NTNU – Trondheim
Norwegian University of
Science and Technology

Biochemical Characterizations of Enzymes from Oil Reservoir Metagenomes

Ronja Kristine Hesthammer

Biotechnology (5 year)
Submission date: June 2014
Supervisor: Martin Frank Hohmann-Marriott, IBT
Co-supervisor: Finn L. Aachmann, IBT
Anna Lewin, SINTEF
Alexander Wentzel, SINTEF

Norwegian University of Science and Technology
Department of Biotechnology

ABSTRACT

Microorganisms have evolved to exist in diverse environments on earth, even under the harsh conditions of oil reservoirs several kilometres below the seafloor. The environmental conditions there are extreme in several aspects, such as high temperature, pressure, salinity, and presence heavy metals or other potentially toxic compounds. In order to live under these conditions, the microbes' enzymes have evolved to withstand the extremes. Enzymes with such traits might be utilized for several biochemical industries.

In a previous project, microbial DNA from two oil reservoirs has been isolated and high-throughput pyrosequencing used to create an extensive metagenomic DNA sequence database. Using bioinformatics tools, three putative gene sequences for thermostable peptidases were identified in this oil reservoir metagenomic sequence database. They showed a high degree of sequence homology to a known peptidase of the enzyme sub-classes, subtilisin-like protease, carboxypeptidase or aminopeptidases, and were named SLP01, CPT01 and ATP01.

Protein production of the recombinant proteins was performed in *Escherichia coli* strains BL21 (DE3) CodonPlus-RIPL and ER2566, using the plasmid vector pET21, at temperatures 16 and 37 °C. Desired protein production was verified using SDS-PAGE analysis. Production of target protein was observed for the APT01 and CPT01 candidate, although in an insoluble form. Some soluble APT01 was observed when the culture had been incubated at 16 °C.

Heat denaturation series was performed on the sample containing soluble APT01, at temperatures of 65 and 80 °C for a duration of 5, 15, 30, 60 and 120 minutes. The results showed APT01 to be thermostable at 65 °C, but none at 80 °C. An activity assay, using skim milk agar (MA) plates, was performed on all samples from the protein production. No activity was observed for any of the three candidates.

ACKNOWLEDGEMENTS

This thesis was performed as the final assignment for the Master of Science study in Biotechnology at the Norwegian University of Science and Technology in Trondheim (NTNU). The practical work for this thesis was carried out at the molecular biology laboratories, using equipment and instruments available in these laboratories.

I especially want to thank all of my supervisors: Professor Finn Drabløs, for helping me through the maze of bioinformatics tools, and how to interpret the findings. Dr. Finn L. Aachmann for initially getting me interested in this project, signing all the necessary forms, and for counselling me in the writing process. Dr. Anna Lewin and Dr. Alexander Wentzel for showing how to use the equipment in the lab, sharing their expertise and answering all the questions I have had at any time of the day, as well as reading several drafts of this thesis.

At last, I want to thank my family and friends for all of their support and all the good memories I have acquired over the five years of this master's degree. My time as a student in Trondheim have truly been terrific, and is something I will be happy and proud of for the rest of my life.

TABLE OF CONTENTS

1. INTRODUCTION

1.1 Microbial diversity and bioprospecting

Microbes have been found to live in most of earths extreme environments, from hydrothermal vents in the deep-sea to the dry valleys of Antarctica. These microorganisms are called extremophiles, or poly-extremophiles if they thrive in (or tolerate[1]) an environment that is extreme for more than one condition (Rothschild, 2001). These extremes can be high or low temperatures and pH, high salinity or pressure, or the presence of toxic or degenerative compounds (e.g. high concentration of heavy metal ions and organic solvents). In order to live and multiply under such conditions these microbes have adapted their metabolism to tolerate the extremes.

There is an increasing demand within modern biotechnology to discover among others new antibiotics and novel biocatalysts for use in biocatalytic industrial processes. By finding enzymes that can catalyse a desired reaction at high temperature or pressure, new and more efficient processes for productions of products such as pharmaceuticals, agrochemicals and food additives can be obtained. Cultivation and isolation of strains containing the desired characteristics by assays have traditionally been used to discover novel enzymes. However, of the microbes present in several environments, as much as 99.8% cannot be grown readily in pure culture and therefore cannot be exploited for biotechnology in a traditional manner (Streit, 2004). The “metagenome technology” provides an alternative approach, by allowing cloning of environmental DNA, screening for desired function and production of protein of interest without cultivation of the original organism. So far, metagenomic analysis have been used to find novel antibiotics and antibiotic resistance genes, lipases, amylases, deoxyribonuclease and lyases, to mention some (Lewin, 2013, Rondon, 2000).

Metagenomic research is important, not only to identify novel enzymes and bioactive compounds, but also to obtain a greater understanding of the vast uncultivable microbial diversity that exists. To analyse the microbial community as a whole and try to understand the underlying metabolic mechanisms and regulatory network making these communities function. With the current technology it is now possible to gain a further and more complete understanding of the evolution of these microbes, and how

[1] For simplicity the term extremophile is not only used for the microbes that “phile” (ie. love) the environment, but also for organisms that tolerate and are able to live under these condition.

they have adapted to live under such poly-extreme conditions (Handelsman, 2004, Lewin, 2013).

1.1.1 Metagenomic libraries

Metagenomics research is the analysis of genomic DNA directly isolated from an environment. DNA libraries are made using isolated DNA from environmental samples, either by use of cloning vectors or by direct sequencing (figure 1.1). To make a clone library, isolated metagenomic DNA is digested and inserted into plasmids, for small inserts (<10 kb), or cosmids, fosmids or bacterial artificial chromosomes (BACs) for larger inserts (>20 kb). Host cells are then transformed with the vector and 16S DNA from each clone is amplified, purified and sequenced using, for instance, Sanger sequencing (Rondon, 2000, Sabree, 2009). A much faster way of making a metagenomic DNA sequence library is by direct sequencing, using next generation sequencing technologies. Commercial pyrosequencing for metagenomic research was developed by the company 454 Life Sciences, so '454 sequencing' is a much-used method for making direct sequence libraries. Sequence reads from the pyrosequencing is finally assembled into contigs, which comprises the DNA library (Kotlar, 2011, Lewin, 2013, Petrosino, 2009).

Metagenomic libraries can be made based on samples from environments never before been analysed, or environments where only a small fraction of the total microbial diversity have been analysed in the past. There are two main strategies to identify genes of interest in metagenomic libraries. The first being function-based analysis, the other using sequence-based searches, each having their strengths and limitations. Using these methods, it allows us to tap into the nearly unlimited resource of novel genes for enzymes and other compounds that potentially can be utilized in the field of medicine (e.g. antibiotics) or as biocatalysts to streamline industrial bioprocesses.

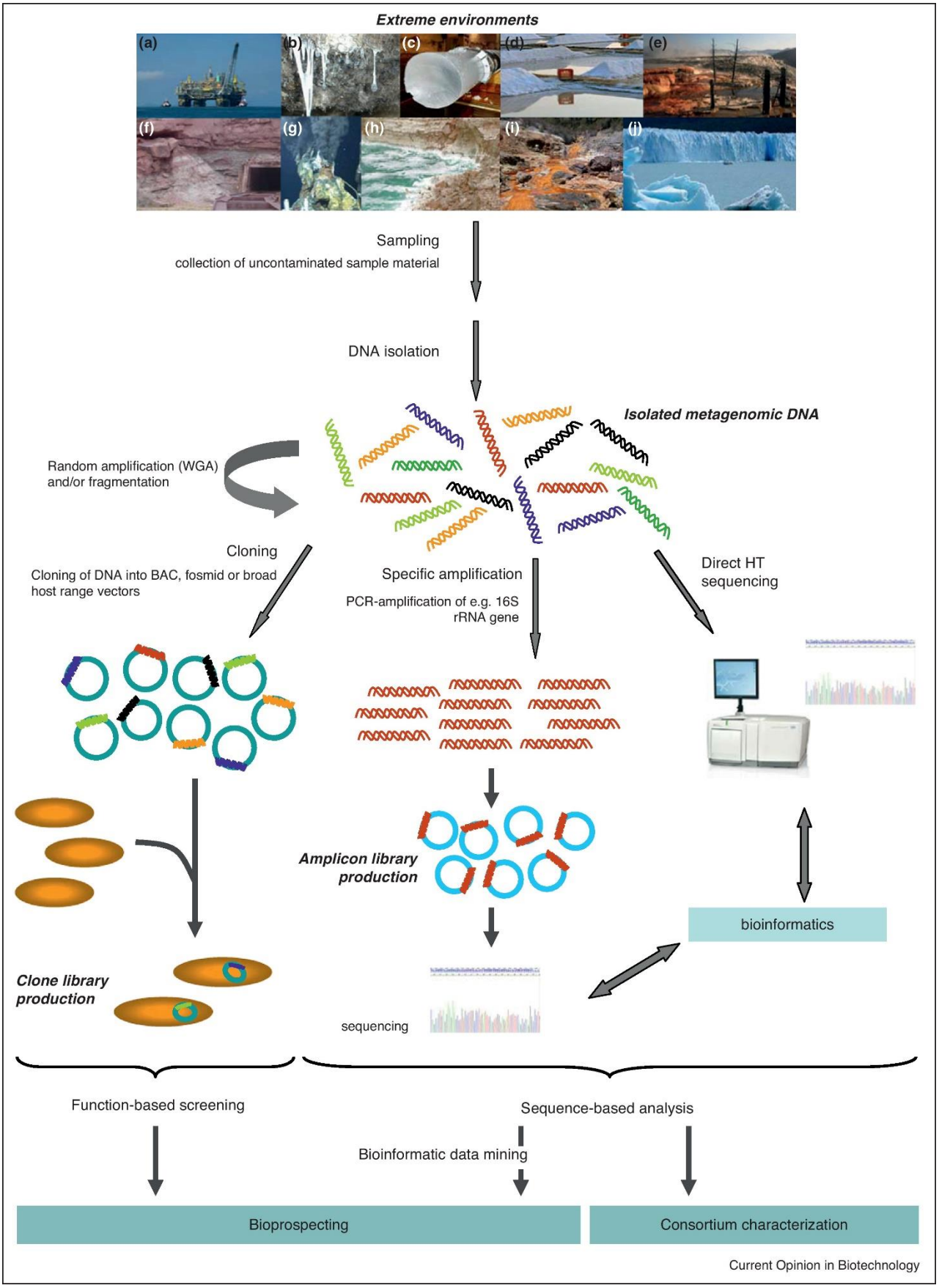

Figure 1.1: *Schematic overview of metagenomic study methodology and construction of DNA libraries from environmental samples (a-j)(Lewin, 2012).*

1.1.2 Function-based analysis

For function-based screening, cloned metagenomic libraries are constructed and expressed in a suitable host organism. The screening process to identify clones with the desired trait is often done by relatively 'simple' activity assays. Identifying clones that express a certain characteristic is independent to sequence homology, which makes it the only approach that can potentially identify entirely new gene classes (Handelsman, 2004). The downside of this strategy is the heterologous expression and low detection frequencies of functional proteins. For example, in a study done by Rondon et al. in 1999 only 29 hemolytically active clones where identified out of a total of over 28 000 clones (Rondon, 2000). In the search for florfenicol[2] resistance genes from Alaskan soil, only a single clone was identified from 13 201 Mb of DNA screened (Lang, 2010). Reasons for this can be found all along the protein synthesis pathway, from inefficient transcription and translation of the metagenomic gene, insufficient secretion of the foreign protein, incorrect folding, or lack of cofactors necessary for the proteins stability and activity (Streit, 2004). The protein product can also be toxic to the host cell at different levels. When expressed under control of a strong promoter the over-expression of the recombinant gene can lead to depletion of the cells energy stores causing the growth rate to stall or, in worst case, kill the organism (Tabor, 2001).

1.1.3 Sequence-based Analysis

Gene mining, direct PCR and hybridization are tools that are used for sequence-based analysis. For gene mining, algorithms and bioinformatics tools are used to search for sequence alignments between the metagenomic DNA library and annotated sequences of interest in a database (e.g. UniProt). Sequence alignment analyses can be performed on any computer with access to the metagenomic DNA library of interest and a protein database (many are available online). This makes this a cheap and effective approach in the search for DNA sequences of interest. PCR and hybridization techniques include the use of arrays and metagenome sequence tags (MSTs) to identify conserved regions of genes. All sequence-based screening is based on conserved and know sequences and cannot be used to detect new gene classes (Streit, 2004). Another limitation of the homology-based identification is that the

[2] Florfenicol is a synthetic fluorinated derivative of chloramphenicol.

functionality of detected putative genes must be demonstrated, which faces the same challenges of recombinant gene expression as described for the function-based analysis. Despite this, it is anticipated that combined with the improvement of bioinformatics tools and protein classification, sequence-based screening of metagenomic databases will have a greater influence on identifying novel biocatalysts compered to function-based analysis (Li, 2009).

The Basic Local Alignment Search Tool (BLAST) is one of the bioinformatic tools used for sequence-based screening of metagenomic DNA libraries (Altschul, 1997). It is a widely used tool to find regions of local similarities between protein or DNA sequences, using annotated databases. BLAST can be used to help identify gene families and to deduce functional and evolutionary relationship between sequences. Different BLAST programs allow you to search a nucleotide, translated nucleotide or protein database using a nucleotide, translated nucleotide or protein sequence as query. The query is compared to sequences in the chosen database and statistical significance of the matches is calculated and displayed. A second much used bioinformatic tool for sequence-based screening is the Open reading Frame (ORF) Finder. This program identifies the possible open reading frames in the DNA sequence query, and lists these along with their protein translation. Information provided by these two tools can be used to identify putative genes interest, for instance genes encoding novel biocatalysts. To increase the chance of identifying an active gene which encodes a functional protein, the DNA sequence need to contain elements such as a start and stop codon and Shine-Dalgarno sequence, used by the cells transcriptional machinery to initiate and terminate gene expression. The encoded protein sequence should contain a conserved active site, which can be checked using BLAST to compare the protein sequence given by ORF Finder to a protein database.

1.1.4 Oil Reservoir Microbiology and Metagenomics

Within the first years of the metagenomic era the main focus has been on easily accessible environmental samples, such as soil (Rondon, 2000) (Lang, 2010), while more extreme environments like hot springs, glacial ice or marine subsurface have remained an untapped source of potentially valuable biocatalysts. Biddle and co-workers (2008) wrote that *"The subseafloor marine biosphere may be one of the largest reservoirs of microbial biomass on Earth […]."* (Biddle, 2008). Deep subsurface oil reservoirs are an example of poly-extreme environments of the subseafloor. These are environments with high temperature, pressure and salinity, and they have often been physically separated to life on the surface for a very long time. The formation of oil utilizes organic material deposited millions of years ago. The microorganisms associated with this material might therefore have been separated from the surface environment for a substantial amount of time, indicating a difference to their relatives on the surface at the DNA sequence level (Kotlar, 2011).

This project is part of an ongoing research activity started in 2008 as a collaboration between SINTEF, NTNU and Statoil. Two samples, one from each of two separate oil reservoir wells on the Norwegian Continental Shelf (NCS) (figure 1.2), were collected using specialized technology allowing contamination-free samples, as well as avoiding potential loss of microbes DNA due to rapid pressure changes. Total DNA of the samples was isolated and direct 454 pyrosequencing was used to construct metagenomic DNA libraries. Bioinformatics analysis on the DNA sequences showed that the reservoirs harbour a rich community of microorganism, but are dominated by a smaller number of taxa (Kotlar, 2011). Bacteria dominate "Well I" with a smaller fraction of Archaea, while it is the opposite for "Well II". However, the two wells appear to contain essentially the same organisms. Considering that the two reservoirs appear to be physically separate, and that they have been separated for a considerable amount of time, Lewin and coworkers (2013) proposed that the average generation times of the organisms are extremely long resulting in a small degree of evolvement since the separation took place.

