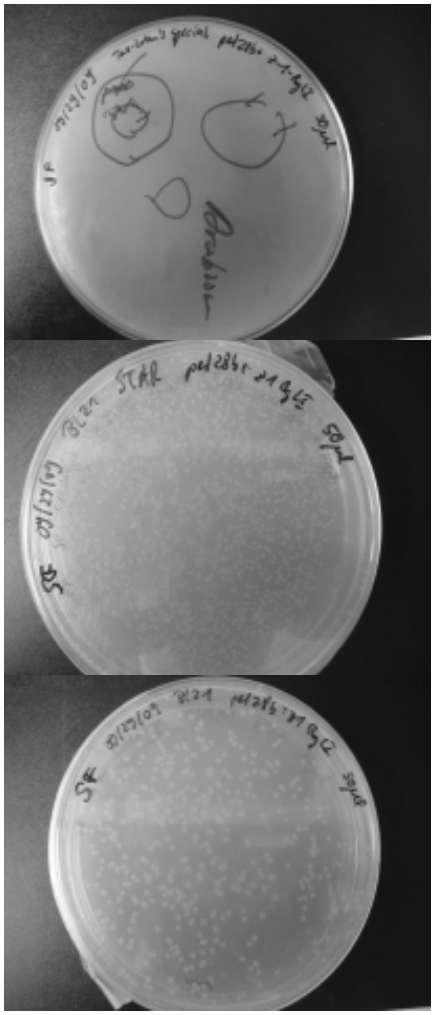


Figure 17 BL21 strain transformants on LB plates

The plates were crowded with colonies of which each one held the genetically modified plasmid pet28b with the gene z1-ßgl.

Compared to the other two plates, the BL21Star (DE3)+GROEL/ES one contained the most colonies. It was even hard to distinguish single dots.

In a next step a small expression study was performed to show evidence of the target protein z1-ßgl. In two steps, an inoculation was prepared: At first 1ml of LB with antibiotics was inoculated with a colony and after a few hours re-inoculated into 5ml of LB with antibiotics. Further steps were carried out according to *2.2.5*.

About 12ml lysate out of each expression was created. Parts of it were used for further analysis, parts of it were stored at -20°C for longer terms.

A SDS-PAGE was run to show evidence of the expressed protein in BL21 strains. The 4-20% Tris-HCl gel was successfully stained and expression levels were sufficiently high enough to proof the presence of z1-ßgl in the insoluble (Figure 18) and soluble fraction (Figure 19).

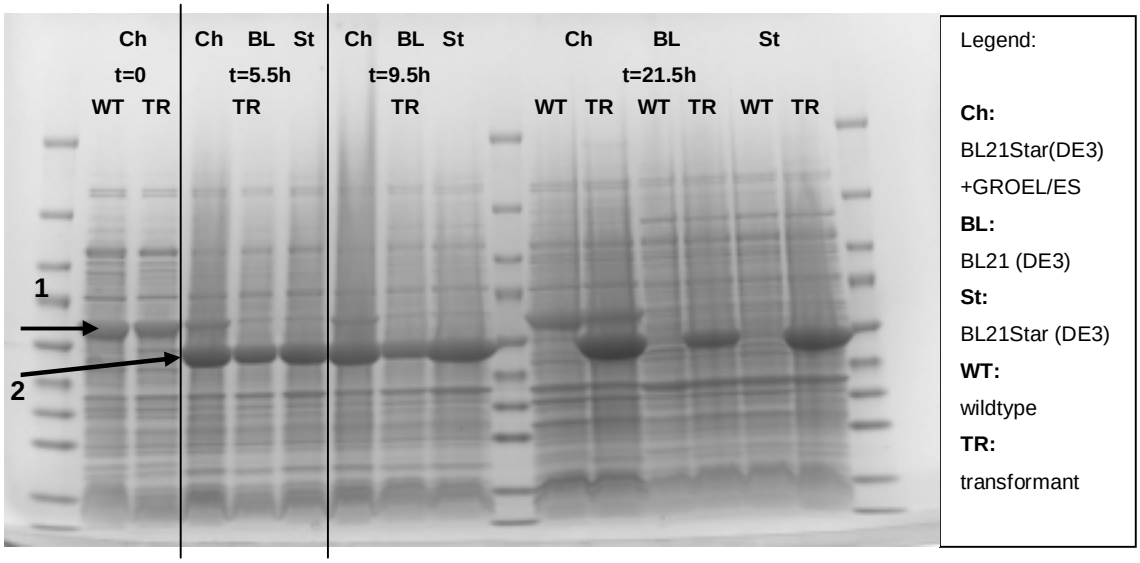


Figure 18 SDS-PAGE: Insoluble fractions of first BL21 strain expressions

The first time point at t=0 was set where the cultures were induced. Two more samples were taken after 5.5h and 9.5h until the expression was finished after 21.5h.

Arrow 1 points at the bands which represents the chaperones. In all BL21Star (DE3)+GROEL/ES strains this band was existing. At t=0 chaperones had already been expressing chaperones which was a good signal. On the other side arrow 2 represents the target enzyme z1-ßgl being misfolded in inclusion bodies. This band is visible in all strains except the wildtypes.

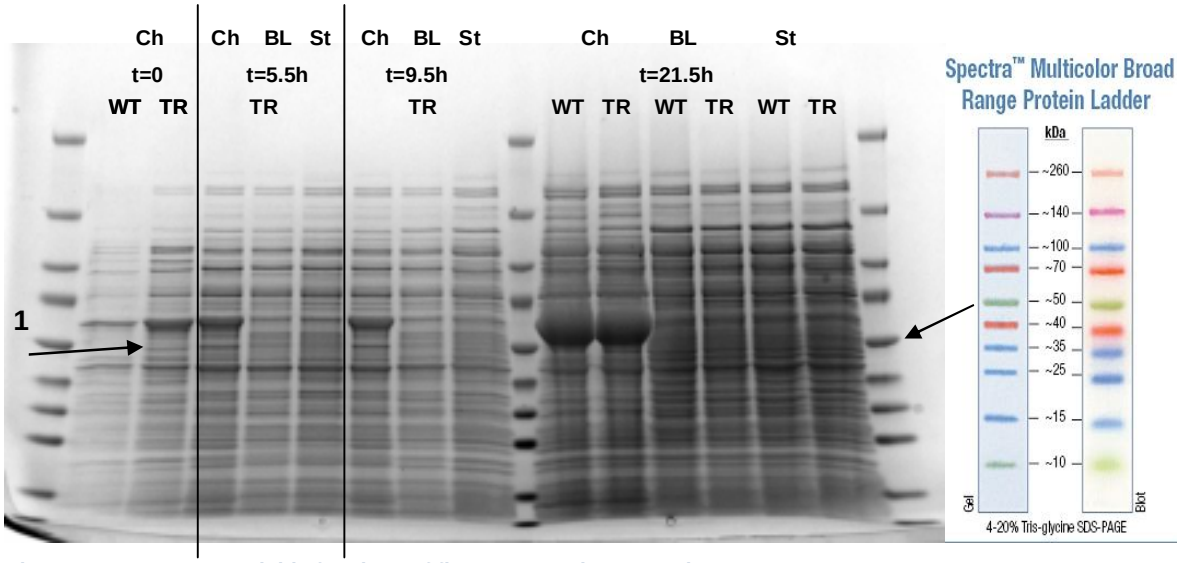


Figure 19 SDS-PAGE: Soluble fractions of first BL21 strain expressions

The soluble fraction of the expression was more important than the insoluble one.

Arrow 1 points here at the soluble z1-ßgl enzyme. An interesting fact is that the band is thicker in the chaperone expressing host than in all the others. BL21 (DE3) and BL21Star (DE3) have an almost fainting band.
These findings led to a more detailed examination of the behavior of the BL21Star (DE3)+GROEL/ES host.

3.3.3. Optimization studies in BL21Star (DE3)+GROEL/ES

3.3.3.1. Impact of lower expression temperatures and different IPTG concentrations

An optimization course was adopted in which the impact of IPTG concentration and temperature on the expression levels of the chaperone expressing BL21 strain were tested.
The expression conditions of *2.2.5* were slightly changed: Instead of using one temperature and a single induction, 3 temperatures and 3 different IPTG concentrations with 9x 14ml tubes were used. Each was filled with 5ml LB/Kan+CAP and 20% L-Arabinose to a final concentration of 0.5%. The media was mixed and inoculated with 100µl of the same glycerol stock. The cultures grew up until OD=1.0, got induced with 0.25mM, 0.5mM and 1mM IPTG and were put on the temperatures 20°C, 25°C and 37°C for 20h. Lysis steps for sonication were carried out according to *2.2.5*. The insoluble (Figure 20) and soluble fraction (Figure 21) were loaded on a 10% SDS-PAGE gel.

