HANNA ARONSSON

Construction of novel plasmid system for efficient co-expression of ADH-A, StEH1 and GroEL/ES

Degree project C in chemistry

May 2018

Department of chemistry – BMC
Supervisor: Prof. Mikael Widersten
Abstract

The enzymes StEH1 and ADH-A have been insufficiently co-expressed previously and the aim of this project was to construct a new plasmid system to improve the expression. These enzymes are interesting to combine since they can produce 2-hydroxyacetophenone from racemic styrene oxide. The cDNA for StEH1 was cut out from pGTacStEH1 and successfully inserted into pGT7ADHC1B1 with digestion followed by ligation. The constructed plasmid was transformed into the BL-21-AI strain of E. coli for expression and purification. SDS-PAGE analysis suggested that ADH-A, which depends on chaperones for proper folding, instead formed inclusion bodies.
### Abbreviations

<table>
<thead>
<tr>
<th>Abbreviation</th>
<th>Definition</th>
</tr>
</thead>
<tbody>
<tr>
<td>ADH</td>
<td>alcohol dehydrogenase</td>
</tr>
<tr>
<td>bp</td>
<td>base pairs</td>
</tr>
<tr>
<td>BSA</td>
<td>bovine serum albumin</td>
</tr>
<tr>
<td>cDNA</td>
<td>complementary deoxyribonucleic acid</td>
</tr>
<tr>
<td>DNA</td>
<td>deoxyribonucleic acid</td>
</tr>
<tr>
<td>dNTP</td>
<td>deoxynucleoside triphosphate</td>
</tr>
<tr>
<td>E. coli</td>
<td><em>Escherichia coli</em></td>
</tr>
<tr>
<td>EDTA</td>
<td>ethylenediaminetetraacetic acid</td>
</tr>
<tr>
<td>IMAC</td>
<td>immobilized-metal affinity chromatography</td>
</tr>
<tr>
<td>IPTG</td>
<td>isopropyl β-D-1-thiogalactopyranoside</td>
</tr>
<tr>
<td>LB</td>
<td>lysogeny broth</td>
</tr>
<tr>
<td>Mw</td>
<td>Molecular weight</td>
</tr>
<tr>
<td>NAD(P)</td>
<td>nicotinamide adenine dinucleotide (phosphate)</td>
</tr>
<tr>
<td>OD&lt;sub&gt;600&lt;/sub&gt;</td>
<td>optical density at 600 nm</td>
</tr>
<tr>
<td>PAGE</td>
<td>polyacrylamide gel electrophoresis</td>
</tr>
<tr>
<td>rpm</td>
<td>revolutions per minute</td>
</tr>
<tr>
<td>SDS</td>
<td>sodium dodecyl sulfate</td>
</tr>
<tr>
<td>StEH1</td>
<td><em>Solanum tuberosum</em> epoxide hydrolase 1</td>
</tr>
<tr>
<td>TAE</td>
<td>Tris, acetic acid, EDTA</td>
</tr>
<tr>
<td>Tris</td>
<td>tris(hydroxymethyl)aminomethane</td>
</tr>
<tr>
<td>wt</td>
<td>wild type</td>
</tr>
</tbody>
</table>
## Contents

<table>
<thead>
<tr>
<th>Section</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>Introduction</td>
<td>5</td>
</tr>
<tr>
<td>Biocatalysis and synthetic biology</td>
<td>5</td>
</tr>
<tr>
<td>Plasmids and chaperones</td>
<td>5</td>
</tr>
<tr>
<td>Restriction enzymes</td>
<td>6</td>
</tr>
<tr>
<td>Epoxide hydrolases</td>
<td>7</td>
</tr>
<tr>
<td>Alcohol dehydrogenases</td>
<td>8</td>
</tr>
<tr>
<td>Project overview</td>
<td>10</td>
</tr>
<tr>
<td>Methods</td>
<td>10</td>
</tr>
<tr>
<td>Transformation for plasmid preparation</td>
<td>10</td>
</tr>
<tr>
<td>Digestion</td>
<td>11</td>
</tr>
<tr>
<td>Protein expression and purification</td>
<td>12</td>
</tr>
<tr>
<td>Results</td>
<td>13</td>
</tr>
<tr>
<td>Digestion</td>
<td>14</td>
</tr>
<tr>
<td>Protein expression and purification</td>
<td>15</td>
</tr>
<tr>
<td>Discussion</td>
<td>16</td>
</tr>
<tr>
<td>Digestion</td>
<td>16</td>
</tr>
<tr>
<td>Protein expression and purification</td>
<td>16</td>
</tr>
<tr>
<td>Conclusion</td>
<td>17</td>
</tr>
<tr>
<td>Acknowledgements</td>
<td>17</td>
</tr>
<tr>
<td>References</td>
<td>18</td>
</tr>
</tbody>
</table>
**Introduction**