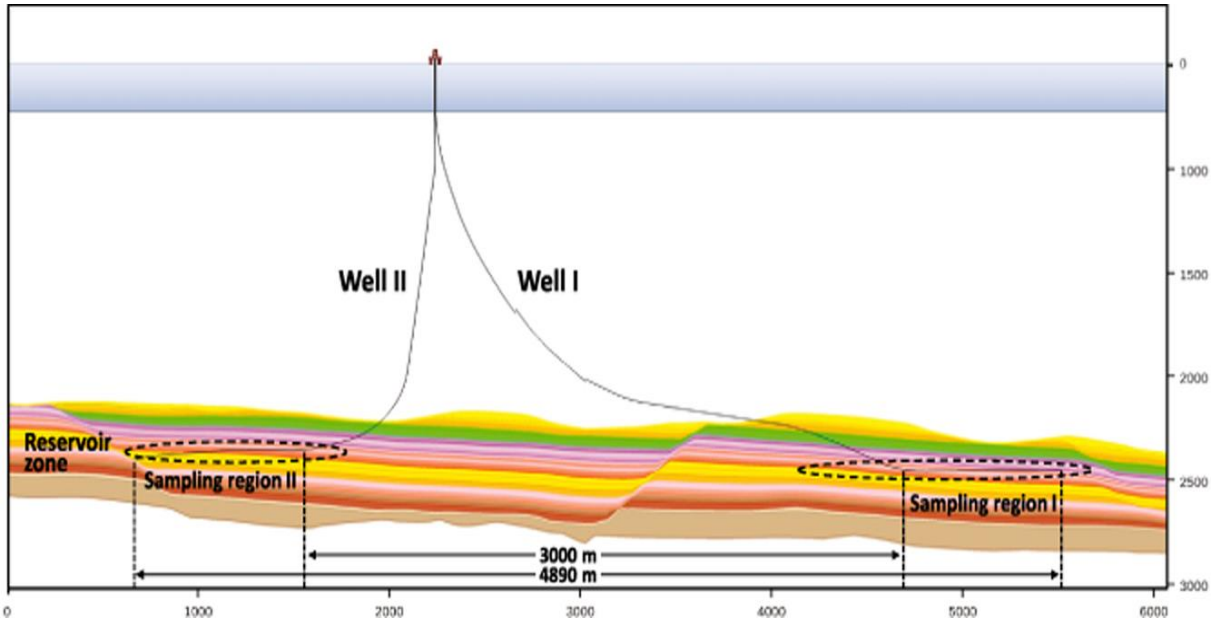

Figure 1.2: *"*Schematic representation of the geological profile of the area for oil reservoir wells sampled. The different lower sediment formations are shown by different colouring, and the areas for the sampled regions for Well I and Well II are indicated." (Lewin, 2013) *Distance in meters are given in the horizontal and vertical axis.*

1.2 Protein classes of interest for industrial processes

The Structure-Function Linkage Database (SFLD) classifies predicted enzymes based on specific sequence-structure features to specific chemical capabilities, hence hypothesizing function. In this database an enzyme family is defined as a group of *"[…] evolutionary related proteins that catalyse the same overall reaction in the same way."* (Hicks, 2014). A protein is clustered into a family of which it has high sequence or structural similarity to the other family members, i.e. they are homologous. Some families are divided into subfamilies when there is evidence for an ancient divergence corresponding to more than 150 point mutations per 100 amino acid residues. This number of point mutations represents a time line of 2500 million years, when calculated with the typical evolutionary rate (0.6 substitutions per amino acid per 1000 million years (Rawlings, 2007). Annotation of enzymes normally includes the assignment of a class within the Enzyme Commission (EC) Scheme presented by the NC-IUBMB (Moss, 1992). An enzyme is assigned four numbers, where the first represents one of the six main classes of reactions catalysed (oxidoreductases, transferases, hydrolyses, lysases, isomerases or lipases, respectively). The second and third number represents the sub and sub-sub class of the overall reaction and the fourth number indicates substrate specificity.

Since ancient time have enzymes found in nature been used on production of foods such as cheese, sourdough, wine and beer. Over the last five decades, there has been a massive development in the enzyme industry. Thanks to manufacturing of purified enzymes and recombinant gene technology, along with protein engineering, it has been made possible to provide enzymes tailor-made for the process of interest. Today, enzymes are used in countless industrial processes, from active components in detergents and hygiene products to textile and paper production. Of all the enzymes used industrially, peptidases remain the dominant type due to its extensive use in detergent and dairy industries. While the second largest group is represented by various carbohydrases, primarily amylases and cellulases (Kirk, 2002).

1.2.1 Peptidases

Peptidases are ubiquitous enzymes, essential for the survival of all living organisms. The digestion of food proteins, antigen presentation of the immune system and regulation of blood clots, as well as post-translational maturation of enzymes, peptide hormones and neurotransmitters, are some biological processes where peptidases play an important role. Peptidases have been used for several industrial processes, from clotting cheese and processing of other foods to being one of the active agents in detergents. Other commonly used names for peptidases are proteases, proteinases and proteolytic enzymes, but peptidases is the term recommended by the Nomenclature Committee of the International Union of Biochemistry and Molecular Biology (NC-IUBMB) and the Human Gene Nomenclature Committee (Kirk, 2002, Rawlings, 2007).

A peptidase is an enzyme that hydrolyses peptide bonds, and therefore belongs to the third EC class. The reaction on peptide bonds is assigned subclass number four, hence all peptidases have the EC number 3.4.X.Y., where the X represents the sub-sub class, and the Y indicates substrate specificity. There are six main catalytic types of peptidases recognized. Serine, threonine, cysteine, aspartic, glutamic and metallo-peptidases, indicating what chemical group is responsible for hydrolysing the peptide bond. The catalytic nucleophile of the serine, threonine and cysteine type is the reactive group of the amino acid side chain, while an activated water molecule is the nucleophile of the metallo- and aspartic type. For metallo- peptidases the metal ion, which most commonly is zinc but it can also be cobalt, copper or manganese, is tetrahedrally coordinated by three amino acid residues and the activated water molecule (Rawlings, 2007). The glutamic peptidases have a catalytic dyad of Glu/Gln. This peptidase type was not recognised until 2005, and there is still much to learn about its mechanism (Rawlings, 2007) (Yabuki, 2004).

Structural analysis of peptidases have shown that the active site is commonly located in a groove on the surface of the enzyme. The catalytic mechanisms of peptidases varies between and within classes (depending on which classification method used), but can be described as an acid-base reaction, where the nucleophile acts as the proton donor and an amino acid residue (called the general base) is the proton acceptor. The active site residues are well conserved for peptidases within a family (Rawlings, 2007).

Classification of peptidases based on chemical mechanism of catalysis is not optimal. For example, even though all serine peptidases contain a serine residue that acts as the nucleophile, they have many different molecular structures and catalytic mechanism. In other words, they are not all homologous to each other (Rawlings, 2007). There exists two other ways of classifying peptidases as well. One is by the kind of reaction they catalyse, i.e. endo- or exopeptidases, while the last and newest method of classification uses molecular structure and homology to group peptidases.

Endopeptidases

The general characteristic of endopeptidases is that they cleave internal peptide bonds in proteins and larger polypeptide chains. Acting distant from the existing N- and C-termini, endopeptidases generate new termini that can function as substrate for exopeptidases, for instance in food digestion. There are endopeptidases that can only hydrolyse peptides that are smaller than proteins, called ‘oligopeptidases’.

The endopeptidase group includes several enzyme classes (as classified by the EC Scheme). One of these are the serine peptidases (EC: 3.4.21.Y). For serine peptidases the proton donor and acceptor along with a third member makes up the catalytic triad of the active site, represented by a serine, histidine and aspartate residue, respectively. The serine residue anchors the acyl-enzyme intermediate during the reaction, while aspartate orientates the imidazolium ring of the proton donor histidine and helps charge one of the nitrogen atoms in the ring. One family of serine peptidases is the subtilisins. Most of the subtilisins are endopeptidases, meaning that they hydrolyse a peptide bond away from the N- or C-terminus. Most subtilisins are extremely thermostable, making them suitable for many industrial applications (Rawlings, 2007).

Exopeptidases

Exopeptidases hydrolyse peptide bonds within three residues from the terminus of a polypeptide, and therefore requires a free amino or carboxyl group, or both.

There are six types of exopeptidases. Amino- and carboxypeptidases release a single amino acid residue from the N’- and C’-terminus, respectively, needing only the one end group free. Dipeptidyl-peptidases and peptidyl-dipeptidases both release a

dipeptide from its substrate N'- or C'- terminus, respectively. Dipeptidases hydrolyse dipeptides and need both end groups free. However, the term 'dipeptidase' is often used for any peptidases that release dipeptides from their substrate. The last type, tripeptidyl-peptidases cleaves off a tripeptide from the N'-terminus (Rawlings, 2007). Since their hydrolytic activity is on the N'-terminus, the term 'aminopeptidase' is also used for the di- and tripeptidyl peptidases. Aminopeptidases are involved in many proteolytic pathways such as protein maturation, hormone level regulation and cell-cycle control, but they seem to act in coordination with other peptidases to fulfil these roles. As they are of such importance to cellular processes, they are of interest to the pharmaceutical industry (Brown, 2003, Sanz, 2007).

1.3 Recombinant expression systems

The use of recombinant proteins have increased a great deal over the past decade, so naturally, the resources spent on understanding and developing better recombinant systems have increased as well. Bacterial systems are mostly used due to high productivity using inexpensive substrate media, as well as their well-characterized genetics, which has resulted in many available cloning vectors and mutant host strains (Terpe, 2006). All bacteria can be used for recombinant protein production, but *Escherichia coli* is the most common. This Gram-negative bacterium is cheap to grow, it grows rapidly, is well known and has been established in most laboratories since the birth of molecular cloning in 1973 (Cohen, 1973). However, a downside with using *E. coli* as a host is that recombinant proteins are often produced in insoluble and non-functional form (Sørensen, 2005).

To obtain a high level of the target protein it is often needed to clone the gene downstream of a well understood and regulated promoter in the vector used. Choosing the best host and promoter system for a specific protein is not always straight forward, and often depends on the heterologous protein itself. Many bacterial systems are not able to perform the correct posttranslational modifications that can be essential for the activity of the target protein (Terpe, 2006). There is no guarantee that the product of the recombinant gene will successfully be produced in its active form, but many modified strains of host cells and vectors utilising different promoter systems have been developed to overcome some of the different challenges.

1.3.1 T7 RNA polymerase system

The naturally occurring *lac* operon of *E. coli* has for many years been used to regulate the bacterium, and most of the promoters developed to induce heterologous production have elements derived from this. The T7 RNA polymerase system is one of the most widely used promoter systems for *E. coli*, and its use in the pET vector has become very popular (figure 1.3). The target gene is cloned into pET plasmids downstream of a strong bacteriophage T7 promoter, so a source of T7 RNA polymerase is needed to induce expression (Tabor, 2001). The host cells, transformed with the pET plasmids, contain a prophage (e.g. λDE3) encoding the T7 RNA polymerase under the control of the isopropyl-β-D-1-thiogalactopyranoside

(IPTG) inducible *lac* derived promoter *lac*UV5 (Baneyx, 1999). This system has several advantages, such as several times faster RNA elongation compared to the bacterium's own RNA polymerase, the termination of transcription is less frequent, and it allows the target gene to be cloned with low transcriptional activity. This is important as even a low expression of the heterologous protein can interfere with the metabolism of the host cell, and in that way be toxic to the cell (Tabor, 2001).

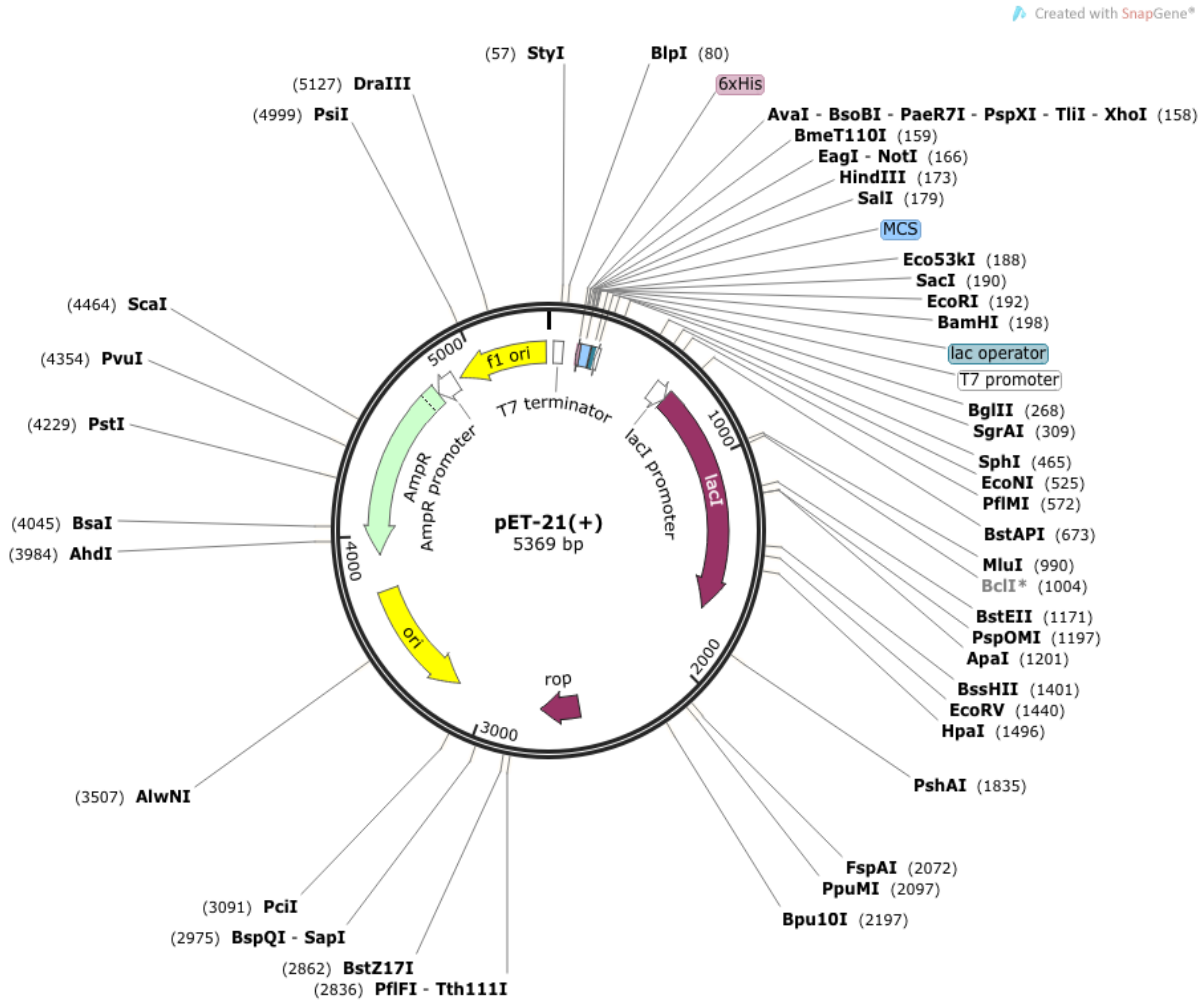

Figure 1.3: *Plasmid map of the pET21(+) vector. This plasmid contains a ampicillin resistance gene (green), two origins of replication (yellow), the* lacI *gene (red), MCS (light blue) and lac operator (turquoise) (www.snapgene.com/resources).*

1.3.2 Host strains

Many mutant host strains have been developed for *E. coli* to address some of the limitations met when producing heterologous proteins. The most commonly used are BL21 and K-12, and their derivatives (Terpe, 2006). BL derivatives are deficient in the *E. coli* endogenous proteases *lon* and *ompT*, which have been proven involved in degradation of recombinant proteins (Jiang, 2001). Other common mutations are regarding tRNAs. The abundances of the various tRNAs differs between organisms. Forced over-expression of heterologous proteins can exhaust the pool of tRNAs rare to the host bacterium, causing the translation to stall. For this reason, several strains of the BL21 have been modified to contain extra copies of tRNA genes. These are collectively given the name “CodonPlus” (Terpe, 2006). As an example, the strain BL21-CodonPlus(DE3)-RIPL contains extra copies of the *argU*, *ileY*, *leuW* and *proL* tRNA genes, which collectively recognises the codons AGA, AGG, AUA, CCC and CUA.