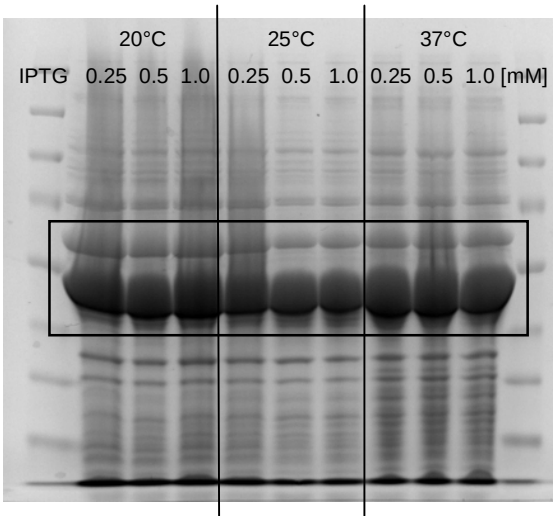


Figure 20 SDS-PAGE: Insoluble fractions of IPTG and temperature optimization study

The labeled area in figure 22 demonstrates the chaperone (upper band row) and protein expression (lower band row). The expression system was absolutely working, although a lot of z1-ßgl enzyme ended up misfolded in the pellet.

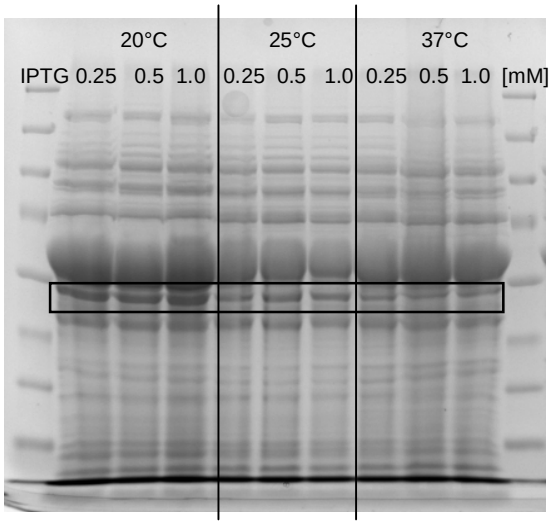


Figure 21 SDS-PAGE: Soluble fractions of IPTG and temperature optimization study

On the contrary, the circled area in figure 23 points out the soluble proteins in the lysate. The most abundant depot was found in the expression samples at 20°C. The higher the temperature was, the lower the expression levels. A difference between the IPTG concentrations couldn't be detected.

An enzyme assay carried out described in *2.2.8* revealed the difference within the different temperatures too (Table 7).

	0,25	0,5	1	0,5	mM IPTG
20C	0,5327	0,4233	0,5105		
25C	0,2749	0,2845	0,1922		
37C	0,2837	0,4193	0,3669		
30C				0,3521	
NegCon	0,3974				

Table 7 Enzyme assay for different temperatures and different IPTG conc.

The 20°C sample was the most active of all. The IPTG concentrations in high ranges didn't make a huge difference within the samples of the same expression temperature. The negative control seemed to be very close to the z1-ßgl transformants. Thus it is assumed that the enzyme activity is rather low.

3.3.3.2. Purification of the optimized expression

For purification the amount of protein was too small so in a next step a large scale expression was performed. 2L of z1-ßgl transformant were set up according to *2.2.5* and then grown up at 37°C until OD=0.8 respectively. The temperature was lowered to 20°C where the cells were developing chaperone proteins. At an OD of 1.3 the culture was induced with 0.05mM IPTG.

The expression finished after 24 hours and the cells were disrupted by french pressing as described in *2.2.5*.

The lysate was analyzed on a SDS-PAGE gel to check the expression level of the target protein (Figure 22).

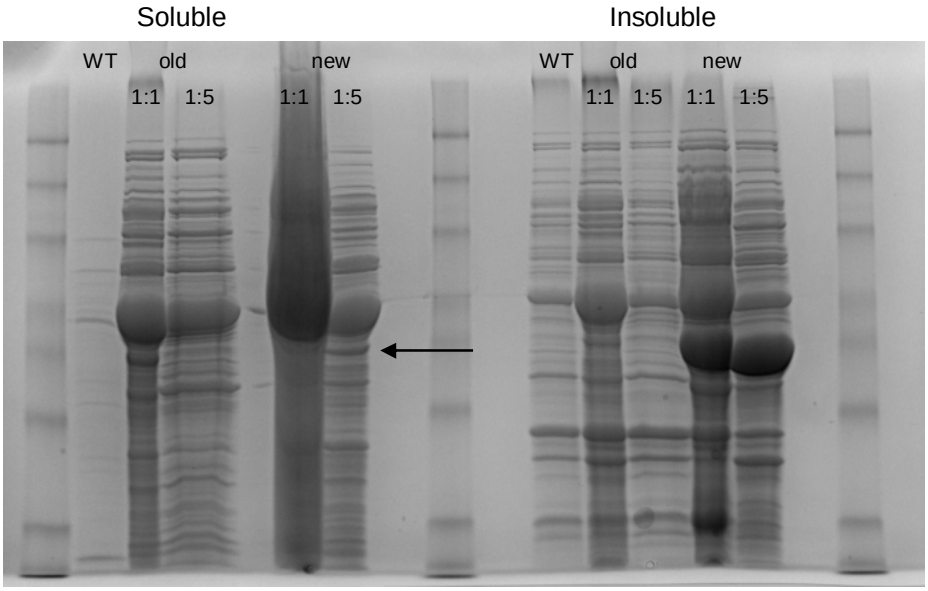


Figure 22 Large Scale expression of z1-ßgl in BL21

The z1-ßgl was well expressed, again more in the insoluble than the soluble fraction. The new 1:1 sample was overloaded, that's why there is a smearing band. The old samples of the insoluble fraction from the last experiment had shown no z1-ßgl band, but it was not important to check the reason for that.

The purification was performed over an Anti-His column according to method I in *2.2.5*.

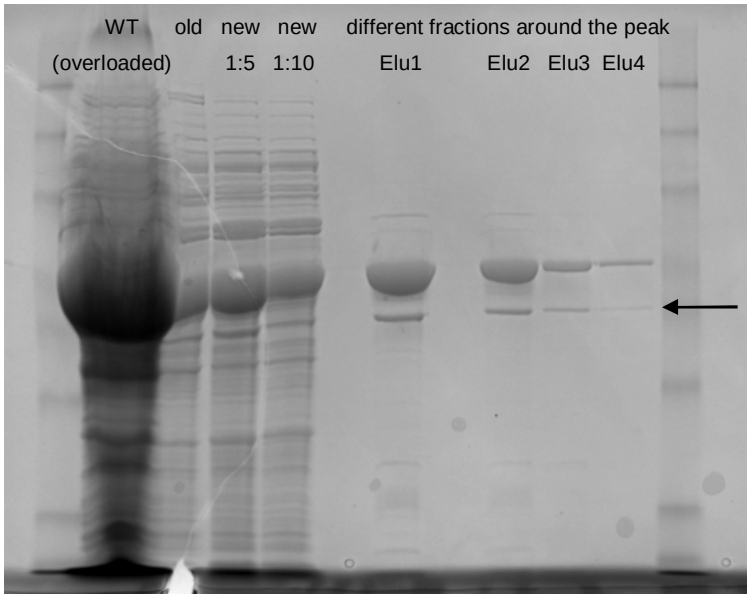


Figure 23 Purification of the his-tagged z1-ßgl in BL21

The fractions were collected around the peak that occurred in the chromatogram and loaded on a SDS-PAGE gel (Figure 23). Elu1 held the most protein while Elu 2, 3 and 4 contained less

and less protein concentrations because these samples were taken in the declining phase. The arrow points out the z1-ßgl enzyme while the bands above represent the chaperone protein.

To purify the samples even more a size exclusion chromatography was introduced. The procedure was performed according to *2.2.9*, although the proteins couldn't be separated (Figure 24).

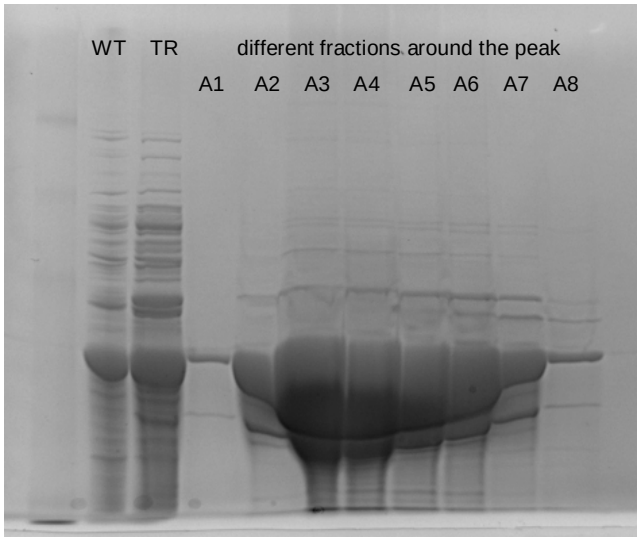


Figure 24 Size exclusion between chaperones and z1-ßgl protein

The elutions were quite overloaded. The chaperones were still remaining in the fractions.