*Biocatalysis and synthetic biology*

Biocatalysis uses enzymes and microbes for synthetic chemistry, which is a more environmentally friendly approach compared to for example organocatalysis. The area has had a big development over the last century: in the beginning it was found that parts of living cells could be used for chemical reactions and now biocatalysis, promoted by the parallel development of protein engineering, can be used for various applications. With protein engineering, industries are able to use enzymes that are customized for a specific applied purpose, instead of the method being customized to fit the enzyme. [1, 2] Protein engineering started out with site-directed mutagenesis, where certain amino acids could be changed. It then developed into techniques where changes of loops or whole domains was possible. [3]

There are different strategies that can be used for protein engineering and they can either be random, focused or mimic the natural system of mutagenesis. Random mutations could be obtained by chemical mutagenesis where the DNA is damaged by chemicals, by mutator strains where cells with disabled proofreading are used, or by error prone PCR where a low fidelity DNA polymerase introduces mutations. Focused mutations could be obtained by site-directed saturation mutagenesis where restriction enzyme cloning is used and advantageous mutations can be predicted by molecular modelling. For mimicking of natural systems, they could be obtained by homologous recombination where different methods could be used, such DNA shuffling, or by non-homologous recombination where sequence homology-independent protein recombination (SHIPREC) is used. In DNA shuffling, a DNase is used to fragmentize the gene and then the fragments randomly prime each other in PCR to produce mutations and in SHIPREC the shuffling can be done to recombine genes from different protein families, but that have similarities. [4]

Synthetic biology is used to engineer biological systems to make them perform new tasks. It is used for several applications such as reprogramming signaling pathways and metabolic engineering. Metabolic engineering is about optimizing desired cellular activities by changing some functions of enzymes or in the regulatory systems. Reactions in the cells are targeted so that the cells will produce molecules that could be useful for different applications. [5-7]

*Plasmids and chaperones*

Plasmids are double stranded DNA that takes a circular shape and can be found in archea and in bacteria. They replicate independently of the genome in the organism and can contain beneficial properties, such as antibiotic resistance or the capability to form toxins. Plasmids can be used for recombinant expression of enzymes, which means that genes for these enzymes can be inserted into plasmids from which the proteins can be produced in the host cells. [8,9]

Plasmids consists of different important elements, starting with the origin of replication, which is where the replication begins. Another element is the antibiotic resistance gene which makes sure that the host can survive an environment with some antibiotic in it. The third element worth mentioning is promoter regions, which drives the transcription of the insert and the last element of importance is the primer binding site, where primers, which initiate PCR amplification, can be bound. [8, 9]

The folding of a protein is an important process that determines its function. If the protein is not properly folded it will lack these functions and therefore not be very useful. Proteins can
fold spontaneously but some needs assistance and this can be achieved by molecular chaperones. Chaperonins are a class of chaperons that has been well studied and can be divided into two groups. Group I chaperonins can be found in bacteria and Group II chaperonins in archaea and in the cytosol of eukaryotes. [10, 11]

The chaperonin GroEL/GroES from *E. coli* is one of the most studied. GroEL consists of two rings that are stacked back to back and each of these rings has seven subunits. The GroES functions as a lid when the protein is folded in a cavity of the GroEL. First, ATP binds to GroEL, which will cause GroEL to bind to GroES and this will in turn cause a conformational change so that the cavity, where the protein is folded, is formed. When a protein is not present, GroES only binds to one of the rings of GroEL, but when a protein is present, it will bind to both rings and function as the lid. [10, 11]