In λDE3 cells, there is observed some T7 polymerase activity, even in the absence of IPTG. This ‘leaky’ expression from the *lac*UV5 promoter can cause the growth rate of the bacteria to decrease, or in worst-case cause cell death due to possible toxic effects of the heterologous protein. The amount of plasmid containing the target gene may therefore not be very high for these strains. For this reason, target genes are initially cloned into a host that does not contain the gene for T7 polymerase, to ensure the establishment and propagation of the gene carrying plasmid (Tabor, 2001). A commonly used host strain for this purpose is the *E. coli* DH5α. This strain carries *rec*A1 and *end*A1 and other mutations for increased insert stability, and has a high transformation efficiency (Taylor, 1993).

1.3.2 Vectors

A vector is a self-replicating DNA molecule which can be used to carry cloned genes or other fragments of DNA between hosts. Minimum requirements for a cloning vector are that it should be a reasonable small DNA molecule, easy to transfer between cells and straightforward generation and purification of large amounts of vector DNA. Plasmids are DNA molecules found in both prokaryotic and eukaryotic cells. Most are circular molecules of double stranded DNA, carrying genetic information but they are not chromosomes or part of the cell's permanent genome. The first vectors to be used were derivatives of small multicopy plasmids of bacteria, and they are still the most widespread. In addition to the natural characteristic, vector used today have been modified with artificial improvement making it easier to detect and select the cells carrying the vector. Genetic information not needed is removed, and genes encoding, for instance, antibiotic resistance are added. Antibiotic resistance is used for direct selection of clones containing the vector, by incubating the cells in media containing the antibiotic of choice. A second feature commonly added to plasmids is a 'multiple cloning site' (MCS). This is a sequence of DNA synthesized to include several sites for frequently used restriction enzymes (see section 1.3.4. below for more details on restriction enzymes). The rest of the vector should not contain any of the restriction sites used in the MCS, which ensures that the insert is cloned in the same location every time (Clark, 2010).

Many derivatives of plasmids and other vectors have been developed for use in recombinant gene technology, and finding the best fit can be difficult and is often done by 'trial-and-error'. For this thesis, two plasmids were initially chosen for the cloning of candidate genes. Plasmid pET21, mentioned above, is a much used plasmid for recombinant gene expression, as it utilizes the T7 RNA polymerase system for strong and regulated expression of the target gene. Other than the T7 promoter, this plasmid also contains the gene for ampicillin resistance (*amp*), a MCS with restriction site for *Ava*I, *Xho*I, *Not*I, *Eag*I, *Hind*III, *Sal*I, *Eco*RI and *Bam*H, a bacterial origin of replication (ori) and a phage ori (f1) (Dana-Farber/Harvard Cancer Center, 2004-2014). The second plasmid is the pUC57 plasmid, a derivative of pUC19. Its MCS contains six restriction sites with protruding 3'-ends, which are resistant to *E. coli* exonuclease III, and it also carries the gene *bla* for ampicillin resistance (GenScript).

1.3.4 Restriction endonucleases

Restriction endonucleases (RE) are enzymes that can be found in prokaryotes. Here they protect the organism from viruses and other infectious DNA molecules by recognising specific sequences in double-stranded DNA (restriction site) and cut the DNA at that site. They are classified into four different types (I-IV), where type II is the largest class with over 350 different prototypes known and over 200 of them being commercially available. It is this type of restriction enzymes that gave rise to recombinant DNA technology after their discovery in the 1970's (Loenen, 2014, Pingoud, 2014).

The choice of which cloning strategy to use can affect the choice of vector, as the needs for certain restriction sites to be present and the reading frame of the target gene to be compatible. The restriction enzymes used for cloning should not have sites within the target gene or more than needed to prepare the vector fragment, as this drastically decreasing the odds of re-ligating a functional vector.

1.4 Purification methods

In order to analyse and characterize recombinant proteins, separating them from other proteins produced in the host cell is often necessary. The utilization of the bio-valuable compounds identified through metagenomic analysis would not have been as successful if individual purification techniques had to be developed for each protein type. Affinity tags, such as the commonly used histidine tag (His-tag), allow for purification of most proteins without knowing anything about its biochemical properties (Arnau, 2006). The His-tag is added by modifying the recombinant DNA sequence to contain nucleotides encoding 5-15 histidine residues on the N'- or C'-terminus of the target gene. The His-tag can be positioned at either end of the gene sequence, but it is argued that the risk of it affecting the post-translational modifications is lower when applied at the stop-codon-end (Sancho, 1992). The purification is possible due to histidine's affinity for the metal ions Ni^{2+}, Co^{2+}, Cu^{2+} and Zn^{2+}, and is performed using immobilized metal-ion affinity chromatography (IMAC).

When the protein of interest is thermostable, as assumed for this study, it is possible to purify the protein sample by heat denaturation. By subjecting the sample to heat, e.g. 65°C, proteins not stable at this temperature will denature and become insoluble while thermostable proteins will remain soluble. Most of the proteins produced by mesophiles, such as *E. coli,* will denature at temperatures above 40 °C (Torrez, 2003).

2. AIM OF PROJECT

The aim of this master project was to identify thermostable peptidases encoded in an oil reservoir metagenomic sequence database, and produce these using recombinant gene technology.

This metagenomic DNA sequence library has previously been exploited to identify thermostable enzymes from other classes. As peptidases are ubiquitous enzymes, essential for all living organism, there was of interest to include gene candidates from this class. If successful production of the enzyme candidates is obtained, functional characterization is needed to further develop the proteins for application in industrial processes (figure 2.1).

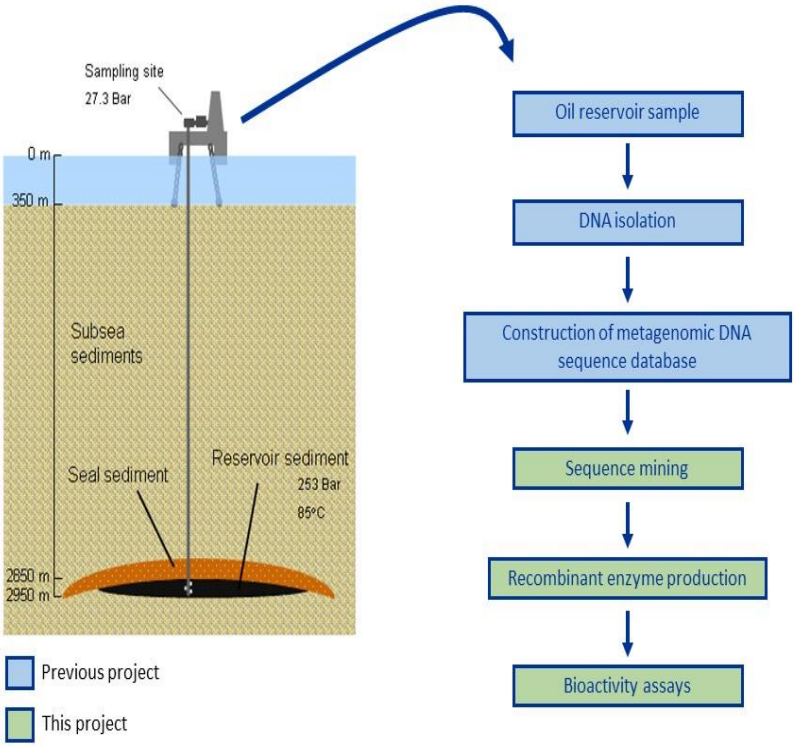

Figure 2.1: *Schematic overview of steps performed in I previous project to make the metagenomic DNA sequence library used in this project to identify putative gene sequences for thermostable peptidases. These candidates were expressed using recombinant gene technology and bioactivity assays was performed for functional characterization.*

3. MATERIALS AND METHODS

3.1 Materials

3.1.1 Chemicals

Chemicals used in this study are listed in table 3.1, together with the purity of the substance, its supplier and country of production.

Table 3.1: Chemicals.

Chemical	Purity	Supplier	Country
Acetic acid	99.9 %	VWR International	France
Agar Bacteriological	N/A	OXOID Ltd.	UK
Ampicillin sodium salt	≥91.0 %	AppliChem	Germany
Calcium chloride	99.8 %	VWR International	Belgium
D-(+)-Glucose anhydrous	≥99.5 %	VWR International	Belgium
EDTA (ethylene-diamine-tetraacetic acid)	≥99.5 %	Sigma-Aldrich	Germany
Ethanol	100.0 %	VWR International	France
Glycerol	≥99.5 %	VWR International	France
IPTG (Isopropyl β-D-1-thiogalactopyranoside)	≥98 %	VWR International	Belgium
iso-Propylalchohol	100 %	VWR International	France
Magnesium sulphate heptahydrate	≥99.5 %	VWR International	Belgium
Manganees II Cholide	≥98 %	Avantor Performance Materials B.V.	Mexico
MOPS (3-morpholinopropane-1-sulfonic acid)	≥99.5 %	Research Organics	USA
Potassium acetate	≥90.0 %	Merck KGaA	Germany
Potassium chloride	≥99.5 %	Merck KGaA	Germany
Rubidium chloride	≥99 %	AMRESCO	USA
Skim milk powder	N/A	OXOID Ltd.	UK
Sodium chloride	99.9 %	VWR International	Belgium
Sodium hydroxide	≥99 %	VWR International	UK
Trichloroacetic acid	≥99.5 %	Sigma-Aldrich	Germany
Tris-(hydroxymethyl) aminomethane	99 %	VWR International	Belgium
Tryptone	N/A	OXOID Ltd.	UK
Yeast extract	N/A	OXOID Ltd.	UK

3.1.2 *Escherichia coli* strains and plasmids

Bacterial strains and plasmids used for this study are listed in table 3.2, below.

Table 3.2: Microbial material.

Material	Relevant traits	Source
***E. coli* strains**		
DH5α	Cloning strain, *F´ proA+B+ lacI^q Δ lacZ M15/ fhuA2 Δ(lac-proAB) glnV gal R(zgb-210::Tn10)Tet^S endA1 thi-1 Δ(hsdS-mcrB)5*	NTNU, institute strain collection
BL21 (DE3) CodonPlus-RIPL	Expression host for pET vectors, *F- ompT hsdS(r_B^- m_B^-) dcm+ Tetr gal λ(DE3) endA Hte [argU proL Camr] [argU ileY leuW Strep/Specr]*	NTNU, institute strain collection
ER2566	Expression host for pET vectors, *F- λ- fhuA2 [lon] ompT lacZ::T7 gene 1 gal sulA11 Δ(mcrC-mrr)114::IS10 R(mcr-73::miniTn10-TetS)2 R(zgb-210::Tn10)(TetS) endA1 [dcm]*	NTNU, institute strain collection
Plasmids		
pUC57	Cloning vector, *bla* (Amp^R)	GenScript
pUC57_FragT1	pUC57 derivate carrying a transposon fragment, positive control in transformation, *bla* (Amp^R)	GenScript
pOD1	pET21 derivate carrying the *seW* gene, T7 expression vector, *bla* (Amp^R)	Other project at NTNU
pET21_APT01	pET21 derivate carrying the APT01 gene, T7 expression vector, *bla* (Amp^R)	This work
pET21_CPT01	pET21 derivate carrying the CPT01 gene, T7 expression vector, *bla* (Amp^R)	This work
pET21_SLP01	pET21 derivate carrying the SLP01 gene, T7 expression vector, *bla* (Amp^R)	This work

3.1.3 Media and solutions

Water purified by Milli-Q purifiers, was used to prepare solutions and media if not otherwise specified. The components of all media prepared and used for this project are listed in table 3.3.

Table 3.3: *Compounds and their amounts needed to prepare the media Luria Broth (LB), LB agar (LA), skim milk agar (MA), Psi* and Super Optimal Broth (SOC).*

	LB (g/L)	**LA (g/L)**	**MA (g/L)**	**Psi* (g/L)**	**SOC (g/L)**
Yeast extract	5	5		5	5
Tryptone	10	10		20	20
NaCl	10	10			0.5
Agar		15	15		
Skim milk powder			16		
$MgSO_4$				10.26	4.8
Glucose					3.603

*Psi represents a simple medium similar to LB, with added $MgSO_4$, to increase cell density (Studier, 2005).

The compounds listed in table 3.3 were weighed in and mixed before the medium was autoclaved. For the SOC medium, the pH was adjusted to 7.6 using KOH, before autoclaving. Media not containing agar were stored at room temperature. When needed, premade solution of ampicillin was added to the medium to a final concentration of 100 µg/ml. To minimize degradation and ensure maximum effect of the antibiotic, Ampicillin was added after the medium had cooled to <50 °C (following autoclaving) and just before inoculation with bacteria.

Media containing agar (LA and MA) were cooled to 50 °C, and ampicillin added, before casting the petri dishes (20-25 ml/dish). The plates were left to solidify before being stored upside down in the dark at 4 °C. For some of the MA plates, 40 µl 0.1 M IPTG was added, after the plates had solidified.

The solutions TFB I and II were used to make competent cells for transformation. Table 3.4 lists the components, and their amount, mixed to prepare these solutions.

Table 3.4: *Components and their amount used to prepare 200 and 100 ml of TFB I and II, respectively.*

	TFB I	**TFB II**
Potassium acetate	0.588 g	-
Rubidium chloride	2.42 g	0.121 g
Calcium chloride	0.294 g	1.1 g
Manganese (II) chloride	15 g	-
MOPS	-	0.21 g
Glycerol (15 v/v)	30 ml	15 ml

The solids were weighed out and mixed with most of the water. The pH was adjusted to 5.8 using diluted acetic acid for TFB I, and pH 6.5 using NaOH for TFB II, before additional water was added to make a total of 200 and 100 ml, respectively. The solutions were then filter sterilized and stored at 4 °C.

Buffers TE, TAE and Tris-HCl was used for this study. Table 3.5 lists the components, and their amount, mixed to prepare these solutions.

Table 3.5*: Components and their amount to prepare the buffers Tris-HCl, TE and TAE. pH was adjusted to 8 using the solutions listed.*

	Tris-HCl (g/L)	**TE (g/L)**	**TAE (g/L)**
Tris-(hydroxymethyl) aminomethane	121.14	12.11	4.85
EDTA	-	7.44	0.37
Acetic acid	-	-	1.20
pH adjusted to	8	8	8
pH adjusted with	HCl	NaOH	NaOH

After mixing the solutions, the Tris-HCl and TE buffer was filter sterilized and stored at room temperature.

3.2 Methods

3.2.1 Bioprospecting for peptidases in the oil reservoir metagenome database

Searches for enzymes that are not membrane bound, belong to the enzyme classification number 3.4, and are smaller than 1000 amino acid, was performed in the protein sequence database UniProt. Peptidase sequence candidates were identified, using only annotated FASTA nucleotide sequences of known proteases. Sequence similarity searches were performed using BLAST, to identify local alignments of the chosen peptidases to nucleotide sequences in the given metagenomic database from oil reservoirs 1 and 2. ORF Finder was used to predict open reading frames of the contigs showing high percentage of identity to the chosen peptidase sequences, and candidates displaying possible truncated (non-complete) protein sequences were eliminated from the candidate list. The FASTA protein sequence was used for BLAST-based homology searches against proteins in the NCBI Protein Data Bank. Highest percentage identity structures were used to check the presence of conserved active site residues in the candidate gene sequence. The presence of start and stop codon and Shine-Dalgarno sequence was also checked before listing the protein as a finale candidate. The final elimination was based on the enzyme class and function, homology to species of interest, as well as the length of the gene (budget limitations).

The synthesis of the putative gene candidates, and cloning into pUC57 vectors, were ordered from GenScript (www.genescript.com). Each gene was constructed with addition of nucleotides providing restriction sites for *Nde*I, *Xho*I, *Bgl*II and *Bcl*I, a His-tag and protective bases flanking the gene, before it was cloned into pUC57 plasmids using EcoRV. The complete DNA sequence, including restriction sites, His-tag and protective bases, for each of the gene candidates can be seen in appendix A.1.