A Western Blot should give precise information about the presence of the z1-ßgl in the *E.coli* lysate. The elutions of the size exclusion chromatography were prepared according to *2.2.7*.
The his-tagged beta-glucosidase was illuminated in all elutions (Figure 25).

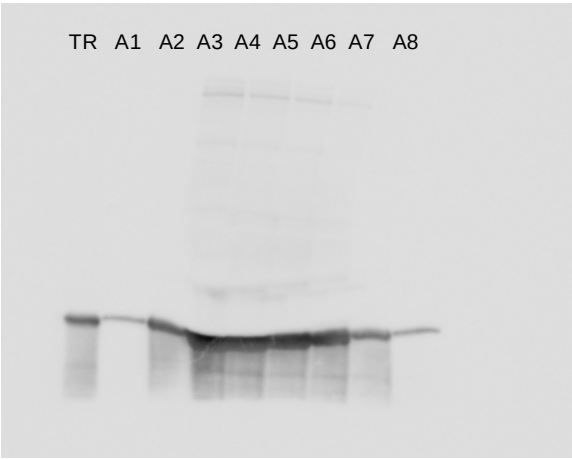


Figure 25 Detecting the z1-ßgl after size exclusion via Western Blot

The ladder was mistaken as a pre-stained marker wouldn't show up in a Western Blot. Although the bands were clear and comparing it to Figure 27 the bands run the same way.

The enzyme activity was tested on all elution fractions according to *2.2.8* and all of them showed a certain activity. Although the numbers were very low and the enzyme seemed to be rather inactive.

3.3.3.3. Inclusion body refolding

In a last experiment the inclusion bodies of the *E.coli* pellet were denatured and refolded as described in *2.2.11*. The enzyme activity was tested again and a positive control of A.niger beta-glucosidase was added, incubated at 50°C and 80°C (Table 8).

neg	0,063			
	at 50C	**at 80C**		
pos	0,7006	0,146		
		Inclusion Body refolding		
	crude	Elu 1	Elu 2	Elu 3
transf	0,085	0,0502	0,0428	0,0363

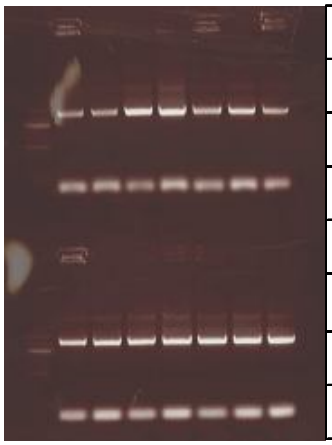
Table 8 Enzyme assay after IB refolding in BL21

Compared with the positive and negative control the z1-ßgl was only expressed in little amounts as the absorbance units were very low. Even the crude enzyme solution exhibits more strength in activity.

3.4. Completion of the vector and expression studies in PichiaPink strains

3.4.1. Completion of the vector pPink-z1-ßgl

The gene for the pichiaPink system was generated in a PCR reaction according to 2.2.1.2 (Figure 26). The annealing temperature was set at 52.7°C.



For Prim MlyI	1,48
3.sfrev1 FseI	1,86
$MgSO_4$	3
10 mM dNTP	1,5
Thermopol Buffer (10X)	5
Deep Vent Polym	0,75
DNA	3
Water	36,41

Figure 26 PCR of the z1-ßgl-pPink gene

These bands were purified according to *2.2.1.3* and double digested with the restriction enzymes *Mly*I and *Fse*I whereas pPinkα-HC plasmid was cut with *Stu*I and *Fse*I as described in *2.2.1.4*.

Two ligation attempts failed until the third one was successful by changing the insert/vector ratio to 9:1. Each ligation mix was prepared twice (Table 9). Two mixes differing in the amount of vector were incubated at 16°C either for 2 hours or overnight.

Mass_vector	30	ng	Mass_vector	60	ng
Length_insert	1290	bp	Length_insert	1290	bp
Length_vector	7900	bp	Length_vector	7900	bp
Insert/Vector Ratio	9	-	Insert/Vector Ratio	9	-
Mass_insert	44,09	ng	Mass_insert	88,18	ng

pet28b + z1-ßgll	**Ligation**	Conc [ng/ul]	**pet28b + z1-ßgll**	**Ligation**	Conc [ng/ul]
Vector	0,43	70,53	Vector	0,85	70,53
Insert	0,60	72,96	Insert	1,21	72,96
Ligation Buffer	1	10	Ligation Buffer	1	10
T4 Ligase	1	-	T4 Ligase	1	-
Water	6,97	-	Water	5,94	-
Total	**10**		**Total**	**10**	

Table 9 Ligation mixes for the pPink system

All 4 ligation reactions were heat inactivated at 65°C for 10min upon completion and immediately transformed into TOP10 Electrocomp *E.coli* and chemically competetent TOP10 *E.coli* as described in *2.2.2.1*. The screening for an error-free clone was carried out according to *2.2.2.2*. Out of 30 electroporated colonies only one turned out to hold the gene and out of 20 chemically transformed colonies also just one contained the gene (Figure 27).

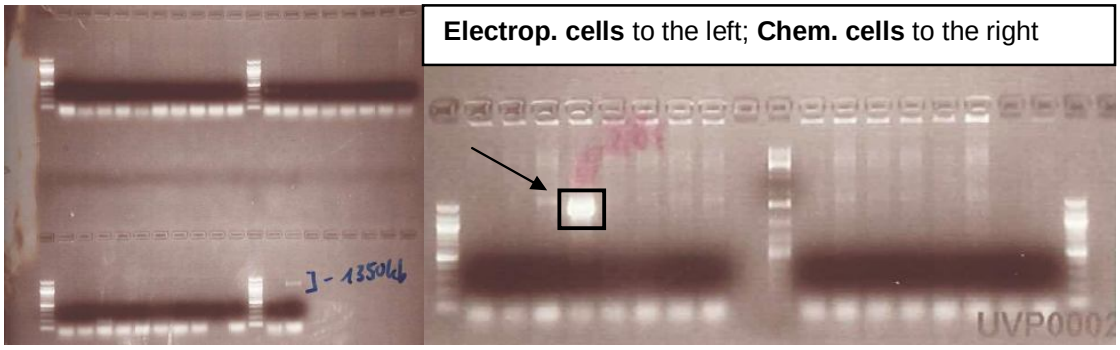


Figure 27 Screening for an error-free clone in the pichiaPink system

The promising clones were re-inoculated for a Miniprep. Then the propagated plasmids were sent for sequencing and the chemically transformed one emerged as the error-free clone. The bases aligned all with the original sequence (Figure 28).

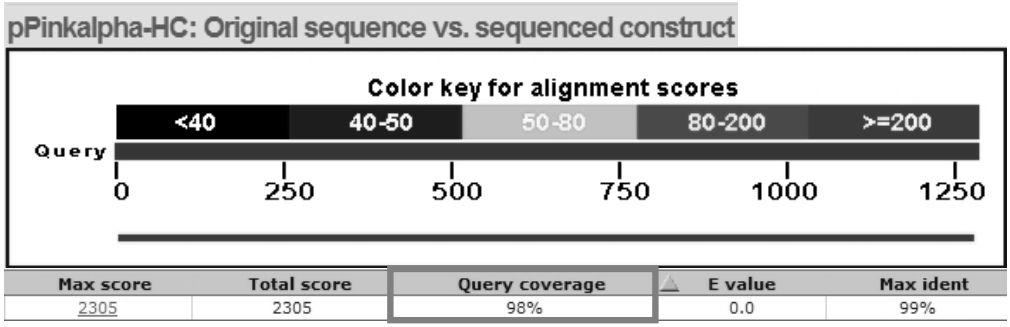


Figure 28 Aligning pPink-z1-ßgl with original sequence

3.4.2. Expression of the z1-ßgl-pPink in *pichia* Pink IV strain

The transformation of the completed vector pPink-z1-ßgl was carried out according to *2.2.4*. All the prearrangements for an electroporation were made. Only the pichia IV strain was used for transformation. 6µl of 500ng/µl linearized plasmid via *Spe*I restriction were added to 80µl of electrocompetent *pichia* IV cells. One transformation was incubated for 2 hours, the other one for 16 hours until they were spread on 4 PAD plates.

After a few days three distinctive white colonies in total occurred on the plates. Red/white selection applied. Those colonies were restreaked on fresh PAD plates and used in the first small scale expression study according to *2.2.6*.

Samples of the secreted proteins (secr) and the lysate supernatant (lys) were loaded on a 10% SDS-PAGE gel (Figure 29).