*Restriction enzymes*

Restriction enzymes, also called restriction endonucleases, can recognize specific sequences in DNA and cleave it at those sites, which can give smaller fragments of the DNA. Restriction enzymes can be found in many bacterial types and the function there is to cut foreign DNA from, for example, viruses. To protect their own DNA, the species have methylations of the DNA which makes the sites unrecognizable for the restriction enzymes. There are three types of restriction enzymes, but the most useful one for recombinant technology is Type II that cleaves the DNA within the restriction sequence. These restriction enzymes can either cleave to give sticky ends or blunt ends (*Figure 1*). *HindIII* is an enzyme that gives sticky ends with four unpaired bp and *EcoRV* is an enzyme that gives blunt ends with no unpaired base pair. The recognition sequences are often 4-6 bp long and palindromic, which makes them easy to recognize when searched for. After a restriction, the fragments can be separated on an agarose gel and fragments of known size can then purified to be used further in for example ligation. [8]

During ligation, an enzyme called ligase can connect the restricted fragment to another fragment or plasmid and these often uses ATP as a cofactor. When fragments are to be ligated, both need to have been cut with either the same restriction enzyme to give matching sticky ends, or with an enzyme that gives blunt ends. This is because the ligase only connects the DNA and does not add any new nucleotides. [8]

*Figure 1.* Example of an enzyme that give blunt ends (*EcoRV*) and an enzyme that give sticky ends (*HindIII*).
Epoxide hydrolases

Epoxide hydrolases are enzymes that catalyze reactions of epoxides into vicinal diols. They participate in different pathways in different organisms, including detoxification, cell signaling and secondary metabolism. The epoxide hydrolases do not use cofactors and some show high catalytic efficiency which makes them desirable in biocatalysis. [12-14]

Most epoxide hydrolases can be placed in the family of α/β-hydrolase fold enzymes which forms vicinal diols as their products. The fold for these epoxide hydrolases consists of a core domain with a β-sheet surrounded by α-helices (α/β-domain) and of a lid domain with mostly α-helices. These domains together form the active site of the enzyme where three residues from the α/β-domain work as a catalytic triad and two tyrosines of the lid works by assisting the opening of the epoxide ring. This is done by hydrogen bonding to the oxygen in the epoxide ring of the substrate that is used. [15]

One epoxide hydrolase from potato, StEH1 (*Solanum tuberosum* epoxide hydrolase 1) accepts a broad set of substrates and is an α/β-hydrolase fold enzyme (*Figure 2*). This epoxide hydrolase has shown to yield the (R)-1-phenylethanediol from racemic mixtures of styrene oxide (*Figure 3*). [16]

![Figure 2. Structure of StEH1 with the lid domain in light blue and the α/β-domain in darker blue (pdb: 2cjp) [14] produced in PyMOL.](image)

![Figure 3. StEH1 wild type produces (R)-1-phenylethanediol as the main product from racemic mixtures of styrene oxide.](image)
Alcohol dehydrogenases

Alcohol dehydrogenases are enzymes that catalyze the reduction of ketones or aldehydes into alcohols or oxidation of alcohols into the corresponding carbonyl compound. They are NAD(P)-dependent enzymes, so they use the cofactors NAD(P)H and NAD(P)+ for the catalysis. The ADHs adds a hydride to carbonyls or removes them from alcohols. The regeneration of the cofactors can be done in several ways, for example with a coupled substrate or with a coupled enzyme (Figure 4). This approach could be cheaper and simpler, which is one of the factors that makes ADHs interesting in biocatalysis. Another approach for regeneration is to run the reactions in vivo and let the cells take care of the coenzyme regeneration. [17, 18]

![Figure 4](image)

**Figure 4.** (A) Regeneration of NAD(P)H or NAD(P)+ with a coupled substrate. (B) Regeneration of NAD(P)H or NAD(P)+ with a coupled enzyme.