3.2.2 Preparing chemically competent cells

Three strains of *Escherichia coli*, DH5α, BL21 (DE3) codon plus RIPL, and ER2566 (table 3.2), were used for preparing chemically competent cells for transformation. Cells of the frozen bacteria strain were used to inoculate 5 ml of psi medium (table 3.3) and incubated at 37 °C at 225rpm overnight. 1 ml of the overnight (O/N = 12-16 hours) culture was inoculated in 100 ml of psi medium, and incubation was continued at 37 °C at 225 rpm until optical density (OD) at 600 nm reached ~0.4. The cells were incubated on ice for 15 minutes before centrifuged (4500 rpm, 5 min, 4 °C) using a Biofuge primoR centrifuge from Heraeus (rotor #7590). While on ice, the cell pellet was re-suspended in 40 ml cold TFB I, centrifuged (4500 rpm, 5 min, 4 °C) and re-suspended in 3 ml cold TFB II. The cells where transferred to sterile 1.5 ml Eppendorf tubes, 100 µl/tube, and snap frozen using liquid nitrogen. All competent cells were stored at -80 °C.

3.2.3 Sub-cloning of candidate genes in pET21

All the pUC57_gene plasmids were shipped from GenScript in tubes containing of 4 µg dried DNA. TE buffer (40 µl) was added to dissolve and dilute the plasmids to a final concentration of 100 ng/µl. Plasmid preparations, already isolated, of pUC57_APT01, pUC57_CPT01 and pOD1 were used for this cloning step. These preparations were obtained using the Wizard Plus SV Minipreps DNA Purification System from Promega, from *E. coli* DH5α cells transformed in the initial attempt to include the His-tag. DNA concentrations of these isolations were measured using the NanoDrop 1000 Spectrophotometer (Thermo Scientific). The third candidate, pUC57_SLP01, was used directly from the tube provided by GenScript.

Enzyme digestion of the pUC57_gene and the pOD1 plasmids was performed using restriction enzymes *Nde*I and *Xho*I as shown in table 3.6. The amounts of plasmids used were adjusted from concentration of DNA in the plasmid preparations, to give a vector-to-gene ratio of 1:9. The reaction mixtures were mixed, and incubated in a 37 °C water bath for 90 minutes.

Table 3.6: *Reaction mixtures used to digest pUC57_APT01, pUC57_CPT01, pUC57_SLP01 and pOD01 using the restriction enzymes NdeI and XhoI.*

	pUC57_ APT01 (µl)	**pUC57_ CPT01 (µl)**	**pOD1 (µl) (for APT01 and CPT01)**	**pUC57_ SLP01 (µl)**	**pOD1 (µl) (for SLP01)**
Plasmid prep.	30	37	50	5	8
NEB Buffer 3.1	6	6	6	1,5	1,5
dH_2O	22	15	2	6,5	3,5
***Nde*I**	1	1	1	1	1
***Xho*I**	1	1	1	1	1
Total	60	60	60	15	15
Loading dye	7.5	7.5	7.5	1.85	1.85

The fragments from the enzyme digest were separated using agarose gel electrophoresis. The gel was prepared using agarose solution containing 0.8 % agarose and 20 µl Gel Green (per 400 ml). Sample loading dye (10mM Tris-HCl pH7.6, 0.03 % bromophenol blue, 0.03 % xylene cyanol FF, 60 % glycerol and 60 mM EDTA) was added to the reaction mixtures as seen in table 3.6, before the samples were added to separate gel wells. Electrophoresis was run at 75 volts for approximately 60 minutes.

A molecular imager (Gel Doc^{TM} XR+ system from BioRad) was used to visualize the DNA bands by UV light. Fragments were excised from the gel and purified using the QIAquick Gel Extraction Kit Protocol (QIAGEN). The standard protocol (for microcentrifuge) provided by the manufacturer was followed with two modifications. The washing step 10 was repeated using 500 µl of Buffer PE, and that the DNA was eluted in step 13 using 30 µl sterile deionized nuclease free water to increase its concentration. The eluted DNA was used directly for ligation of candidate genes into pET21 expression vector to form pET21_*gene*. The ligation mixture contained 30 µl of the gene fragment, 10 µl of the vector fragment, 1 µl T4 DNA ligase and 4 µl T4 DNA ligase buffer. For the negative control reaction, dH_2O was added in place of the gene fragment. The ligation mixtures were incubated O/N over ice. After incubation, the ligation mixtures were stored at -20 °C until transformation later the same day.

3.2.4 Transformation of *E. coli* DH5α with pET21_gene

Chemically prepared competent *E. coli* DH5α cells were transformed with recombinant pET21_gene plasmids, using heat shock transformation. Tubes containing frozen competent cells (section 3.2.2) were thawed on ice before 50 µl each were transferred to new pre-chilled 1.5 ml Eppendorf tubes for each transformation reaction. While on ice, the ligation mix for pET21_APT01 (10 µl), pET21_CPT01 (10 µl), pET21_SLP01 (5 µl) and re-ligated pOD1 (10 µl), were added to the cells and incubated for 30 minutes on ice. For the negative control, competent cells with no addition of plasmid DNA was used, while pUC57 plasmids carrying a transposon fragment (pUC57_FragT1) obtained from GenScript for a previous project, was used as a positive control. The cells were exposed to a heat shock in a water bath at 42 °C for 45 seconds, then incubated for two minutes on ice. SOC medium (250 µl), pre-warmed to 37 °C, was added, and the cells were incubated at 37 °C at 225 rpm for one hour, before being plated on LA plates for growth and selection of successful transformants. The plates were incubated at 37 °C O/N, and then stored in at 4 °C until further use.

3.2.5 Verification of correct insert

Verification of correct pET21_gene construct was confirmed by restriction digestion and separation by gel electrophoresis. Using sterile toothpicks, twelve single colonies for each of the three pET21_gene plasmids, and the pOD1 control, were streaked out on a new plate and inoculated into 13 ml tubes containing 5 ml of LB medium and ampicillin. Both plates and tubes where incubated at 37 °C O/N, the tubes in a shaking incubator at 225 rpm. The plates were moved to 4 °C the next day, and stored at this temperature. Plasmid preparations were made for each of the O/N cultures using the Promega Wizard® Plus SV Minipreps DNA Purification System Protocol. The centrifugation, in step 3.B.1, was preformed directly on the 13 ml tubes used for incubation of the O/N cultures. The supernatant was discarded and 250 µl Cell Re-suspension Solution was used to completely re-suspend the cells, before transferring the entire sample to a sterile Eppendorf tube. The DNA was eluted using 100 µl nuclease free water.

Restriction digestion was performed on the plasmids, using *Bgl*II restriction enzyme for pET21_APT01 and pET21_CPT01, and *Nde*I and *Xho*I for pET21_SLP01, due to differences in restriction sites. All samples were incubated in a water bath at 37 °C for 2 hours, before the samples were separated using gel electrophoresis as described above.

A larger volume of plasmids preparation stock was prepared for recombinant protein production. Based on the results from the verification step, single colonies originating from one of the newly streaked CFUs carrying the correct insert were used to prepare 4x 5 ml O/N cultures, for each pET21_gene candidate. Plasmid preparation was performed as described above. The concentrations of the plasmid preparations were measured using NanoDrop™ 1000 Spectrophotometer (Thermo Scientific).

3.2.6 Recombinant protein production

The BL21 (DE3) codon plus RIPL strain of *E.coli* was chosen for production of the recombinant proteins. Cells of this strain were made competent for transformation by the procedure explained in 3.2.2. The final plasmid preparations from 3.2.5 were used to transform these cells by heat shock transformation as in 3.2.4, but changing the heat shock duration to 20 seconds. Single colonies of transformed BL21 cells, were used to prepare O/N pre-cultures (5 ml LB +amp, at 37 °C, 225 rpm). From these, 1 ml was used to inoculate 100 ml LB with ampicillin in 500 ml Erlenmeyer flasks with baffles. The cultures were incubated in a shaking incubator at 37 °C and 225 rpm. The optical density was measured spectrophotometrically at a wavelength of 600 nm. When the OD_{600nm} value was close to one, the 100 ml culture in the flasks was divided into 2x 50 ml and transferred to two 250 ml flasks. IPTG was added, to a concentration of 0.5 mM, to one of the two parallels, while the respective other 50 ml culture was handled equally further without induction and served as a negative control. The eight flasks where subsequently incubated (37 °C, 225 rpm) and the absorbance was measured every hour until the bacteria cultures reached a stationary growth phase. The cultures were then transferred to 50 ml tubes and centrifuged at 6500 rpm for 10 min at 23 °C, using a Biofuge primoR centrifuge from Heraeus (rotor #7590). The supernatant was discarded and the pellet was stored at -20 °C until cell lysis was performed the following day.

3.2.7 Isolation of proteins

To isolate the proteins produced intracellularly, the cells were lysed by sonication. The pellet samples obtained in 3.2.6 were thawed and cells of each re-suspended in 10 ml 0.1 M Tris-HCl buffer (pH 8) and sonicated for 7x 1 minute with intermittent mixing. Sonication was performed using a Branson Sonifier equipped with a double stepped microtip (3 mm), set to duty cycle at 40% and output control set to 4. Centrifugation for 30 minutes at 20 000 g and 4 °C followed. Due to uncertainties regarding the solubility of the produced target proteins, both the supernatant and pellet samples were saved for further analysis. The supernatant was distributed into 1.5 ml Eppendorf tubes, some stored at 4 °C and some at -80 °C. The pellet, representing the insoluble fraction, was stored at -80 °C.

3.2.8 Detection of protein using SDS-PAGE

SDS-PAGE (sodium dodecyl sulphate polyacrylamide gel electrophoresis) was used to detect production of target proteins. Premade gels (12 % acrylamide, 12 wells) and running buffer (FB60500), provided from C.B.S. Scientific, was used for all SDS-PAGE experiments preformed in this study. The protein size reference marker used was the Precision Plus Protein Unstained Standards from Bio-Rad, and the BioSafe™ Coomassie G-250 Staining Solution from the same supplier was used to stain the proteins separated in the gel.

Samples from both pellet and supernatant of induced and non-induced cultures of all three candidates and the pOD1 control were prepared for electrophoresis. For each of the eight supernatant samples, one tube stored at -80 °C was thawed and 24 µl was transferred to a new sterile Eppendorf tube. SDS-PAGE 5x DTT loading buffer (0.25 % Bromophenol blue, 0.5 M DTT, 10 % SDS and 50 % glycerol) was added (6 µl). For the pellet samples, a small amount of frozen pellet was suspended in an Eppendorf tube containing 80 µl of 0.1 M Tris-HCl (pH 8), and 20 µl of loading buffer was added to each. All 16 samples were incubated on a heat block at 95 °C for 10 minutes, then centrifuged using a 5415R microcentrifuge from Eppendorf, for 5 minutes at 5000 rpm at 23 °C. For each sample, 5-25 µl was loaded to a well and the SDS-PAGE was run at 175 volts for 35-45 minutes.

To increase the concentration of protein in the supernatant samples, TCA precipitation was performed prior to the preparation of the samples with DTT loading dye. A protocol from the Björkman Group at the California Institute of Technology (BjörkmanGroup, 2001) was followed, using 800 µl of protein sample and 200 µl of TCA stock (5 g Trichloroacetic acid dissolved in 3.5 ml H_2O). When the precipitation procedure was complete, 50 µl of 0.1 M Tris-HCl (pH 8) and 15 µl of DTT loading buffer was added and the samples where heated, centrifuged and run on a gel as explained above.

3.2.9 Recombinant protein production using other parameters

An alternative production host strain was tested for the SLP01 candidate. The plasmid pET21_SLP01 and the control plasmid pOD1 were used to transform competent cells of *E. coli* ER2566, as described in 3.2.4, with a heat shock at 42 °C for 20 seconds. The second attempt of recombinant protein production was carried out in BL21 cells for all three constructs and ER2566 SLP01 candidate. The method was equal to that presented in 3.2.6, with the exception that the temperature during incubation after induction with IPTG was set to 16 °C. Incubation was ended and cells were centrifuged (as above) when OD_{600nm}-values stabilized. Lysing of cells using sonication and detecting of proteins by SDS-PAGE was done as described in 3.2.7 and 3.2.8.

A third attempt was performed using 1 litre cultures. Transformed *E. coli* BL21 cells with pET21_APT01, pET21_CPT01 and pOD1 was inoculated into 5 ml of LB broth containing 100 µg/ml ampicillin at 37 °C for 16 hours. The following day, all of the O/N culture (5 ml) was added to a 3000 ml Erlenmayer flask, with baffles, containing 1 litre of 3x LB and ampicillin (three times the amount of tryptone and yeast extract, but identical amounts of NaCl per litre). The cultures grown in a shaking incubator at 37 °C at 180 rpm until OD_{600} was close to three. IPTG was added to each flask to a total concentration of 0.5 mM, and the cultures were subsequently incubated at 16 °C and 180 rpm for five hours. Absorbance was measured once every hour. Centrifugation was done at 12 000 g, for 30 minutes at 14-16 °C. The supernatant was discarded, and the pellet were re-suspended using 40 ml, 1 M Tris-HCl (pH 8) and stored at -20 °C, until sonication the following day. Isolation and detection of proteins was performed as described in 3.2.7 and 3.2.8, with the duration of centrifugation after heat block treatment increased to 20 minutes.

3.2.10 Heat denaturation series

A frozen tube of the induced supernatant sample of the BL21 pET21_APT01 culture, from the protein production using 50 ml cultures and incubation at 16 °C (section 3.2.9), was thawed and 200 µl was transferred to sterile Eppendorf tubes. Using heat blocks, samples were subjected to either 65 or 80 °C for different durations of time, as shown in table 3.7, and subsequently placed on ice until further processing.

Table 3.7: *Heat denaturation at 65 or 80 °C for 5, 15, 30, 60 or 120 minutes. Sample 11 was a control sample, and was not subjected to heat prior to SDS DTT treatment.*

Time	**Temperature**	
(min)	**65 °C**	**80 °C**
5	1	6
15	2	7
30	3	8
60	4	9
120	5	10

After heat treatment, the 11 samples were centrifuged using a 5415R microcentrifuge from Eppendorf for 5 minutes at 13 000 rpm and 23 °C. Each sample was then prepared and analysed using SDS-PAGE as explained in 3.2.8.

3.2.10 Activity assay using skim milk agar

To investigate if the *E. coli* BL21 cells transformed with the pET21_gene (section 3.2.6) secreted active proteases, single colonies were streaked out on MA plates with added IPTG. BL21 cells transformed with pOD1, and purified proteinase K from *Tritirachium album* (Sigma-Aldrich) was applied as a negative and positive control, respectively. Three parallels where made, each incubated at three different temperatures (37, 45 and 60 °C) for up to 48 hours.

This activity assay was also used for the supernatant samples obtained from the protein production and isolation (sections 3.2.6-3.2.9). Drops of supernatant samples were added to MA plates for both induced and non-induced samples for all candidates, to test for presence of active peptidases. The negative and positive controls were equal to those mentioned above. The plates were incubated at 37 °C for 48 hours.

4. RESULTS AND DISCUSSION

4.1 Data mining

The aim of this study was to identify and produce new and thermostable peptidases. Many steps are included in the process of identifying novel enzymes and producing these in host cells. Figure 4.1 shows a flow diagram of the major steps involved in the method used for this thesis to search for and select gene candidates of putative peptidases. By aligning gene sequences encoding known peptidases to the contigs in the given metagenomic DNA sequence library, many different candidates were identified. Selecting contigs that were not truncated or genes that was not likely to encode membrane bound enzymes increased the possibility of finding candidates that could successfully be produced in E. coli in a functional form. Also for this reason, sequences that had high sequence identity and coverage to known peptidase genes, encoded conserved active site residues, as well as presence of start- and stop codons and Shine-Dalgarno sequence were selected.