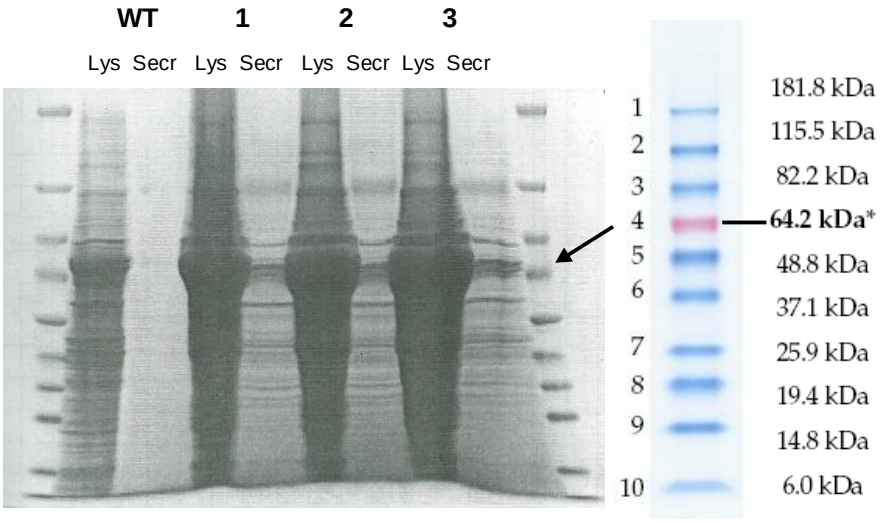


Figure 29 SDS-PAGE: First expression in *pichia* Pink IV

From the first SDS-PAGE it was hard to tell which band was the z1-ßgl in particular.

A Western Blot according to 2.2.7 and an enzyme assay as described in 2.2.8 were performed on the samples of this SDS-PAGE. The Western Blot couldn't detect any bands on the membrane. However, the enzyme assay detected beta-glucosidase activity in the samples, high in the lysate and low in the supernatant (Table 10).

lysate	**WT**	**0,5631**
	1	1,1795
	2	1,2819
	3	1,7147
secreted	**WT**	**0,123**
	1	0,2033
	2	0,1886
	3	0,1956

Table 10 Enzyme assay: First expression in *pichia* Pink IV

To substantiate those findings further a small scale z1-ßgl purification of both fractions was carried out according to Method II of *2.2.9*.

The elutions and some controls were loaded on a 10% SDS-PAGE (Figure 30).

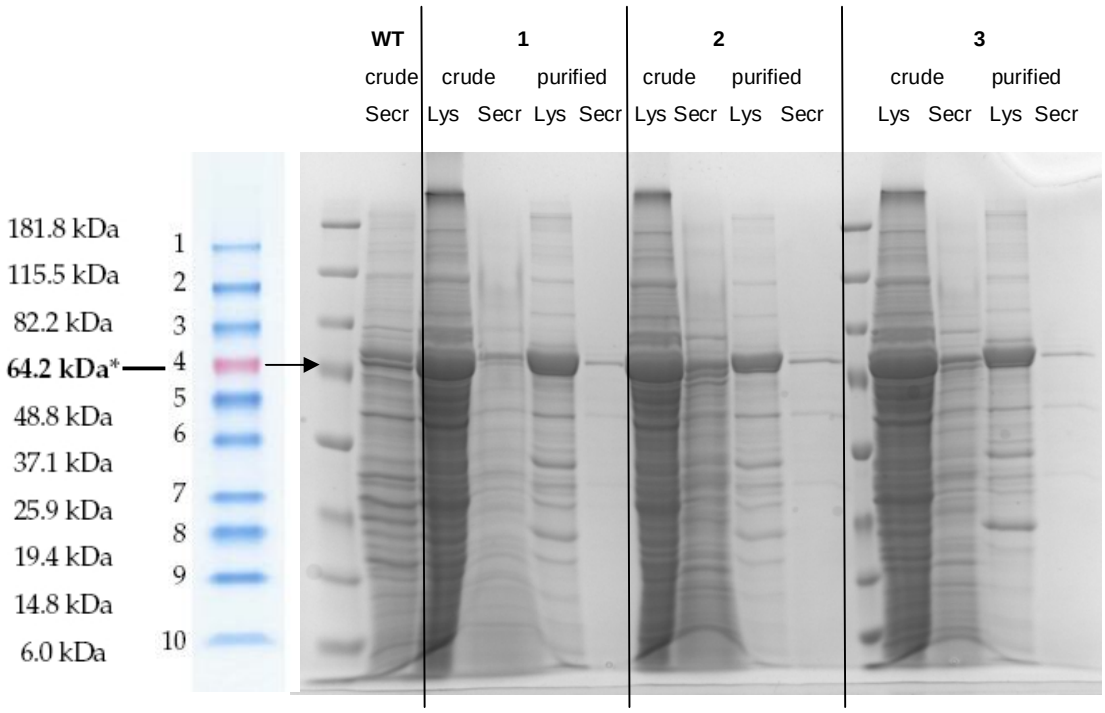


Figure 30 Small scale purification of z1-ßgl in *pichia* Pink IV

Another enzyme assay on the elutions was performed according to *2.2.9*. This one revealed even a lower activity than before. Summing up, the purification was not successful. Some expression optimization studies were performed to increase the protein yield and to check if the protein was expressed intracellularly.

3.4.3. Optimization studies in *pichia* Pink IV

3.4.3.1. Testing the z1-ßgl activity within secreted proteins

First of all three additional media were tested to observe the expression levels in the different environments. The procedures for the expression are described in *2.2.6* with some little adjustments: Five cultures in BMGY where prepared, the wildtype in 100ml, three transformants in 100ml and one transformant in 1L. The transformants came all from one glycerol stock. The cultures were induced at OD=2.0. The wildtype and the 1L transformant were induced in BMGY medium, whereas the three 100ml transformants were changed to MMY, BMM and MM, respectively. The final methanol concentration was increased to 1%.

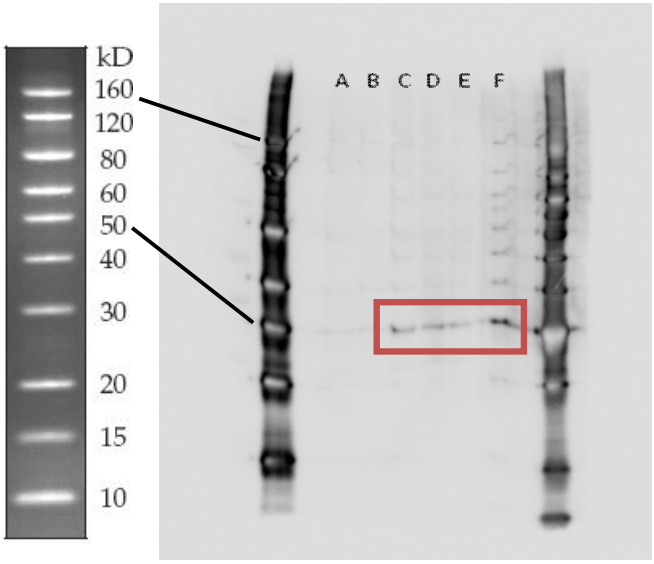
Samples of the secreted fractions were heat treated and concentrated by ammonium sulfate precipitation, both described in *2.2.10*. A fast Bradford assay without standard curve was carried out to tell the difference in protein production between the different media. BMMY and MMY of the transformants had the most, the wildtype in BMMY less, BMM and MM no proteins secreted.

15ml of the BMMY transformant were run over a nickel column according to method I in *2.2.9*. A following Bradford assay the 20-fold concentrated elution revealed no protein (no blue color change).

3.4.3.2. Testing the z1-ßgl activity within intracellular proteins

In a next step the cells were disrupted by French pressing as described in *2.2.9*. An enzyme assay was carried out on the supernatant of the lysates according to *2.2.8*. The supernatant of the MMY lysate emerged as the medium with the highest activity of beta-glucosidase.

In addition to those lysates, the second of the three distinctive transformants was cultured and expressed in MMY to the same conditions as described in *3.4.3.1*. This transformant was compared to the other one in a Western Blot which was performed according to *2.2.7*. The Western Blot included the lysate supernatants of the four cultures in different media, the second transformant in MMY and the wildtype. Four distinctive bands in MMY, BMM, MM and the second transformant did develop (Figure 31).



Legend:

A wildtype
B BMMY
C MMY
D BMM
E MM
F 2nd transformant in MMY

Figure 31 Western Blot: z1-ßgl detection in intracellular fraction of *pichia* Pink IV

To get a clearer picture of the z1-ßgl activity, the protein concentrations of all four media expression lysates were measured via Bradford assay according to *2.2.8*. The same amount of protein was then added in each enzyme reaction. A “no enzyme” and “wildtype” control were included. The enzyme assay was carried out as described in *2.2.8* (Figure 32).