Alcohol dehydrogenase A from *Rhodococcus ruber* DSM 44514 is NADH- and zinc-dependent and can be overexpressed in *E. coli*. This enzyme has shown high tolerance towards organic compounds, which has been shown by regenerating NADH with high concentrations of 2-propanol (up to 99% (v/v)) [19]. Furthermore, ADH-A has shown to be regio- and enantioselective and also demonstrates activity with many different alcohols and ketones. The wild type (wt) ADH-A shows a high efficiency for (S)-1-phenylethanol but a low efficiency for (R)-1-phenyl-1,2-ethanediol (Figure 5). The C1B1 mutant, which was used in this project, shows approximately 8-fold higher $k_{cat}$ for the (R)-1-phenyl-1,2-ethanediol compared to the wild type, but $K_M$ also increases, so the overall efficiency ($k_{cat}/K_M$) is approximately the same. Although the overall efficiency is low, an increase in $k_{cat}$ is more desirable than a low $K_M$. In cells, the efficiency is best evaluated through $k_{cat}/K_M$ because the substrate concentrations are low and $K_M \gg [S]$, which makes $v_0$ dependent on both $k_{cat}$ and $K_M$. When running the reactions for biocatalysis, a lot of substrate can be added and then the efficiency is best evaluated through $k_{cat}$ and $K_M << [S]$, which makes $v_0$ dependent on $k_{cat}$. Therefore, the C1B1 mutant gives a more satisfactory result for the (R)-1-phenyl-1,2-ethanediol than the wild type. [19-21]

The structure of ADH-A C1B1 takes a tetrameric form when crystallized (Figure 6) but is dimeric in solution. Each subunit contains two Zn$^{2+}$ ions, where one is stabilizing and the other one is a part of the active site. The F43H and Y54L exchanges makes the histidine side chain pointing in another direction as compared to the phenylalanine (Figure 7). This change in direction enlarges the binding-site for the substrate, which could be one of the reasons for the increase of $k_{cat}$ for (R)-1-phenyl-1,2-ethanediol. The Y54L exchange was made after the F43H and the result indicated that these two mutations together made the difference in efficiency, as the leucine also provide more space for the substrate binding. [21]
Figure 5. The ADH-A wt shows high efficiency for ((S)-1-phenylethanol) while ADH-A C1B1 shows a more satisfying efficiency for (R)-1-phenyl-1,2-ethanediol compared to the wt.

Figure 6. Structure of ADH-A C1B1 tetramer and structure of one subunit of ADH-A C1B1 (pdb: 6ffz) [21] produced in PyMOL.

Figure 7. Structure of the active site in ADH-A C1B1 (pdb: 6ffz) [21] and structure of the active site in ADH-A wt (pdb: 3jv7) [22] produced in PyMOL.
Project overview

The aim of the project was to insert the genes for StEH1 and ADH-A C1B1 into the same plasmid and to express them in the BL21 AI strain of *E. coli* to see if the expression would be sufficient for the intended use. ADH-A depends on the co-expression of the chaperonins GroEL/ES for proper folding and therefore the plasmid was transformed with the pRep4GroEL-ES which was already present in the used strain of bacteria. The starting plasmid, pGT7ADHC1B1 already contains the gene for ADH-A and the gene for StEH1 was to be inserted into this. Another plasmid, pGtacSTEH1 [23], was the source of the StEH1 cDNA. Both plasmids were identical with the exception of promoter and gene (Tac promoter for StEH1 and T7 promoter for ADH-A C1B1).

Both these genes had been inserted into the same plasmid (pETDuet-1) in previous work of Widersten group and the expression of ADH-A and the chaperonins was in that system insufficient. The hypothesis was that the two plasmids pETDuetADHC1B1StEH1 and pRep4GroEL-ES were not perfectly compatible and therefore a new plasmid was to be constructed to check if this could provide a more satisfying result.

The activities of these enzymes are interesting to combine, since the formed ketone could be a starting molecule in other reactions or be used for production of pharmaceuticals. The advantage here is that this could be done from a simple starting molecule and without organocatalysts which would make the synthesis more environmentally friendly and cheaper. The reaction that is studied and carried out by these enzymes of interest converts styrene oxide into 2-hydroxyacetophenone (*Figure 8*). StEH1 wt produces the (R)-1-phenylethanediol as the main product which is a preferred substrate for ADH-A C1B1.

![Figure 8. Reaction scheme for the reaction carried out by the enzymes expressed in this project.](image)

pGtacStEH1 was sent for DNA sequencing, so that the sequence could be observed for planning of the plasmid construction. Appropriate restriction sites were chosen for digestion of both plasmids containing the genes with help of *Genious® 11.1.2*.