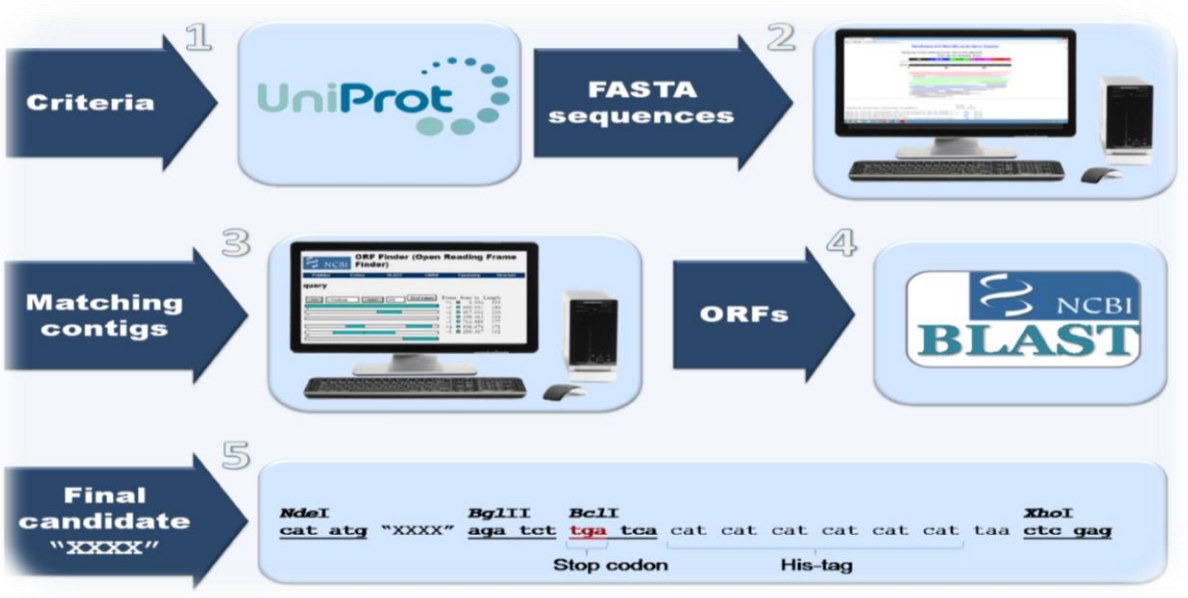

Figure 4.1: *The five major steps used for the bioprospecting of novel peptidases in the metagenomic database. 1: Selected criteria were used to search for annotated peptidases in UniProt. 2: BLAST was used to match the nucleotide FASTA sequence of these peptidases to contigs in the metagenomic DNA library. 3: The nucleotide sequence of the matching contigs was subjected to NCBI's ORF Finder to locate probable open reading frame of gene of interest. 4: Identified ORFs were validated by the use of BLAST against the NCBI database. 5: Selected candidates were synthesized by GenScript to include the restriction sites and His-tag shown. "XXX" represents the candidate gene sequence.*

Through bioprospecting of the metagenomic DNA sequence library, administrated by the Core facility bioinformatics at NTNU (BioCore), three protease candidates were chosen for cloning, production and characterization. These three putative peptidase genes each have a high sequence similarity to a known peptidase of the enzyme sub-classes: subtilisin-like protease, carboxypeptidase or aminopeptidase. The three gene candidates are listed in table 4.1.

Table 4.1: *Predicted function and enzyme class, given name for this study and sequence length in bp of the three gene candidates chosen.*

Enzyme class	Given name	Function	Sequence length (bp)
Subtilisin-like protease	SLP01	Cleaves peptide bonds using serine as a nucleophile	2028
Carboxypeptidase	CPT01	Cleaves peptide bonds at C-terminus	1542
Aminopeptidase	APT01	Cleaves peptide bonds at N-terminus	1530

A minimum requirement when selecting the gene candidates was that they included the necessary elements described earlier (start/stop codon, Shine-Dalgarno sequence, active site), and had a high – but not too high – sequence identity to a known peptidase. The greater the homology, the better chance of identifying genes encoding functional peptidases. However, as the aim of this project is to identify novel enzymes, some differences in sequence should be present. It was also the intention to choose candidates from different sub-classes of peptidases, to have the possibility to characterize the different activities.

4.1.1 Candidate SLP01

The peptide sequence of the putative peptidase SLP01 is characterized by 100 % coverage and 85 % identity with an annotated subtilisin-like serine protease (662 aa, ~71 kDa) of *Thermococcus kodakaraensis KOD1*. *T. kodakaraensis* is a hyperthermophile archaeon that has an optimum temperature of 85 °C, indicating that its enzymes are adapted to high temperatures. The SLP01 peptide sequence showed 88 % identity with 92 % coverage to the crystallographic structure of chain A of the protein Pron-Tk-Sp from the same archaea species. Both start and stop codon as well as a Shine-Dalgarno sequence could be clearly identified within the contig, and the serine, histidine and aspartate residues of the active site are conserved (H180, S359, D147, respectively) (Foophow, 2010). Having a great percentage of coverage, and quite high result for identity to a hyperthermophile for both nucleotide and protein sequence, makes this gene a promising candidate, taking into account that the aim of this project was to identify thermostable peptidases. This candidate is the only endopeptidase of the three candidates chosen.

4.1.2 Candidate CPT01

The CPT01 gene is part of a contig in the metagenomic database annotated to the bacterium *Kosmotoga olearia*, strain TBF 19.5.1, and has 100% coverage and 100% id with Carboxypeptidase Taq (500 aa, ~58 kDa) from this bacterium. *K. oleria* was isolated from oil production fluid from a platform in the North Sea. It is a thermophile with optimum temperature of 65 °C (DiPippo, 2009). It is not surprising that the CPT01 gene had this high degree of homology to the *K. oleria*, as it has previously been observed that the DNA sequences annotated to *K. oleria* earlier in this project, have high similarities to the published species reference genome of TBF 19.5.1 (Kotlar, 2011). The CPT01 gene sequence is of interest, despite it having 100 % identity and coverage with the enzyme from *K. oleria,* since this organism is newly sequenced and has not been studied to great extent, and is known to be associated with oil reservoirs.

The gene's putative protein structure has 98 % coverage and 45 % identity with Chain A of the *Thermus thermophilus* M32 carboxypeptidase, which is a metallo-peptidase. The start and stop codons are located within the contig, but parts of the Shine-Dalgarno sequence seem to be missing as the start of the contig is only 10 base pairs upstream of the start codon. All residues in the *T. thermophilus* M32 carboxypeptidase

associated with the active and its metal binding site are conserved in the CPT01 sequence (Lee, 2009).

4.1.3 Candidate APT01

The gene candidate APT01 has high percentage coverage and identity, 97 % and 77 %, respectively, to the annotated gene for cytosolic aminopeptidase (497 aa, ~53 kDa) of the DSM 2380 strain of *Pelobacter carbinolicus*. This bacterium is a mesophile, thriving at temperatures between 25-40 °C. Both the start and stop codon as well as the Shine-Dalgarno sequence are well defined within the contig. The suggested protein structure has 97 % coverage and 46 % identity with the structurally solved chain A of the Leucine Aminopeptidase from *Pseudomonas putida*. The active site residues of the homologue Leucine aminopeptidase are conserved within the sequence of candidate APT01 (Kale, 2010). This candidate has less homology to annotated nucleotide and peptide sequences, compared to the other two. Which makes it an interesting choice due to less certainty when it comes to activity and function. Aminopeptidases are involved in many cellular processes and are therefore promising candidates for the pharmaceutical industry.

4.2 Cloning of candidate genes into pET21 plasmid

A set of vector and host strains was selected for cloning and recombinant gene expression. The plan was to clone the gene candidates into the pET21 vector using two different routes: One where the genes were directly excised from pUC57 and ligated into pET21, and the second where a modification step to include a His-tag was first to be carried out before the cloning into pET21. These two routes and further steps planned for production and isolation of the enzymes are shown in figure 4.2.

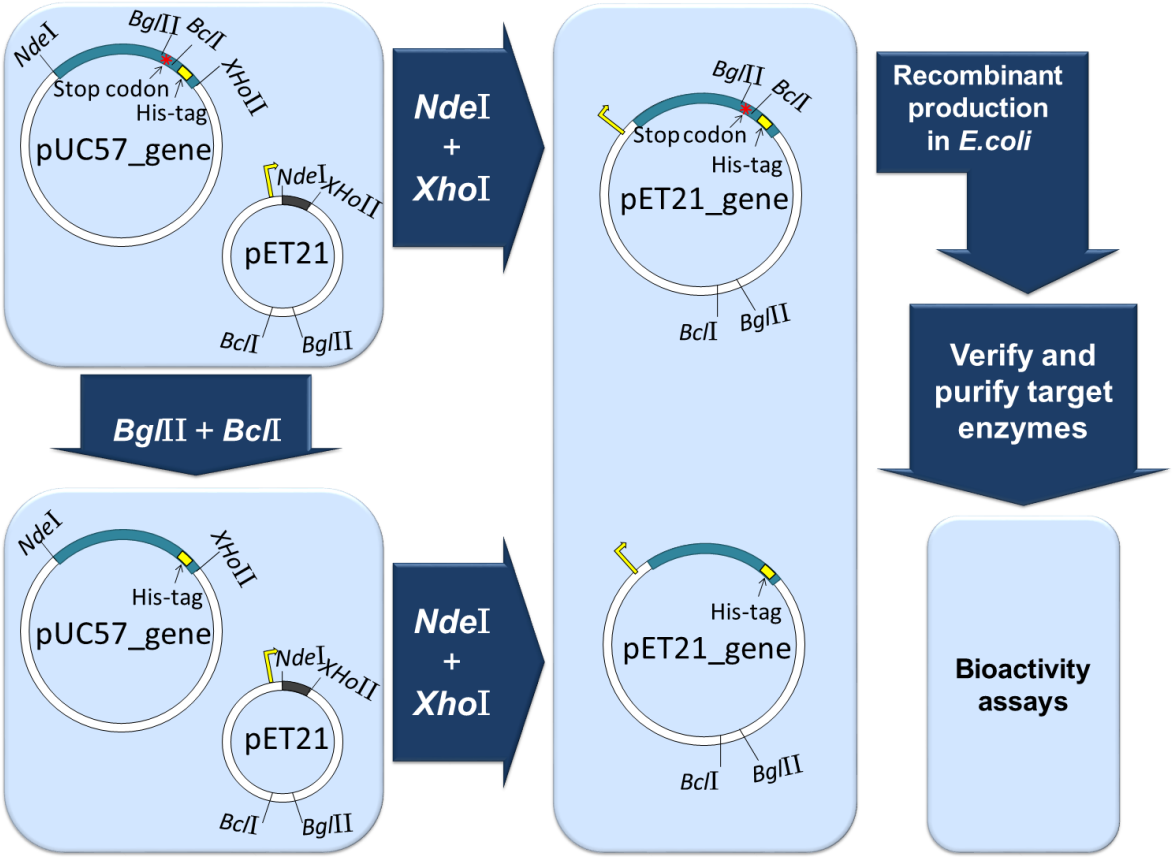

Figure 4.2: *Steps for the cloning and expression of the recombinant proteins. Depending on if the gene of interest, marked in dark blue, was cut with restriction enzymes* BglII *and* BclI *and re-ligated before being excised from pUC57, with* NdeI *and* XhoI*, and ligated into pET21, the resulting peptide would or would not contain a His-tag, respectively.*

The chosen gene candidates were synthesized and cloned into pUC57 plasmids by GenScript (www.genscript.com). Nucleotides making up the active site for the restriction enzyme *Nde*I were added to the candidate genes where the last three bases made up the start codon initiating transcription, and the same was done for the restriction enzyme *Xho*I following the stop codon. This allowed the genes to easily be cloned into a different plasmid at a later stage. A modification possibility for a His-tag

was added at the end of the gene. This included the restriction site for the enzymes *Bgl*II and *Bcl*I followed by a sequence coding for six histidine residues. The three first bases of the *Bcl*I restriction site also represents a stop codon, terminating the transcription before the His-tag if not removed from the sequence. Before each candidate was cloned into the plasmid pUC57 by using *Eco*RV, bases were added to either end of the candidate gene for protection. GenScript performed codon optimization of the ordered genes for optimized expression in *E. coli*. Distribution of the codon usage frequencies along the sequence was adjusted to 0.89, 0.87 and 0.90 for APT01, CPT01 and SLP01, respectively. Here values over 0.8 are considered good, and 1.0 is perfect. The GC content was also adjusted to give average values between 30-70 percent.

Due to difficulties with ineffective digestion by the restriction enzyme *Bcl*I, likely due to methylation of the restriction sites, this modification step was put on hold. The second (His-tag free) route of the study was prioritized. The plasmid pOD1, containing a gene for selenoprotein SelW, was utilised in this project. When digested with restriction enzymes *Nde*I and *Xho*I the *selW* gene was removed, leading to the vector backbone of pET21. For this report, the vector will be referred to as pOD1 as long as it is carrying the *selW* gene, and pET21 when the gene is removed. Although the initial attempt on obtaining the His-tag modifications did not succeed, the plasmid preparations of pUC57_APT01, pUC57_CPT01 and pOD1 that were made, were used for cloning the APT01 and CPT01 gene into pET21. The concentration of these plasmid preparations can be seen in table 4.2. The concentration of pUC57_SLP01 was 100 ng/µl.

Table 4.2: *Concentration of plasmid DNA for the preparations made of the pUC57_APT01, pUC57_CPT01 and pOD1 plasmid produced in* E. coli *DH5α.*

Plasmid	Concentration (ng/µl)
pUC57_APT01	111,7
pUC57_CPT01	87,1
pOD1	21,7

4.2.1 Vector construction

After both the pUC57 plasmids, containing the genes of interest, and the pOD1 plasmids had been digested by the restriction enzymes *Nde*I and *Xho*I, the DNA fragments were separated using gel electrophoresis. Images of the obtained gels are shown in figure 4.3 and 4.4.

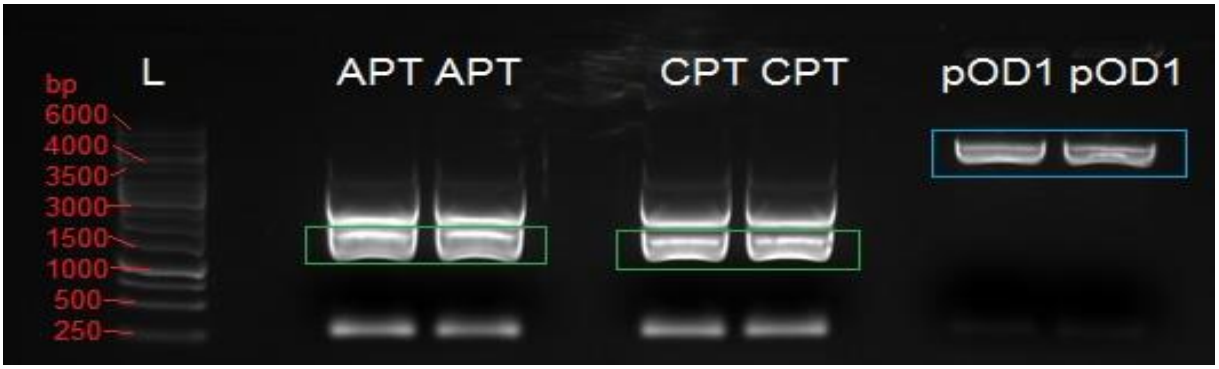

Figure 4.3: *Gel electrophoresis run with fragments of plasmids cut with restriction enzyme* NdeI *and* XhoI*. The 1kb DNA ladder from GeneRuler ™ was a run in the lanes marked "L", the size of the given fragments is given in number of base pairs in red. The plasmid pUC57 containing APT01 and CPT01 was run in the lanes marked "APT" and "CPT", respectively, while the plasmid pOD1 was run in the lanes marked "pOD1". The gene fragments of interest is marked with green boxes, while the vector fragments of interest is marked with blue boxes.*