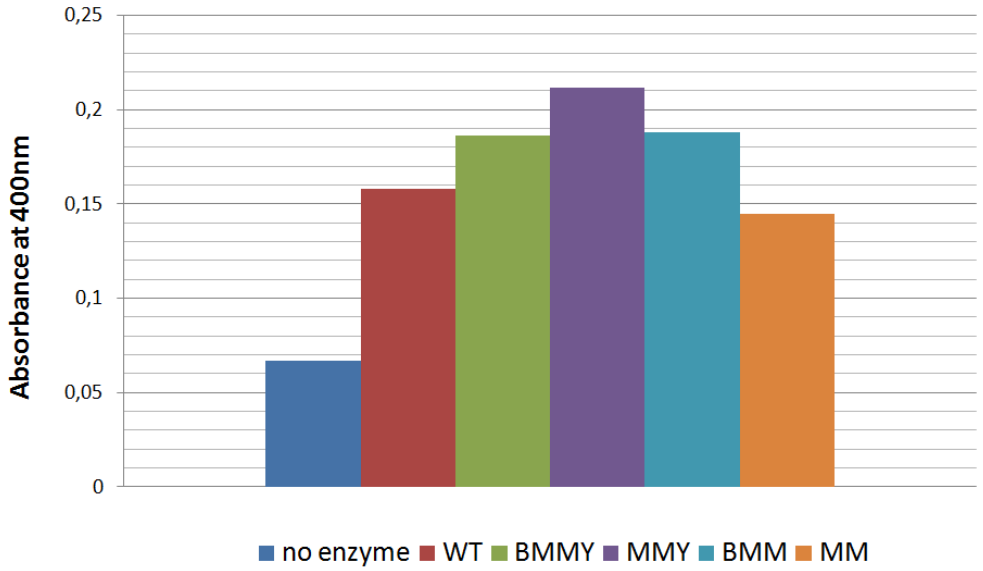


Figure 32 Enzyme Assay: High beta-glucosidase activity in the wildtype

As expected, the MMY media expressed the most active beta-glucosidase. However, the wildtype creates a significant magnitude of beta-glucosidase activity. *P. pastoris* seems to produce an endogenous beta-glucosidase which is released when breaking the cells.

3.4.3.3. Purification of z1-ßgl-pPink under nature and denaturing conditions

In a fresh attempt to purify the z1-ßgl-pPink enzyme the same expression from before was repeated with a 1.5L MMY culture. The cells were disrupted via French press method and the supernatant of the lysate was saved.

1ml of the lysate was used for purification under nature and 6ml for purification under denaturing conditions according to method II and method III in *2.2.9*. The denaturing experiment was done twice, one in the standard way, the other one by adding $MgCl_2$ to 1M final concentration. This should have diminished the affect of EDTA on the nickel beads.

The results were loaded on two separate 4-15% SDS-PAGE gels, one for the nature (Figure 33) and one for the denaturing purification (Figure 34).

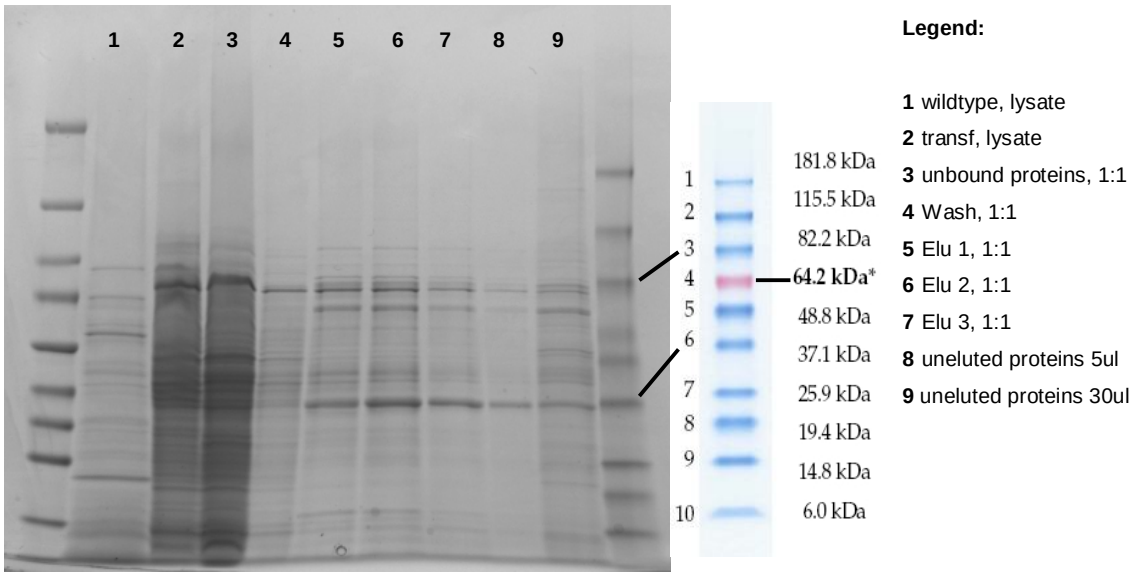


Figure 33 SDS-PAGE: Nature purification of z1-ßgl-pPink

Three proteins seemed to be purified, two over 65kDa, one lower than 37kDa. None of them was lying around 50kDa, which is the size of the fully enlonged z1-ßgl.

Legend:

1	wildtype, lysate	**10**	unbound proteins, 1:1
2	transf, lysate	**11**	Wash 1, 1:1
3	unbound proteins, 1:1	**12**	Wash 7, 1:1
4	Wash 1, 1:1	**13**	Elu 1, 1:1
5	Wash 7, 1:1	**14**	Elu 2, 1:1
6	Elu 2, 1:1	**15**	Elu 3, 1:1
7	Elu 3, 1:1	**16**	uneluted proteins, 1:1
8	Elu 2, 1:1		
9	uneluted proteins, 1:1		

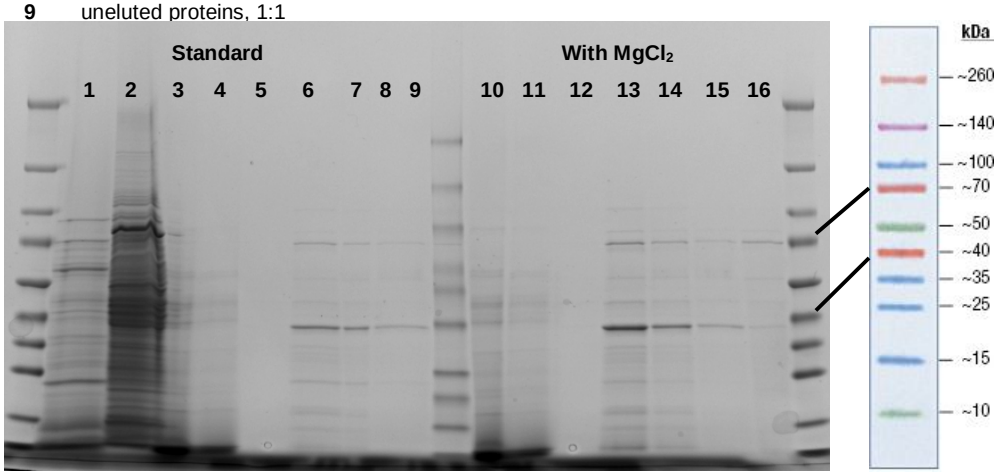


Figure 34 SDS-PAGE: Denaturing purification of z1-ßgl-pPink

Both procedures delivered the same scheme of bands, although the elutions of the one with $MgCl_2$ contained more proteins. The yield was reasonably bigger than in the standard procedure which proofed that EDTA in the Breaking buffer has an influence on the binding capacity of the nickel beads.

4. Discussion

The properties of the thermophilic beta-glucosidase of *Thermus* sp. Z-1 make it a very interesting enzyme which could be applied to cellulose degradation reactions at high temperature. However, it is still a challenge to heterologously produce this enzyme in sufficient amounts. Two fundamentally different host systems – *E.coli* and *P. pastoris* – were tested with different conditions to locate the most appropriate environment for the z1-ßgl.

4.1. Testing the first bacterial expression in OrigamiB

The bacterial gene was especially designed for an expression in OrigamiB via the pet303/CT-His plasmid. During the first expressions it turned out that the vector didn't contain the gene. Before any transformation into an expression host it had to be absolutely paid attention to the correctness of the DNA sequencing. To have a presentiment which plasmid information could be error-free, further controls like vector/colony-PCR were introduced

The re-screening was carried out in a new plasmid propagation system, the One Shot® TOP10 Electrocomp™ *E.coli*. It assures high-efficiency transformation [12] and was used whenever a chemical transformation didn't work out several times

A new expression was carried out with altered IPTG concentrations and expression temperature according to *3.1.3* as it was known that protein folding may improve at lower temperatures like 25°C [13]. Using only 0.2mM IPTG would overexpress the enzyme less so that inclusion bodies would develop only in small scale. However, the changes didn't upgrade the expression quality at all (Figure 6).

4.2. Optimizing the accessibility of the enzyme with chaperones

Out of this reason a more protein solubilizing method was searched. A solution was found in the co-expression of the target enzyme with the chaperonin GroEL/ES. According to Machida et. al [3] the usage of GroEL/ES had a drastic impact on the proper folding of the heterologously produced enzyme. Solubility and enzyme activity would increase.