Methods

Transformation for plasmid preparation

Transformations for plasmid preparations were made with the XL1-Blue strain of *E. coli*. 50 μL of cells were mixed with 1-1.25 μL autoclaved MQ-water (negative control) or 1-1.25 μL plasmid (20 – 150 ng) in an electroporation cuvette. The solutions were then electroporated at 12.5 kV/cm. After electroporation, 1 mL of 2TY medium (1.6 % (w/v) trypotone, 1.0 % (w/v) yeast extract, 0.5 % (w/v) NaCl) was added to both cuvettes and the cells were transferred into culture tubes and incubated at 37 °C and 200 rpm for 1 h. The cells were then diluted 100, 1000 and 10 000-fold before 100 μL was spread out on LB-plates (1.0 % (w/v) trypotone,
0.5 % (w/v) yeast extract, 1.0 % (w/v) NaCl, 1.5 % (w/v) agar) containing 100 μg/mL ampicillin. The plates were incubated overnight at 37 °C. Colonies were picked and inoculated into ~ 2 mL of 2TY media containing 100 μg/mL ampicillin in culture tubes and incubated at 37 °C and 200 rpm overnight or at 25 °C and 200 rpm over the weekend. The plasmids were then purified from the cells with a GeneJET plasmid miniprep kit and eluted with 50 μL autoclaved MQ-water. The concentrations of purified plasmids were measured with NanoDrop® ND-1000 spectrophotometer. Purified plasmid was sent for DNA sequencing for complete mapping of the plasmid sequences.

**Digestion**

All the restriction enzymes and the buffers used in the digestions were supplied by Thermo Fisher Scientific. A schematic figure over the digestions can be seen in Figure 9.

The digestions were carried out with approximately 500 ng of plasmid and carried out on both pGtacStEH1 and pGT7ADHC1B1. For the digestion of the pGtacStEH1 plasmid, the first digestion was followed by a fill-in reaction by DNA Polymerase I, Large (Klenow) Fragment to obtain blunt ends. Some digestions were carried out in two tubes in parallel to see if the digestion could work in different buffers. The digestions were performed at 37 °C for 1 h and were heat inactivated at 65-80 °C for 20 minutes. The digestion products were then loaded onto 0.8% agarose gels containing 1x GelRed™ together with a GeneRuler 1 kb ladder. The bands that were present at the right size were cut out and the DNA was extracted with GeneJET gel extraction kit and eluted with 50 μL autoclaved MQ-water. The concentrations of purified DNA were measured with NanoDrop® ND-1000 spectrophotometer.

The digestion of pGtacSTEH1 (ca 4000 bp) was performed by starting the cleaving with 10 U of HindIII, followed by addition of 0.5 U of DNA Polymerase I, Large (Klenow) Fragment and dNTPs (final concentration 2.1 mM) to obtain blunt ends. Then precipitation for change of buffer was carried out, before another digestion performed by 10 U of SacI, in Buffer SacI (10 mM Bis-Tris Propane-HCl pH 6.5, 10 mM MgCl2, 0.1 mg/mL BSA). The digestion performed by HindIII was carried out in two different buffers, 1X Buffer Klenow (50 mM Tris-HCl pH 8.0, 10 mM MgCl2, 500 mM NaCl) and Buffer R (10 mM Tris-HCl pH 8.5, 10 mM MgCl2, 100 mM KCl, 0.1 mg/mL BSA), to examine which would suit the best.

The digestion of pGT7ADHC1B1 (ca 5000 bp) was performed by 10 U of SmaI (which gives blunt ends) and 10 U of SacI. These were carried out in two different buffers, Tango Buffer (33 mM Tris-acetate pH 7.9, 10 mM magnesium acetate, 66 mM potassium acetate, 0.1 mg/mL BSA) and Buffer SacI (10 mM Bis-Tris Propane-HCl pH 6.5, 10 mM MgCl2, 0.1 mg/mL BSA) to examine if both enzymes could work in the other’s preferred buffer.

The precipitations were carried out by adding 0.1 volume of 3 M sodium acetate pH 5.2 and two volumes of ice-cold 99.5% ethanol. This was then incubated in -80 °C for 1 h followed by centrifugation at 17000 x g for 20 min. The supernatant was discarded and the pellet was washed twice with 70% ethanol before being dissolved in autoclaved MQ-water.