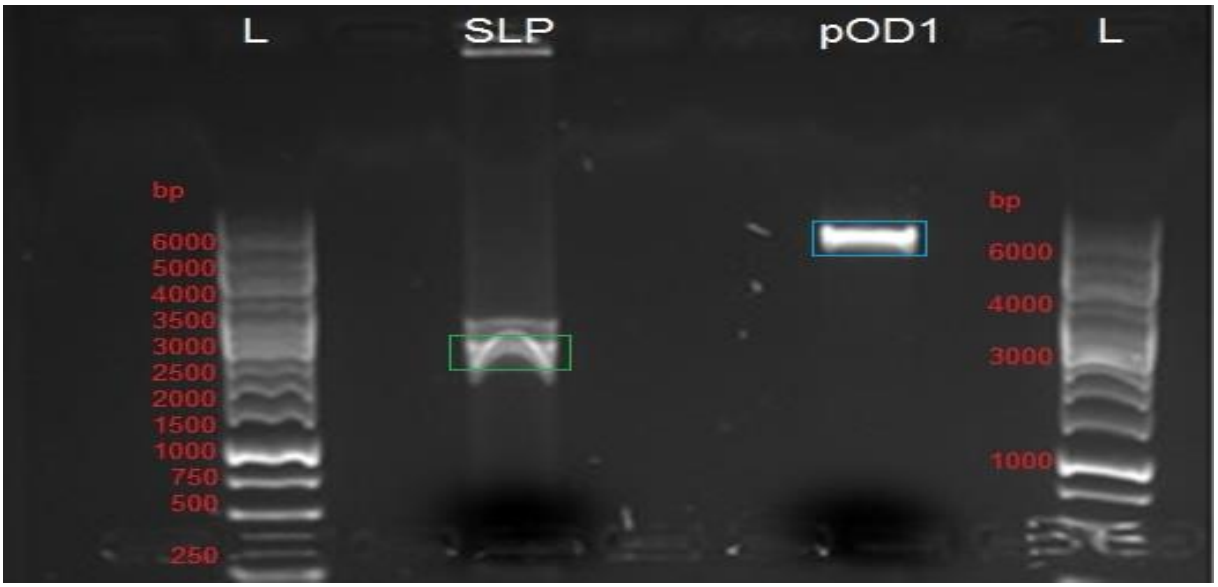

Figure 4.4: *Gel electrophoresis run with fragments of plasmids cut with restriction enzyme* NdeI *and* XhoI*. The 1kb DNA ladder from GeneRuler ™ was a run in the lanes marked "L", the size of the given fragments is given in number of base pairs in red. The plasmid pUC57 containing SLP01 was run in the lane marked "SLP", while the plasmid pOD1 was run in the lane marked "pOD1". The gene fragment of interest is marked with a green box, while the vector fragment of interest is marked with a blue box.*

The expected fragment sizes after this restriction enzyme digestion and gel separation for the pUC57 plasmids were; 1523 bp for APT01, 1535 bp for CPT01, 2021 bp for SLP01 and 2727 bp for the remaining vector (pUC57). For pOD1, the fragment of interest, the vector itself, was expected to have a size of 5365 bp, while the excised fragment should have showed up as a much smaller band with a size of 266 base pairs. With its additional gene excised, the pOD1 vector was equal to, and will from now on be referred to, as pET21.

The visible bands in figure 4.3 and 4.4 match the expected results well. Two fragments of expected size for pUC57_APT01 and pUC57_CPT01 are visible in figure 4.3, were the gene fragment of interest is marked with green boxes. This suggests that the digestion was successful. In addition, a band of ~300 bp size is visible for both samples. This band can represent fragments made by 'star activity' of the restriction enzyme(s). Star activity represents less specific activity for restriction enzymes, resulting in cleavage of sequences similar but not identical to their restriction site. For pUC57_SLP01 it is difficult to distinguish between the two bands representing the gene and the vector. The large amount of DNA applied to the single well makes each band very strong, and considering the fact that the vector is only 706 bp longer than the SLP01 gene, these two fragments may overlap. For pOD1, only one band is visible (~5000 bp) in both figures. This could indicate that one of the restriction enzymes was not able to digest the DNA successfully, and that pOD1 have become linearized but still carries the *seM* gene. On the other hand, the smaller the DNA fragment, the lesser material there is for GelGreen to bind to. Hence, a certain number of short fragments will result in a weaker band compared to the same number of longer DNA fragments, as the longer fragments consists of more nucleotides. Some pOD1 vectors can have been correctly digested, without making a visible band for the *seM* gene fragment due to the low amount of DNA to detect.

The fragments of interest, marked with green and blue boxes, were excised from the gel and rinsed before the DNA of the three different gene fragments was isolated and ligated with the pET21 plasmid, separately, to make pET21_APT01, pET21_CPT01 and pET21_SLP01.

4.2.2 Transformation and verification

To verify that correct restriction digestion and ligation took place, the pET21_gene plasmids were transformed into competent *E. coli* DH5a cells by heat shock transformation. The number of CFU formed after the transformation can be found in table 4.2. Resulting CFUs were picked and incubated O/N in LB containing ampicillin, and DNA was purified for each of the three gene candidates. These clones were checked by restriction mapping, where the DNA was digested again with *Bgl*II restriction enzyme for pET21_APT01 and _CPT01, and *Nde*I and *Xho*I for pET21_SLP01. The resulting fragments where separated using gel electrophoresis. The obtained fragments of this verification step can be seen in figure 4.5 and 4.6.

Table 4.2: *The number of CFU after the transformation of* E. coli DH5a *with pET21_gene. All negative controls gave zero CFU.*

Gene	Culture (µl)	CFU
APT01	50	138
	200	579
CPT01	50	51
	200	296
SLP01	50	16
	200	25

Compared to the bacterial cells transformed with APT01 and CPT01, the number of CFU from the transformation of SLP01 was low. This can be due to few correct assemblies of pET21_gene, unsuccessful enzyme digestion or ligation. SLP01 can be toxic to the host cell (as mentioned in section 1.3.1). If this was the case, and the pET21 T7 promoter was 'leaking' and SLP01 was produced, then few correct clones would be able to grow. It might have been that few cells were successfully transformed, indicating low efficiency of competent cells. However, this is contradictory to a high CFU count (>>500) for the positive control in a previous transformation using cells from the same batch. It is also possible that the cells have lost the pET21 plasmid (e.g. due to metabolic strain), they will then not have been able to grow on the ampicillin containing agar plates.

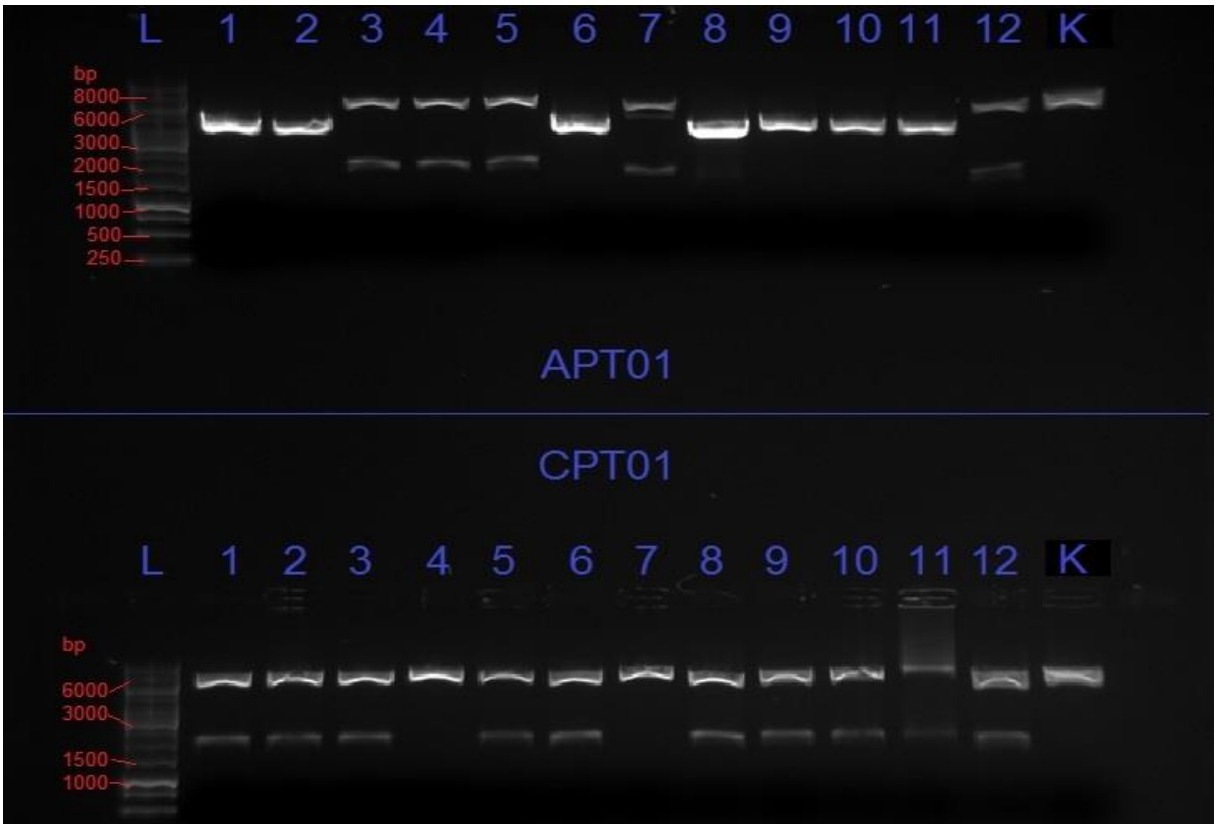

Figure 4.5: *Restriction mapping of 12 CFU of DH5a containing pET21_APT01 and pET21_CPT01 after digestion with* BglII*, each CFU marked with a number 1-12. The 1kb DNA Ruler from GeneScript™ was run in the lanes marked 'L'. 'K' represents the control plasmid, which was not cut prior to electrophoresis.*

In figure 4.5 two bands can be observed in five of the 12 samples for pET21_APT01 (3, 4, 5, 7 and 12) and in ten of the 12 samples for pET21_CPT01 (all except 4 and 7). Expected band sizes for correct insert of the candidate gene into the pET21 vector is either 1523 bp for APT01 or 1535 bp for CPT01, together with 5365 bp for the remaining plasmid. The upper of the two bands visible (for those samples with two visible bands), seem to be of roughly the correct size to represent the vector, while the lower bands seem to contain fragments of 2000-2500 bp size, compared to the DNA ruler. The reason for this has not been identified.

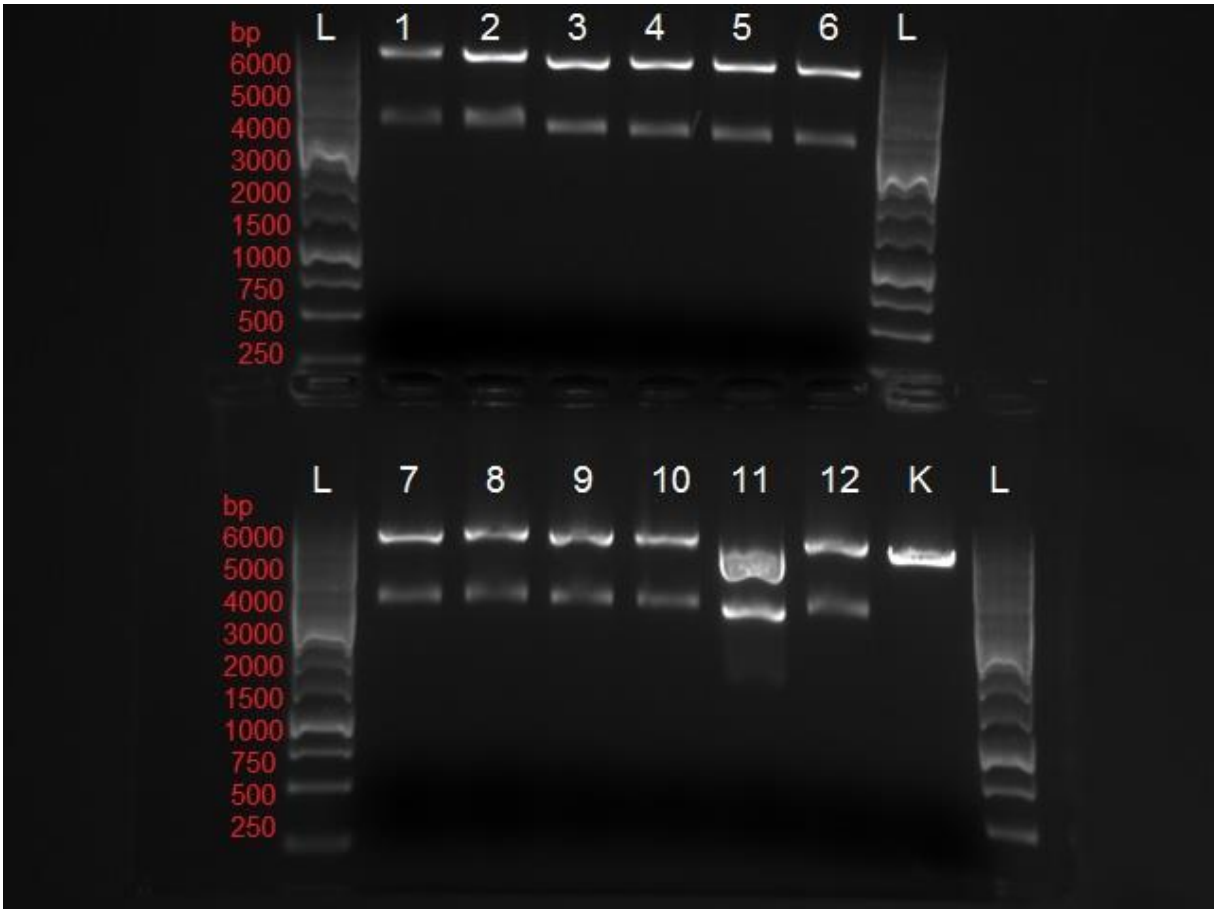

Figure 4.6: *Restriction mapping of 12 CFU of DH5a containing pET21_SLP01 after being cut with* NdeI *and* XhoI*, each CFU marked with a number 1-12. The 1kb DNA Ruler from GeneScript™ was run in the lanes marked 'L'. 'K' represents the control plasmid, which was not cut prior to electrophoresis.*

Correct insert of SLP01 into the pET21 plasmid was expected to give two fragments, one with 5365 bp and one with 2021 bp. For the control sample (K), two bands of 5365 and 266 bp were expected. Two bands are visible in figure 4.6 for all of the twelve samples for pET21_SLP01, both over 2000 bp in size. Only one band, over 5000 bp in size, is visible for the control sample. The size of fragments larger than 3000 bp are hard to estimate, as the bands from the DNA ruler are hard to distinguish. However, all of the bands visible in figure 4.6 seem to be of larger sizes than expected for both correct and incorrect vector construction. The reason for this is unclear.

As incorrect insertion of the gene would result in only one band, as seen for the control sample ('K'), it was assumed that those samples that gave two bands represented successful cloning of the candidate gene into pET21 (for both figure 4.5 and 4.6). Both figures showing aberrant run behaviour compared to the size

standard, for unknown reason. A second double digestion, using different restriction enzymes, could have been performed to verify correct plasmid construction.

4.2.3 Purification and measurement of concentration of the new vector constructs

When correct cloning of the candidate genes had been verified, a larger volume of plasmids preparation stock was prepared for recombinant protein production. The concentration of DNA obtained for each of the three candidates was measured using a Nanodrop^{TM} Spectrophotometer. The plasmid measured concentrations can be seen in table 4.3, below.

Table 4.3: *The concentration of DNA in the samples of plasmid purifications for pET21_APT01, pET21_CPT01 and pET21_SLP01.*

Plasmid	Concentration (ng/µl)
pET21_APT01	43.3
pET21_CPT01	18.1
pET21_SLP01	51.3

The obtained plasmid concentration was not high for any of the three candidates, and quite low for pET21_CPT01. This can indicate that the *E. coli* DH5α transformed with pET21_gene did not grow and replicate optimally, resulting in fewer cells, hence, fewer plasmids. Comparing with the results in table 4.2 strengthens this theory, as a much higher concentration of DNA was obtained from plasmid preparations from DH5α cells transformed with pUC57 plasmids than pOD1. As transformations using the low concentrated pOD1-isolate have proven successful by adjusting the amount used, and limitations in time for the project, it was decided to continue with the plasmid preparations made instead of preparing new ones to try to obtain a higher concentration.