Three different BL21 strains served as new host for the z1-ßgl. In contrary to the former pet303/CT-His the new plasmid pet28b/NT-His had a N-polyhistidine tag. The exact folding of the z1-ßgl was unknown thus an N-terminal His-tag could increase purification qualities.

The first expression was performed in all three BL21 strains. The overall protein levels were almost the same, but the chaperonin co-expressing host recovered more overexpressed enzyme in the soluble fraction (19).

Optimization studies with the BL21Star(DE3)+GroEL/ES to increase yield and activity of the z1-ßgl were carried out according to *3.3.3*.

In a small scale run the expression temperatures were lowered and the IPTG concentrations altered. As expected the proteins were more stable and actively folded at 20°C, compared to

25°C, 30°C and 37°C [8] [13]. The influence of a change in IPTG concentrations could not be observed.

The optimized expression was repeated in large-scale and the enzyme first purified over a nickel-column. Beside of any expectations the chaperones stuck to the column and were eluted together with the z1-ßgl. A size exclusion chromatography was not able to separate the two proteins (Figure 23) and a size cut-off membrane couldn't be used as the sizes of z1-ßgl and GroEL/ES were almost the same (50kDa and 60kDa). A Western Blot was performed on the elutions to give a picture of the z1-ßgl sources. Since the enzyme was reasonably solubilized further optimizations and different purification techniques would be recommended.

Inclusion body refolding according to *3.3.3.3* was not successful because no activity was measured.

4.3. Testing the expression in *Pichia* pastoris

Parallel to *E.coli* the *P. pastoris* expression system offers attractive levels of a variety of intracellular and extracellular heterologously produced proteins [14]. The kits EasySelect™ Pichia Expression Kit and PichiaPink™ Expression System with the hosts X33 and *pichia* Pink IV were chosen. The red/white selection technique of the PichaPink system and the protease knockout strain IV easies the screening and increases the yield of the produced protein.

Both the pPICZαC and the pPinkα-HC contained the *Saccharomyces cerevisiae* α-mating factor for secreted protein expression. Any gene that is being ligated into the vector with correct framing should pass the endoplasmatic reticulum and be secreted into the medium.

4.3.1. Findings in the X33 strain

On the first SDS-Pages after the X33 expression the bands could be barely seen, because the samples were diluted too much. In selecting the best expression media the MMY medium emerged as the one with the highest protein concentration. However, a Western Blot couldn't detect any sources of the z1-ßgl in both the secreted and lysed fraction (Figure 11). The colonies picked were presumably transformed with an empty plasmid. More investigations were omitted because the PichiaPink system from Invitrogen was more advanced and recommended by the laboratory.

4.3.2. Testing the expression in the PichiaPink system

The screening for an error-free clone was performed in chemically- and electrocompetent TOP10 *E.coli* cells. From all screened colonies just one turned out to hold the gene in frame.

The first expression results were poor as the SDS-PAGE (Figure 29) and the Western Blot couldn't state the presence of the enzyme. An enzyme assay was showing activity of beta-glucosidase though. Even if the transformant samples had a higher activity, the lysate of the

wildtype still had a huge background. This attributes to an endogenous beta-glucosidase which is set free after lysing the cells [15].
A small-scale purification step of the first expression was decreasing the overall activity almost to a zero-level. The endogenous beta-glucosidase was washed out and the recombinant z1-ßgl had not been expressed in sufficient amounts.
The first large-scale expression was aiming for the secreted proteins in different media. Various treatments were applied but none of these could show an evidence of z1-ßgl. The purification was unsuccessful as the Bradford assay indicated no blue color change.
In the next step the target fraction was changed and new cultures were lysed and purified. From the SDS-Page it was impossible to point at the z1-ßgl band. So a Western Blot was performed on all media samples and an old freezer stock (Figure 31). The 50kDa band showed up in some of the media but not BMMY. As expected the wildtype wasn't showing any band either. Thus a bigger expression culture is essential to create at least a detectable amount of this enzyme.
Purifications over a nickel column had never resulted in a pNPG active elution. Sometimes the enzyme folds in a way so that the ligand can't access the His-tag. Therefore a denaturing step was carried out to break up the structure of the enzyme and then slowly refold it to its nature state. That experiment was not having any difference to the nature purification.
It was very strange to see the Western Blot detecting the enzyme but the purification never worked.

5. Conclusion

To sum up, the project's goal was to express the thermophilic beta-glucosidase from *Thermus* sp. Z-1 in its active status. Two systems were available, bacterial and eukaryotic. While using *Pichia* pastoris it was hoped that the protein was secreted, but it wasn't. The expression yield of the enzyme in both systems was rather low and so was the activity of the enzyme. A co-expression with the chaperonin GroEL/ES in *E.coli* created more soluble proteins which were more active. However, more optimizations of the expression and purification are highly recommended so a substantial amount of z1-ßgl enzyme can be used for kinetic studies.

References

[1] Saha B.C. (2003) Hemicellulose bioconversion. J Ind Microbiol Biotechnol 30:279-91.

[2] Takase M., Horikoshi K. (1988) A thermostable β-glucosidase isolated from a bacterial species of the genus *Thermus* Springer Berlin 29:55-60.

[3] Takase M., Horikoshi K. (1989) Purification and Properties of a β-Glucosidase from *Thermus* sp. Z-1. Biosci Biotechnol Biochem 53:559-560.

[4] Akiyama K., Takase M., Horikoshi K., Okonogi S. (2001) Production of Galactooligosaccharides from Lactose Using a β-Glucosidase from *Thermus* sp. Z-1. biosci Biotechnol Biochem 65:438-441.

[5] Cereghino J.L., Cregg J.M. (2000) Heterologous protein expression in the methylotrophic yeast Pichia pastoris. FEMS Microbiol Rev 24:45-66.

[6] Aoki H., Nazmul Ahsan M., Watabe S. (2003) Heterologous expression in Pichia pastoris and single-step purification of a cysteine proteinase from northern shrimp. Protein Expr Purif 31:213-21.

[7] Tilly K., VanBogelen R.A., Georgopoulos C., Neidhardt F.C. (1983) Identification of the heat-inducible protein C15.4 as the groES gene product in Escherichia coli. J Bacteriol 154:1505-7.

[8] Machida S., Yu Y., Singh S.P., Kim J.D., Hayashi K., Kawata Y. (1998) Overproduction of beta-glucosidase in active form by an Escherichia coli system coexpressing the chaperonin GroEL/ES. FEMS Microbiol Lett 159:41-6.

[9] Laemmli U.K. (1970) Cleavage of structural proteins during the assembly of the head of bacteriophage T4. Nature 227:680-5.

[10] Towbin H., Staehelin T., Gordon J. (1992) Electrophoretic transfer of proteins from polyacrylamide gels to nitrocellulose sheets: procedure and some applications. 1979. Biotechnology 24:145-9.

[11] Aminlari M., Vaseghi T. (1987) A new colorimetric method for determination of creatine phosphokinase. Anal Biochem 164:397-404.

[12] Dower W.J., Miller J.F., Ragsdale C.W. (1988) High efficiency transformation of E. coli by high voltage electroporation. Nucleic Acids Res 16:6127-45.

[13] Broeze R.J., Solomon C.J., Pope D.H. (1978) Effects of low temperature on in vivo and in vitro protein synthesis in Escherichia coli and Pseudomonas fluorescens. J Bacteriol 134:861-74.

[14] Cregg J.M., Vedvick T.S., Raschke W.C. (1993) Recent advances in the expression of foreign genes in Pichia pastoris. Biotechnology (N Y) 11:905-10.

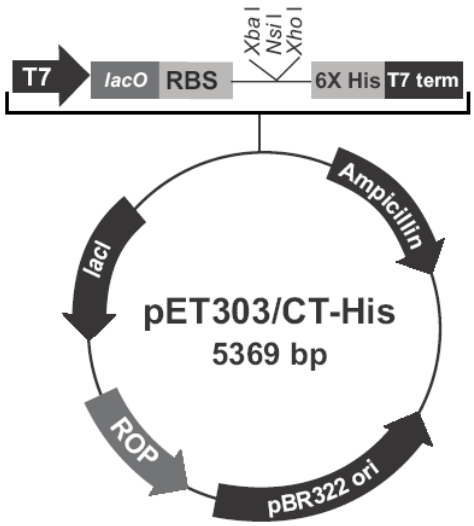
[15] Turan Y., Zheng M. (2005) Purification and characterization of an intracellular beta-glucosidase from the methylotrophic yeast Pichia pastoris. Biochemistry (Mosc) 70:1363-8.