Ligations of digested gene fragments and digested plasmids were performed with 4-8 ng plasmid and 10-14 ng gene fragment. This was ligated in 1x T4 ligase buffer (40 mM Tris-HCl, 10 mM MgCl2, 10 mM DTT, 0.5 mM ATP), 0.4 mM ATP, 0.5 weiss u/µl T4 ligase in a total volume of 10 µL. “Negative” ligations were performed by adding autoclaved MQ-water instead of the gene fragments. Incubation of the ligation mixtures took place in the dark and at room temperature overnight. The enzymes were heat inactivated at 75 °C for 10 min.
The ligation mixtures were then transformed into the XL1-Blue strain of *E. coli* by mixing 50 μL cells with 1-2 μL DNA followed by electroporation at 12.5 kV/cm and then addition of 1 mL 2TY medium. This was transferred into culture tubes which were put on incubation at 37 °C and 200 rpm for one h. 100 μL of cells were then spread evenly on LB-plates containing 100 μg/mL ampicillin and these were put on incubation overnight at 37 °C. Colonies were inoculated into ca 2 mL of 2TY medium containing 100 μg/mL ampicillin in culture tubes and incubated at 37°C and 200 rpm overnight or at 25 °C and 200 rpm over the weekend. Plasmid DNA was extracted with a *GeneJET plasmid MiniPrep Kit* and sent for sequencing. The concentrations of purified DNA were measured with *NanoDrop® ND-1000 spectrophotometer*.

![Figure 9. Schematic illustration of construction of expression plasmid (the figure is not in scale). The size of the fragment with the StEH1 gene is around 1000 bp.](image)

**Protein expression and purification**

The constructed plasmid pGStEH1ADHC1B1 was transformed into the BL21-AI strain of *E. coli* in parallel with the pETDuetADHC1B1StEH1 for a direct comparison to see if the new construct would give a more efficient expression of both enzymes. The transformation was carried out as for the plasmid preparation, with some exceptions. The LB-plates and 2TY contained 100 μg/mL ampicillin and 30 μg/mL kanamycin and the BL21-AI strain of *E. coli* was used. Also, the colonies that were inoculated into 2 mL 2TY, were incubated for 4 h instead of overnight or over the weekend.

After 4 h, 350 μL of cells were transferred into 35 mL 2TY containing 100 μg/mL ampicillin and 30 μg/mL kanamycin followed by incubation at 30°C and 220 rpm overnight. 5 mL of these cells were then transferred into 500 mL of 2TY containing 50 μg/mL ampicillin and 15 μg/mL kanamycin and this was put on incubation for 3 hours, until the OD<sub>600</sub> reached >0.4. Protein expression was induced by adding IPTG to a final concentration of 1 mM and L-arabinose to a final concentration of 0.2% (w/v). The induction was followed by incubation at 30 °C and 220 rpm overnight.
The cells were harvested by centrifugation at 4 °C, 5000 x g, 15 min and the pellets were resuspended in 20 mL lysis buffer (20 mM imidazole, 500 mM NaCl, 20 mM sodium phosphate, pH 7.5, 1 complete tablet EDTA free, 0.02 mg/mL DNase 1) before being transferred and homogenized in a homogenizer. The cells were lysed with a cell disruptor Z plus series from Constant systems Ltd. and then centrifuged at 4 °C, 15000 rpm, 1 h. The supernatant (lysate) was transferred to a conical tube and an equilibrated IMAC gel was added before putting the tube on incubation on a vipping table at 4 °C for 1 h. The pellet was resuspended in 25 mL binding buffer (20 mM imidazole, 500 mM NaCl, 20 mM sodium phosphate) and saved in a conical tube.

The IMAC gel lysate slurry was poured into an empty PD-10 column and the flow-through was collected. The gel was then washed three times with 20 mL washing buffer (100 mM imidazole, 500 mM NaCl, 20 mM sodium phosphate, pH 7.5) and all three fractions were collected. The proteins were eluted from the gel by adding 5 mL elution buffer (300 mM imidazole, 500 mM NaCl, 20 mM sodium phosphate, pH 7.5) twice to the gel followed by collection of the fractions. To obtain the final product of the proteins, the eluted fractions were pooled and concentrated down to 2.5 mL by centrifugation at 4 °C, 4000 x g, 7 min in a spin-column (Vivaspin 20, 100000 M_w cut-off, Sartorius Stedim Biotech). Lastly, the concentrated samples were desalted in a Sepharose G-25 PD-10 column and eluted with 3.5 mL desalting buffer (100 mM sodium phosphate, pH 7.5, 10 μM ZnSO4).