4.3 Protein production

The *E. coli* strain BL21 codon plus RIPL was chosen for expression and production of the gene candidates. Cells of this strain were made competent and transformed with the pET21_gene plasmids. An overview of obtained CFUs after transformation can be seen in table 4.4.

Table 4.4: *Number of CFU formed on LA plates using 50 and/or 200 µl of culture from transformed* E. coli *BL21 cells with pET21_APT01, pET21_CPT01, pET21_SLP01, pOD1 (control), no plasmid (negative control) and pUC57_FragT1 (positive control).*

Plasmid	Culture (µl)	CFU
pET21_APT01	50	4
	200	27
pET21_CPT01	50	56
	200	329
pET21_SLP01	50	25
	200	149
pET21	50	2
	200	48
No plasmid	50	0
	200	0
pUC57_FragT1	200	45

Adjusting the amount of plasmid preparation used based on concentration of DNA seemed to work, as the transformed BL21 cells with pET21_CPT01 resulted in the most CFU, despite the low concentration of plasmid (table 4.3).

4.3.1 Protein production in 50 ml culture

Protein production and isolation was carried out as described in sections 3.2.6 and 3.2.7. Graphs showing the growth rates and measured absorbance for each of the cell cultures can be found in figure 4.7 and 4.8 (the raw data for the graphs can be found in appendix A.3).

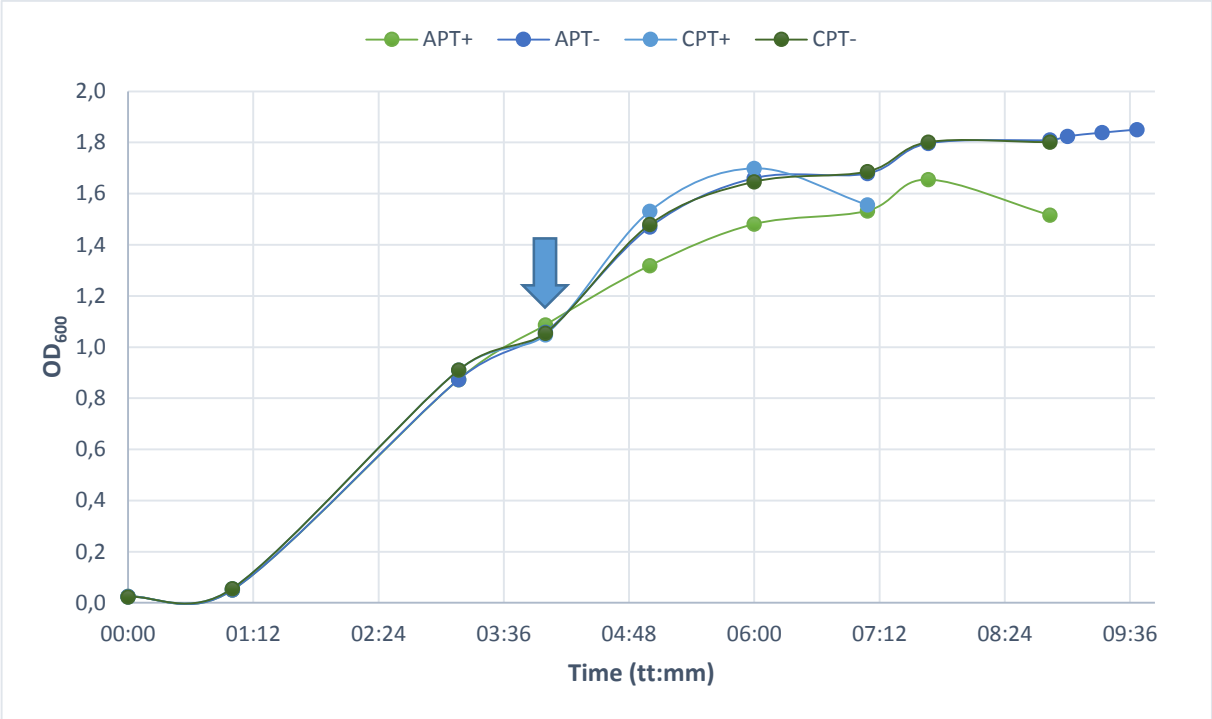

Figure 4.7: *Measured absorbance at OD_{600nm} over time for cultures of BL21 (DE3) cells transformed with pET21_APT01 or pET21_CPT01, incubated at 37 °C. The points represent time and value of absorbance measured, while the line between the point serve as a guide for the eye. Expression of target gene was induced by IPTG at time 4:00, when OD_{600nm} was close to 1, marked by the arrow.*

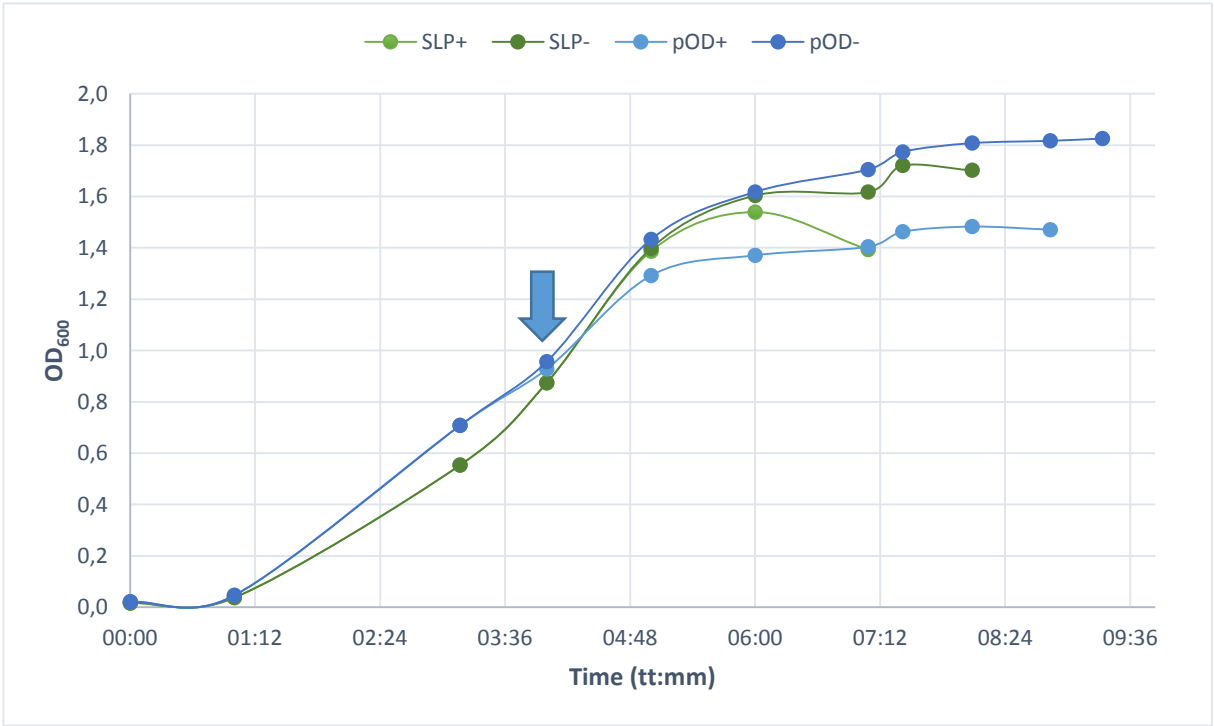

Figure 4.8: *Measured absorbance at OD_{600nm} over time for cultures of BL21 (DE3) cells transformed with pET21_SLP01 or pOD1, incubated at 37 °C. The points represent time and value of absorbance measured, while the line between the point serve as a guide for the eye. Expression of target gene was induced by IPTG at time 4:00, when OD_{600nm} was close to 1, marked by the arrow.*

The growth of the *E. coli* BL21 transformed cell cultures in figure 4.7 and 4.8 seem to stabilize at en OD_{600nm} value of 1.2-1.4. A decrease in absorbance can be observed for the induced cultures before their respective non-induced control cultures, for all candidates including the pOD1-control. This can indicate that high degree of expression of the target genes caused too much stress for the host cells, and they started to die within 3-5 hours after IPTG was added. The only exception being the non-induced BL21 cell culture transformed with pET21_APT01 or pOD1, which did not have a decrease in OD_{600nm} when the last measurement was performed 6 hours after induction.

Strain BL21 transformed with pOD1 was cultured and induced to function as a control. However, that was for the verification of target protein production (SDS-PAGE and activity assays). The induced activity of the T7 promoter in pOD1 caused over-expression of the SelW protein in the host cells.

Samples from pellet and supernatant of both induced and non-induced cultures for each of the candidates and the control were prepared and analysed by SDS-PAGE. The first SDS-PAGE gave results that were difficult to read, due to too much protein in the pellet samples and not enough in the supernatant samples (can be seen in appendix A2). To correct for this, TCA protein precipitation was performed on the supernatant samples (section 3.2.8), and a lower volume of the pellet samples was applied on the gel. The following results can be seen in figure 4.7 and 4.8.

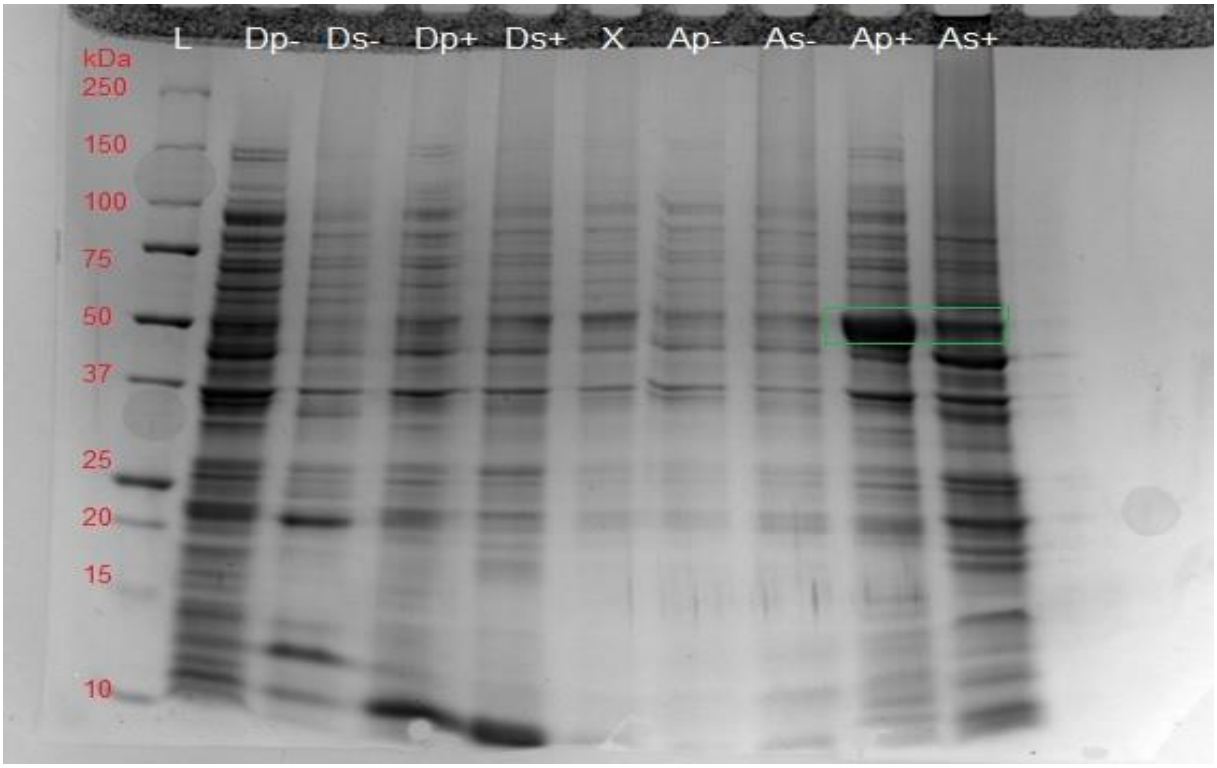

Figure 4.9: *SDS-PAGE of pellet (p) and supernatant (s) of induced (+) and non-induced (-) BL21 DE3 codon plus cells transformed with control plasmid pOD1 (D), or pET21 plasmid carrying the gene for APT01 (A). Biomarker from BioRad was used (L), with sizes in kilo Dalton shown in red. The green box marks bands that are believed to represent the APT01 protein. "X" represents a well were no sample was added.*

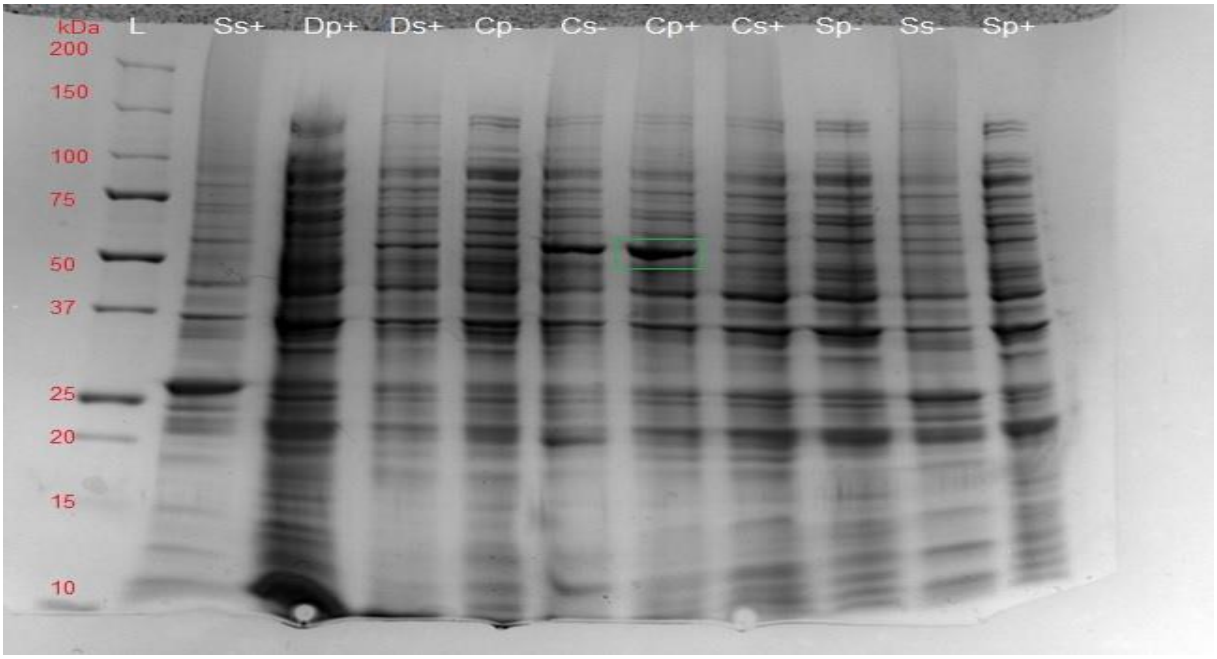

Figure 4.10: *SDS-PAGE of pellet (p) and supernatant (s) of induced (+) and non-induced (-) BL21 DE3 codon plus cells transformed with control plasmid pOD1 (D), or pET21 plasmid carrying the gene for CPT01 (C) or SLP01 (S). Biomarker from BioRad was used (L), with sizes in kilo Dalton shown in red. The green box marks a band that might represent the CPT01 protein.*

The sizes for the three peptidases are roughly 53, 57 and 73 kDa for APT01, CLP01 and SLP01, respectively. The pellet sample from the induced pET21_APT01 transformed cells seem to have a clearly broader band at about 53 kDa in figure 4.9 (left in the green box), compared to the pellet sample of the non-induced APT01 cells and the control. The band of the same respective size is also more distinct for the supernatant sample of the same culture (right in the green box in figure 4.9). This indicates that APT01 is produced in BL21 DE3 cells following IPTG induction, although most of it remains in an insoluble form. A more distinct band of a size close to 57 kDa can be observed for the CPT01 candidate in the pellet sample of the induced cells (marked with a green box in figure 4.10), compared to the non-induced and control samples. This indicates some production of insoluble CPT01 protein. However, a slightly broader band of the same size can also be observed for the non-induced supernatant sample for CPT01. Since this sample originates from a non-induced culture, no CPT01 production should have taken place. As no accumulation of CPT01 is visible for the induced supernatant sample, it is more likely that the ‘Cs-‘ well was contaminated with some of the induced pellet (Cp+) sample, rather than some production of CPT01 did occur without induction. No clear difference can be seen for the induced and non-induced SLP01 samples, indicating that the protein is not successfully produced.