List of figures

Index of Tables

Attachments

- Pet303/CT-His plasmid map and polylinker



Comments for pET303 CT-His
5369 nucleotides

T7 promoter: bases 20-36
T7 promoter priming site: bases 20-39
lac operator (*lac*O): bases 39-63
Ribosome binding site (RBS): bases 95-100
6X His Tag: bases 119-136
T7 reverse priming site: bases 186-206
T7 transcription termination region: bases 147-277
F1 origin: bases 287-742
bla promoter: bases 775-879
Ampicillin (*bla*) resistance gene: bases 874-1734
pBR322 origin: bases 1945-2678 (c)
ROP ORF: bases 2920-3011 (c)
lacI ORF: bases 3914-5032 (c)

(c) = complementary strand

T7 promoter/priming site — Lac operator

```
11  CCGCGAAATT AATACGACTC ACTATAGGGG AATTGTGAGC GGATAACAAT TCCCCTCTAG TAATAATTTT
    GGCGCTTTAA TTATGCTGAG TGATATCCCC TTAACACTCG CCTATTGTTA AGGGGAGATC ATTATTAAAA
```

Ribosome binding site — *Xba*I — *Nsi*I *Xho*I — 6X His Tag

```
81  GTTTAACTTT AAGAAGGAGG TCTAGAATGC ATCTCGAGCA CCACCACCAC CACCACTGAG ATCCGGCTGC
    CAAATTGAAA TTCTTCCTCC AGATCTTACG TAGAGCTCGT GGTGGTGGTG GTGGTGACTC TAGGCCGACG
```

T7 terminator — T7 reverse priming site

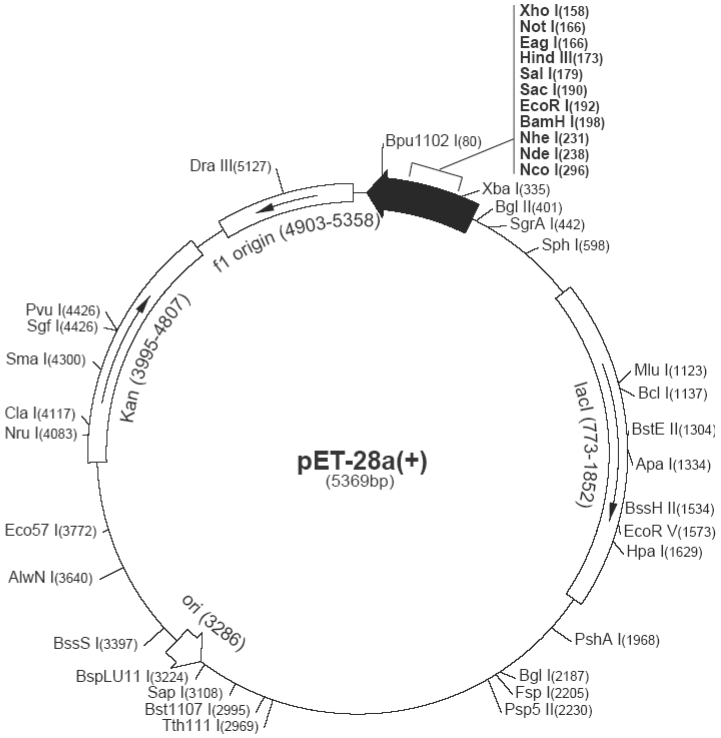
```
151 TAACAAAGCC CGAAAGGAAG CTGAGTTGGC TGCTGCCACC GCTGAGCAAT AACTAGCATA ACCCCTTGGG
    ATTGTTTCGG GCTTTCCTTC GACTCAACCG ACGTCGGTGG CGACTCGTTA TTGATCGTAT TGGGGAACCC
```

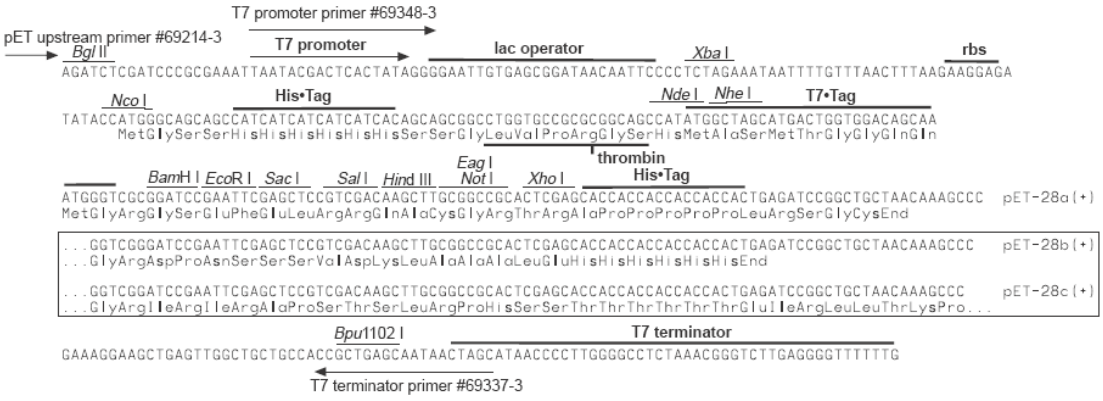
```
221 GCCTCTAAAC GGGTCTTGAG GGGTTTTTTG CTGAAAGGAG GAACTATATC CGGATTGGCG AATGGGACGC
    CGGAGATTTG CCCAGAACTC CCCAAAAAAC GACTTTCCTC CTTGATATAG GCCTAACCGC TTACCCTGCG
```

- pet28b/NT-His plasmid map and polylinker

pET-28a(+) sequence landmarks	
T7 promoter	370-386
T7 transcription start	369
His•Tag coding sequence	270-287
T7•Tag coding sequence	207-239
Multiple cloning sites	
(*Bam*H I - *Xho* I)	158-203
His•Tag coding sequence	140-157
T7 terminator	26-72
lacI coding sequence	773-1852
pBR322 origin	3286
Kan coding sequence	3995-4807
f1 origin	4903-5358

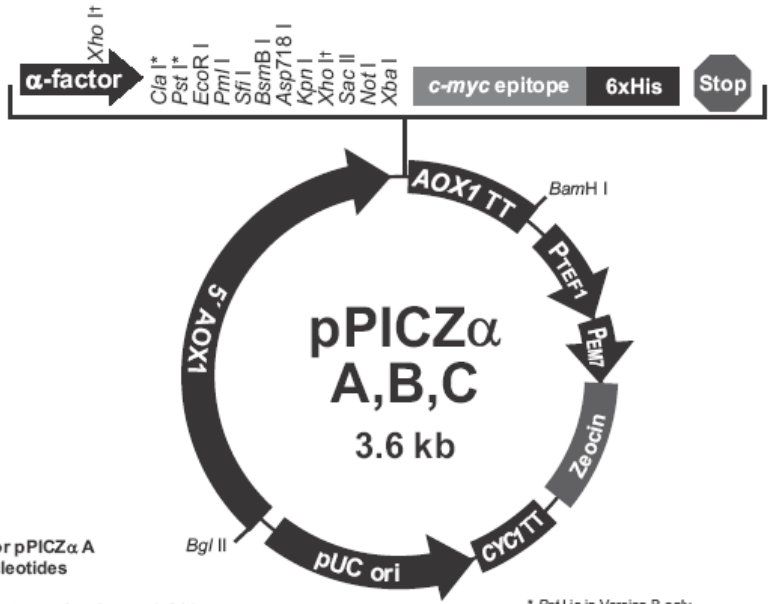
The maps for pET-28b(+) and pET-28c(+) are the same as pET-28a(+) (shown) with the following exceptions: pET-28b(+) is a 5368bp plasmid; subtract 1bp from each site beyond *Bam*H I at 198. pET-28c(+) is a 5367bp plasmid; subtract 2bp from each site beyond *Bam*H I at 198.





pET-28a-c(+) cloning/expression region

- pPICZαC plasmid map and polylinker



Comments for pPICZα A
3593 nucleotides

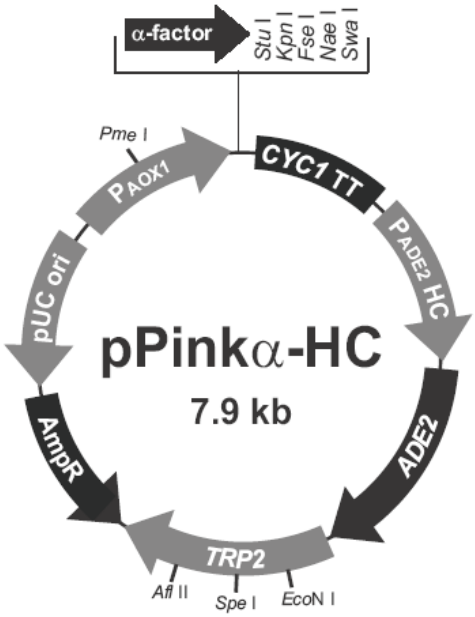
5´ *AOX1* promoter region: bases 1-941
5´ *AOX1* priming site: bases 855-875
α-factor signal sequence: bases 941-1207
Multiple cloning site: bases 1208-1276
c-myc epitope: bases 1275-1304
Polyhistidine (6xHis) tag: bases 1320-1337
3´ *AOX1* priming site: bases 1423-1443
AOX1 transcription termination region: bases 1341-1682
TEF1 promoter: bases 1683-2093
EM7 promoter: bases 2095-2162
Sh ble ORF: bases 2163-2537
CYC1 transcription termination region: bases 2538-2855
pUC origin: bases 2866-3539 (complementary strand)