From all the different fractions obtained in the purification, 2 μL of sample was mixed with 8 μL MQ-water and 10 μL sample buffer (0.5 M Tris-HCl pH 6.8, 25 % (v/v) glycerol, 2 % (w/v) SDS, 0.01 % (w/v) bromophenol blue, 5 % (v/v) β-mercaptoethanol). These were heated at 95 °C for five minutes and then loaded onto an SDS-PAGE gel. Also, an Mw marker (3 μL from GE healthcare, 7 μL MQ-water, 10 μL sample buffer) (M) and references of purified StEH1 (1 μL 35.9 μM StEH1, 9 μL MQ-water, 10 μL sample buffer) (RS) and ADH-A (0.5 μL 35.9 μM StEH1, 9 μL MQ-water, 10 μL sample buffer) (RA) were loaded. The gel was run at 200 V for 50 minutes. The fractions that were collected were: Lysate (Ly), Pellet (P), Flow-Through (FT), Wash 1 (W1), Wash 2 (W2), Wash 3 (W3), Elution 1 (E1), Elution 2 (E2) and Final (F). The gels were then stained with staining solution (1 tablet Coomassie Brilliant Blue R-250 supplied from VWR, 60 % (v/v) methanol diluted 1:1 with 20 % (v/v) acetic acid) and microwaved until boiling followed by incubation at room temperature for 10 minutes. After staining, the gels were destained with destaining solution (10 % (v/v) acetic acid, 20 % (v/v) methanol) and microwaved until boiling followed by incubation at room temperature for 10 minutes. The destaining procedure was repeated three times.

Results

Digestion

The first digestion of pGtacStEH1 suggested that the gene was successfully cut out by the enzymes, since a fragment at slightly more than 1000 base pairs could be seen and this should represent the StEH1 gene (Figure 10). This digestion was carried out in two separate tubes with different buffers and both seemed to have worked for HindIII. After the extraction and purification of the DNA, a low concentration of 1.7 ng/μL could be detected in the tube that contained Buffer Klenow for the HindIII digestion. This digestion was repeated several times and yielded low concentrations in all cases.
The digestion of pGT7ADHC1B1 was carried out in different buffers and at least one restriction enzyme has cleaved the plasmid, since the bands show linear fragments and not circular (Figure 11). The results from this gel do not reveal if both enzymes have worked, but instead this would be determined after the ligation analysis.

*Figure 10.* Result of one gel after digestion of pGtacStEH1. Fragments could be seen at around 3000 bp and 1000 bp suggesting that the gene was successfully cleaved out. The numbers next to the fragment length marker represent the number of bp that fragment contains.

*Figure 11.* Result of the gel after digestion of pGT7ADHC1B1. The plasmid has been cleaved which can be seen on the gel since the bands look like linear DNA. The numbers next to the fragment length marker represent the number of bp that fragment contains.

The ligation products were sent for sequencing and the result of these suggested that the ligation had worked in one case, where the cleaved plasmid from the Buffer Tango experiment was ligated with one of the StEH1 fragments which had a concentration of 2.4 ng/μL. Also, a digestion of this ligation mixture was carried out and showed that some larger fragment was cut out, also suggesting that the gene for StEH1 had been inserted (Figure 12). *Figure 12* also shows one ligation that was unsuccessful and contained the cleaved plasmid from the Buffer SacI experiment.
Figure 12. Result of the gel after digestion of some of the ligation products. The product called BT2.4 shows that a larger fragment (green arrow) has been cleaved out when HindIII and SacI was used, which suggests that the ligation was successful and that this represents the size of both genes. The product called 4BSac1.7 suggest that the plasmid has self-ligated since a fragment that corresponds to the size of the ADH-A gene is present (red arrow). The numbers next to the fragment length marker represent the number of bp that fragment contains.

Protein expression and purification

The analysis of the purification of the proteins was made with SDS-PAGE and two gels were run, one for the pETDuetADHC1B1StEH1 and one for the pGSTEH1ADHC1B1. The gels do not differ from each other significantly (Figures 13 and 14). The gels show that a lot of the proteins are found in the pellet for both plasmids and in both cases, a larger proportion of StEH1 is found in the soluble fractions compared to ADH-A. The chaperonins were successfully expressed in both cases which shows in the lysate and the flow-through. References for both enzymes were used to be able to see what band represents what enzyme, since the bands are rather close to each other, with the StEH1 being a bit smaller than ADH-A.

Figure 13. SDS-PAGE of the fractions collected during the purification from the pETDuetADHC1B1StEH1 plasmid. M=Mw marker, Ly=lysate, P=pellet, FT=flow through, W1=wash 1, W2=wash 2, W3=wash 3, E1=elution 1, E2=elution 2, F=final, RS=reference StEH1, RA=reference ADH-A.
Discussion

Digestion

The digestions were carried out in different buffers to be able to see what buffers could work for more than one restriction enzyme. The digestion to obtain the StEH1 fragment seems to have worked in both buffers that were used, so in further digestions of this fragment, they were carried out in Klenow buffer to make sure that the fragment obtained blunt ends. For the opening of the plasmid where the gene was inserted, this seems to have worked in buffer Tango, since this gave a correct ligation while for buffer SacI, the digestion seems to not have worked properly, since the StEH1 fragment was not inserted during the ligation.

For the ligation that worked, the sequence was examined, but neither of the gene sequences could be read in a whole and therefore, it could only be assumed that no mutations occurred during the ligation and that the purification could still work. The purification depended on that the His-tag after the StEH1 gene was intact in the constructed plasmid and this could be confirmed after the purification and SDS-PAGE analysis.

Protein expression and purification

Since both proteins are shown on the SDS-PAGE analysis, the His-tag can be assumed to be intact after the digestion and ligation. The same problem seems to be present for both plasmids, since both SDS-PAGE gels show that a big amount of ADH-A and some amount of StEH1 are still present in the pellet. This suggests that the proteins, especially ADH-A has some issues during the folding and instead of forming the correct structure, it formed inclusion bodies. The chaperonins might be helpful, since some ADH-A can be seen in the elution fractions and the final fraction, but to verify this, the expression and purification would have to be carried out in the absence of GroEL/ES.

The assumption previously made about that the pETDuetADHC1B1StEH1 and pRep4GroEL-ES are not compatible might not be correct. If this would have been the case, the change of plasmid should have given a more satisfying result, where both proteins would be present in a larger amount outside of the pellet after cell lysis. The problem also does not seem to be in the expression, but in the folding itself, since a broad band still could be seen for ADH-A in the pellet and that this is the case for both plasmids.
A new approach to the problem is therefore needed and one thing to try could be to induce the proteins at different timepoints, since they are induced by different compounds. The induction needs to start with IPTG to begin expression of the chaperonins and StEH1 if the pGSTEH1ADHC1B1 is used, since ADH-A will not fold properly if the induction is started with L-arabinose, since the chaperonins are not expressed at this point. Another approach that could be attempted, is to change the strain of bacteria that is used. It might be beneficial for the folding if the bacteria grow a bit slower than the one currently used.

The insufficient folding of the ADH-A might be dependent on that it is co-expressed with the StEH1. StEH1 doesn’t need the chaperonins to fold properly, but it might still take advantage of them and therefore inhibit the folding of ADH-A instead. The ADH-A has been successfully expressed and folded in the pGT7ADHC1B1 (the plasmid which was used for insertion of the StEH1 gene) in other work within the Widersten group. This suggests that the plasmid is not the issue, but rather that the co-expression is.

**Conclusion**

It could not be proven that the issue with co-expression of StEH1, ADH-A and GroEL/ES was because of the incompatibility of pETDuetADHC1B1StEH1 and pRep4GroEL-ES. It rather seems like the issue lies in the folding of the proteins. This project could take some other directions instead, by trying to change something other than the plasmids that are used.

**Acknowledgements**

I would like to thank my supervisor prof. Mikael Widersten for giving me this project and for helping me, answering my questions and discussing my project. Isac Söderlund has also been of big help during the project, especially with the lab work and showing me the different techniques. Lastly, I would like to thank the rest of the Widersten group for welcoming me and for making my project time fun and meaningful.
References