* *Pst* I is in Version B only
Cla I is in Version C only

†The two *Xho* I sites in the vector allow the user to clone their gene in frame with the Kex2 cleavage site, resulting in expression of their native gene without additional amino acids at the N-terminus.

```
                5´ end of AOX1 mRNA                                  5´ AOX1 priming site
811  AACCTTTTTT TTTATCATCA TTATTAGCTT ACTTTCATAA TTGCGACTGG TTCCAATTGA

871  CAAGCTTTTG ATTTTAACGA CTTTTAACGA CAACTTGAGA AGATCAAAAA ACAACTAATT

931  ATTCGAAACG ATG AGA TTT CCT TCA ATT TTT ACT GCT GTT TTA TTC GCA GCA
                Met Arg Phe Pro Ser Ile Phe Thr Ala Val Leu Phe Ala Ala

983  TCC TCC GCA TTA GCT GCT CCA GTC AAC ACT ACA ACA GAA GAT GAA ACG GCA
     Ser Ser Ala Leu Ala Ala Pro Val Asn Thr Thr Thr Glu Asp Glu Thr Ala

                                  α-factor signal sequence
1034 CAA ATT CCG GCT GAA GCT GTC ATC GGT TAC TCA GAT TTA GAA GGG GAT TTC
     Gln Ile Pro Ala Glu Ala Val Ile Gly Tyr Ser Asp Leu Glu Gly Asp Phe

1085 GAT GTT GCT GTT TTG CCA TTT TCC AAC AGC ACA AAT AAC GGG TTA TTG TTT
     Asp Val Ala Val Leu Pro Phe Ser Asn Ser Thr Asn Asn Gly Leu Leu Phe

                                                                    Xho I*
1136 ATA AAT ACT ACT ATT GCC AGC ATT GCT GCT AAA GAA GAA GGG GTA TCT CTC
     Ile Asn Thr Thr Ile Ala Ser Ile Ala Ala Lys Glu Glu Gly Val Ser Leu

          Kex2 signal cleavage                Cla I   EcoR I   Pml I        Sfi I          BsmB I
1187 GAG AAG AGA GAG GCT GAA GC ATCGAT GAATTCAC GTGGCCCAG CCGGCCGTC TCGGA
     Glu Lys Arg Glu Ala Glu Ala
                      Ste13 signal cleavage
     Asp718 I Kpn I Xho I    Sac II  Not I            Xba I          c-myc epitope
1244 TCGGTACCTC GAGCCGCGGC GGCCGCCAGC TTTCTA GAA CAA AAA CTC ATC TCA GAA
                                               Glu Gln Lys Leu Ile Ser Glu
                                                    polyhistidine tag
1301 GAG GAT CTG AAT AGC GCC GTC GAC CAT CAT CAT CAT CAT CAT TGA GTTTGTA
     Glu Asp Leu Asn Ser Ala Val Asp His His His His His His ***

1353 GCCTTAGACA TGACTGTTCC TCAGTTCAAG TTGGGCACTT ACGAGAAGAC CGGTCTTGCT

                          3´ AOX1 priming site
1413 AGATTCTAAT CAAGAGGATG TCAGAATGCC ATTTGCCTGA GAGATGCAGG CTTCATTTTT

                                                  3´ polyadenylation site
1473 GATACTTTTT TATTTGTAAC CTATATAGTA TAGGATTTTT TTTGTCATTT TGTTTCTTCT
```

- pPinkα-HC plasmid map and polylinker



Comments for pPinkα-HC
7898 nucleotides

5´ *AOX1* promoter region: bases 1-940
α-mating factor secretion signal: 941-1194
Multiple cloning site: bases 1196-1222
CYC1 transcription termination region: bases 1228-1502
ADE2 promoter region: bases 1517-1529
ADE2 ORF: bases 1530-3215
TRP2 gene: bases 3439-5250
pUC ori: 5619-6292 (c)*
Ampicillin (*bla*) resistance gene: bases 6437-7297 (c)

*(c): complementary strand

```
                    5´ end of AOX1 mRNA                                  5´ AOX1 priming site
811 AACCTTTTTT TTTATCATCA TTATTAGCTT ACTTTCATAA TTGCGACTGG TTCCAATTGA

871 CAAGCTTTTG ATTTTAACGA CTTTTAACGA CAACTTGAGA AGATCAAAAA ACAACTAATT

                    ATG AGA TTT CCT TCA ATT TTT ACT GCT GTT TTA TTC GCA
931 ATTCGAAACG      Met Arg Phe Pro Ser Ile Phe Thr Ala Val Leu Phe Ala

980 GCA TCC TCC GCA TTA GCT GCT CCA GTC AAC ACT ACA ACA GAA GAT GAA
    Ala Ser Ser Ala Leu Ala Ala Pro Val Asn Thr Thr Thr Glu Asp Glu

                                α-factor signal sequence
1028 ACG GCA CAA ATT CCG GCT GAA GCT GTC ATC GGT TAC TCA GAT TTA GAA
     Thr Ala Gln Ile Pro Ala Glu Ala Val Ile Gly Tyr Ser Asp Leu Glu

1076 GGG GAT TTC GAT GTT GCT GTT TTG CCA TTT TCC AAC AGC ACA AAT AAC
     Gly Asp Phe Asp Val Ala Val Leu Pro Phe Ser Asn Ser Thr Asn Asn

                                        α-factor priming site
1124 GGG TTA TTG TTT ATA AAT ACT ACT ATT GCC AGC ATT GCT GCT AAA GAA
     Gly Leu Leu Phe Ile Asn Thr Thr Ile Ala Ser Ile Ala Ala Lys Glu

                                  Stu I          Kpn I    Nae I       Swa I
1172 GAA GGG GTA TCT CTC GAG AAA AGG CCTCAGGTAC CGGCCGGCCATT TAAATACAGG
     Glu Gly Val Ser Leu Glu Lys Arg

                          CYC1 transcription termination region
1228 CCCCTTTTCC TTTGTCGATA TCATGTAATT AGTTATGTCA CGCTTACATT CACGCCCTCC
```

- Gesncript optimized beta-glucosidase sequence from Thermus sp. Z-1

GenBank AB034947.1

Gene: `Thermus sp. Z-1 gene for beta-glucosidase`

Note: This sequence has been gene optimized by Genscript! Restriction sites are in bold

http://www.uniprot.org/uniprot/Q9RA58

```
   1 tctagaatgaccgaga acgccgaaaa attcctttgg ggagtggcca ccagcgccta ccagattgag
  61 ggggccaccc aggaggacgg ccgggggcct tccatctggg acgccttcgc ccagcgcccg
 121 ggggccatcc gggacgggag cacaggggag cccgcctgcg accactaccg ccgctacgag
 181 gaggacatcg ccctgatgca atccctcggg gtgggggcct accgcttctc cgtggcctgg
 241 ccccggatcc tccccgaggg ccgggggcgg atcaacccca agggcctcgc cttctacgac
 301 cgcctggtgg accggcttct cgcttccggg atcacgccct ttctcaccct ctaccactgg
 361 gacctgcctt tggcccagga ggagcgggga ggctggcgga gccgggagac cgccttcgcc
 421 ttcgccgagt acgccgaggc ggtggcccgg gccctcgccg accggttgcc cttcttcgcc
 481 accctgaacg agccctggtg ctcggccttc ctcgggcact ggacgggggа acacgccccc
 541 ggcctcagga acctggaagc ggccctccgc gccgcccacc acctcctcct gggccacggc
 601 ctcgccgtgg aggccttgag ggccgcgggg gcgaagcggg tggggatcgt cctcaacttc
 661 gccccgcgct acggcgagga ccccgaggcg gtggacgtgg ccgaccgcta ccacaaccgc
 721 ttcttcctgg accccatcct gggcaagggg tatcccgaaa gcccccttccg agacccccccg
 781 cccgtcccca tcctctcccg cgacctggag ctcgtggcaa ggcccctgga cttcctgggg
 841 gtgaactact acgccccccgt ccgcgtggcc ccgggggacgg ggaccttgcc cgtgcgctac
 901 cttcccccgg aagggccggc cacggccatg gggtgggagg tctaccccga ggggctttac
 961 cacctcttga agcgcctcgg ccgggaggtg ccctggcccc tttacgtcac ggaaaacggg
1021 gccgcctacc ccgacctctg gacgggagag gccgtggtgg aggaccccga gcgggtggcc
1081 tacctcgagg cccacgtgga gggccgcctc cgggcccggg aagaaggggt ggacctccgg
1141 ggctacttcg tctggagcct catggacaac tttgagtggg cgttcggcta cacccggcgc
1201 ttcggcctct actacgtgga cttccccagc cagaggcgca tccccaaaag gagcgccctc
1261 tggtaccaactcgag
```