4.3.2 Incubation temperature at 16 °C

One possible reason non-successful production of the three peptidases in soluble form can be that production at 37 °C did not allow the proteins to fold correctly relative to their production by the synthesis machinery. To test for this theory the cultures were incubated at 16 °C after IPTG was added, while all other production factors remained the same as the cultivation at 37 °C. As there was no visible production of SLP01 using the BL21 strain, the pET21_SPL01 plasmid was transformed into a different strain of *E. coli*, the ER2566 strain, which also has a T7 promoter system induced by IPTG. Empty vectors (pOD1 not carrying any of the three gene candidates) were also used to transform ER2566 cells to serve as control. Graphs showing growth rates can be seen in figure 4.11-4.16 (raw data can be found in appendix A.3).

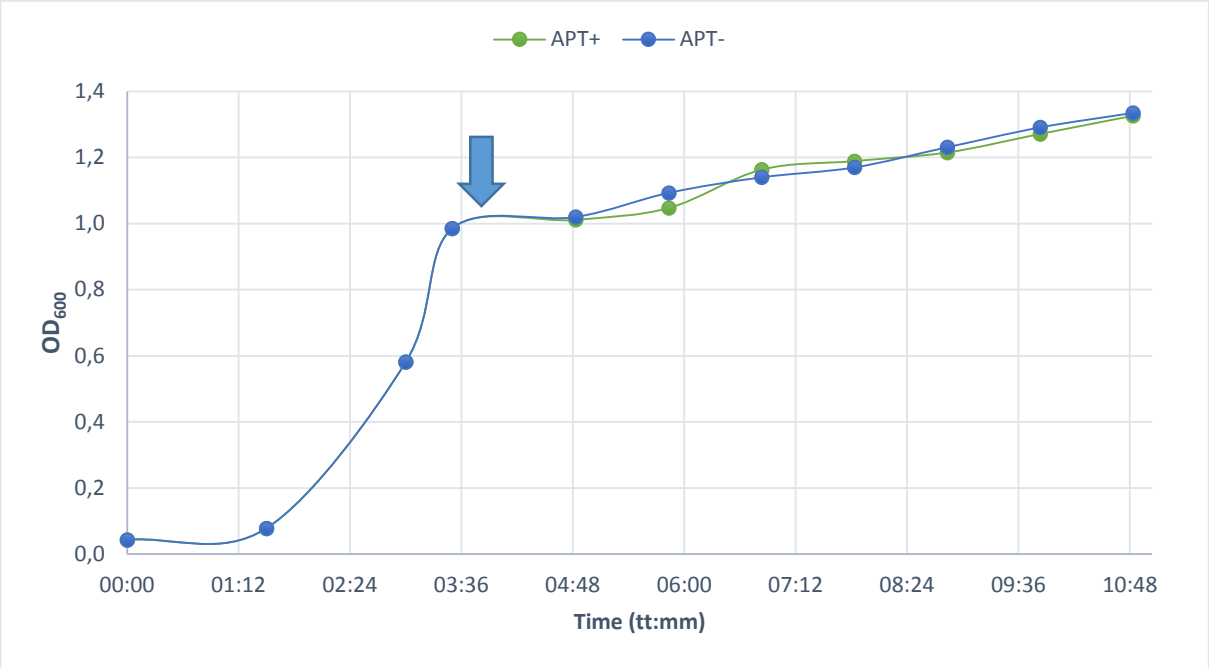

Figure 4.11: *Measured absorbance at OD_{600nm} over time for cultures of BL21 cells transformed with pET21_APT01, incubated at 37 °C. After induction with IPTG at 03:50, marked by the arrow, both the induced (APT+) and non-induced (APT-) culture was incubated at 16 °C. The points represent time and value of absorbance measured, while the line between the point serve as a guide for the eye.*

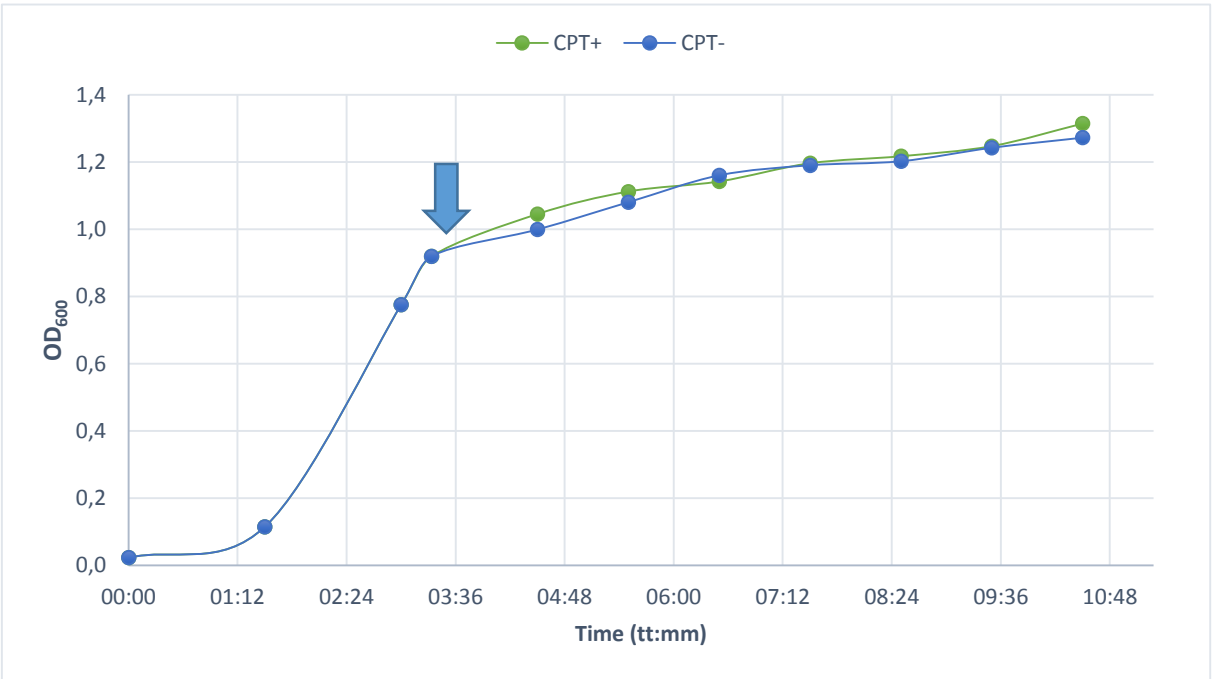

Figure 4.12: *Measured absorbance at OD_{600nm} over time for cultures of BL21 cells transformed with pET21_CPT01, incubated at 37 °C. After induction with IPTG at 03:30, marked by the arrow, both the induced (CPT+) and non-induced (CPT-) culture was incubated at 16 °C. The points represent time and value of absorbance measured, while the line between the point serve as a guide for the eye.*

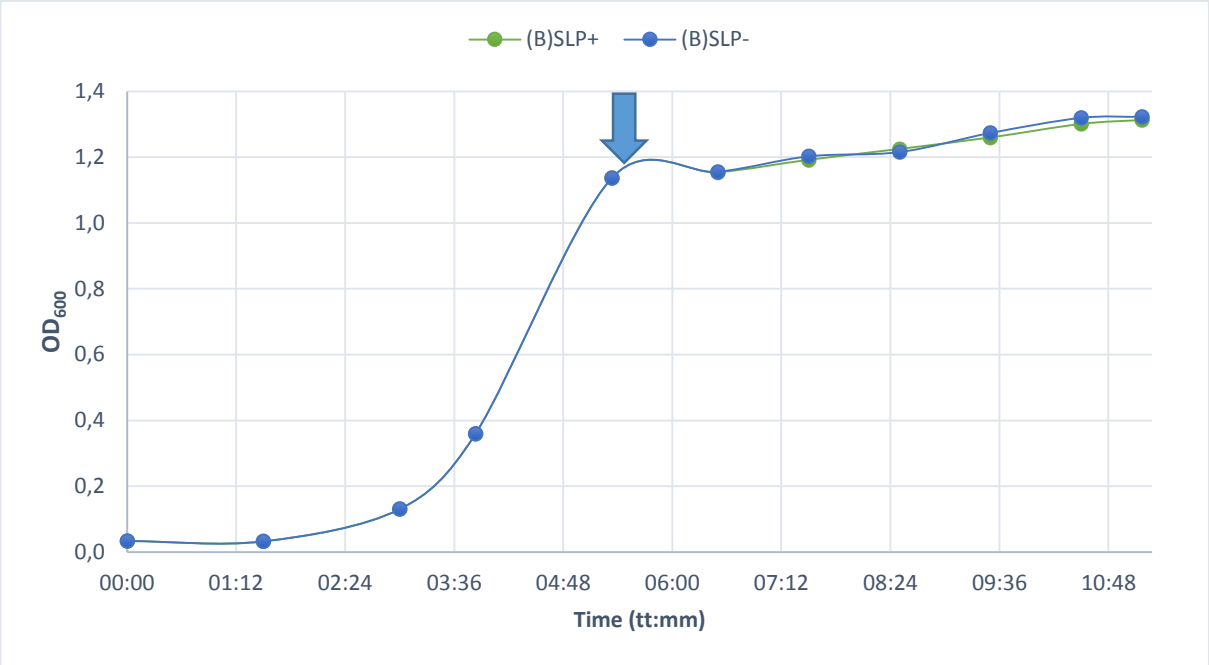

Figure 4.13: *Measured absorbance at OD_{600nm} over time for cultures of BL21 cells (B) transformed with pET21_SLP01, incubated at 37 °C. After induction with IPTG at 05:30, marked by the arrow, both the induced (SLP+) and non-induced (SLP-) culture was incubated at 16 °C. The points represent time and value of absorbance measured, while the line between the point serve as a guide for the eye.*

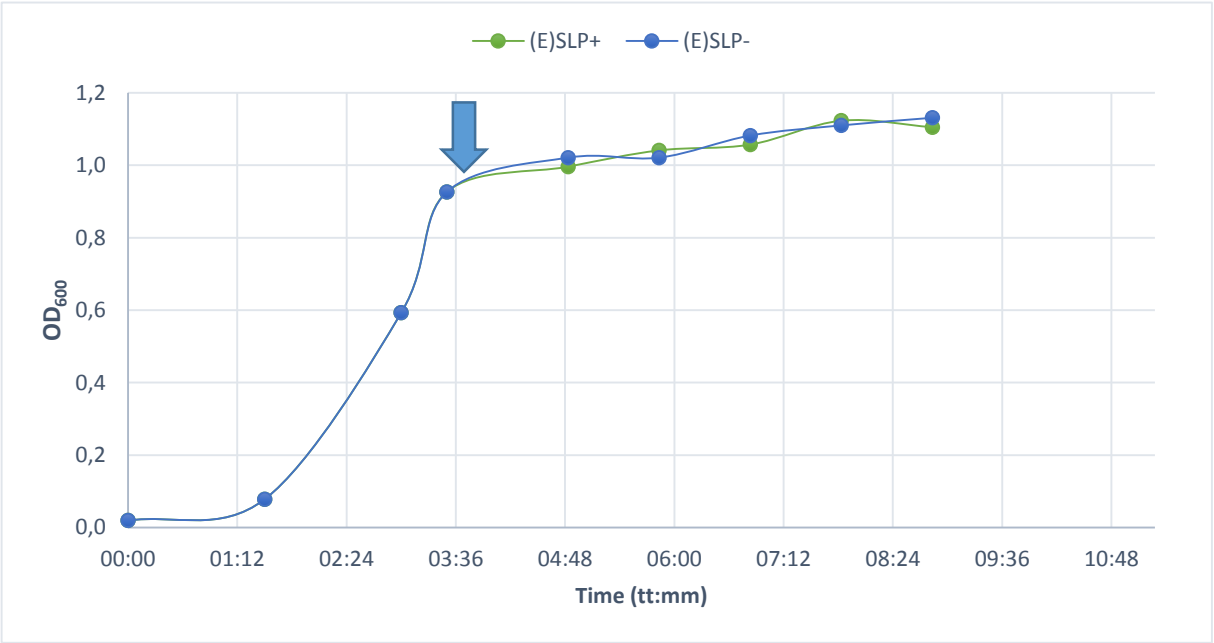

Figure 4.15: *Measured absorbance at OD_{600nm} over time for cultures of ER2566 cells (E) transformed with pET21_SLP01, incubated at 37 °C. After induction with IPTG at 03:50, marked by the arrow., marked by the arrow, both the induced (SLP+) and non-induced (SLP-) culture was incubated at 16 °C. The points represent time and value of absorbance measured, while the line between the point serve as a guide for the eye.*

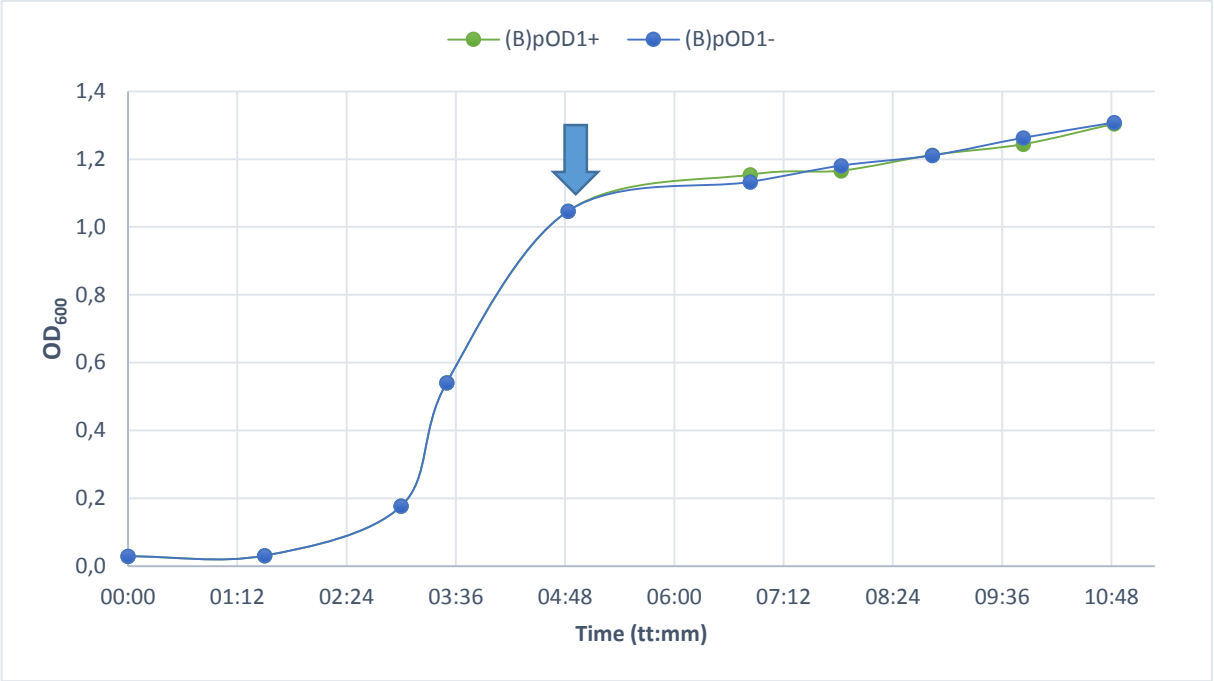

Figure 4.14: *Measured absorbance at OD_{600nm} over time for cultures of BL21 cells (B) transformed with pOD1, incubated at 37 °C. After induction with IPTG at 04:50, marked by the arrow, both the induced (pOD1+) and non-induced (pOD1-) culture was incubated at 16 °C. The points represent time and value of absorbance measured, while the line between the point serve as a guide for the eye.*

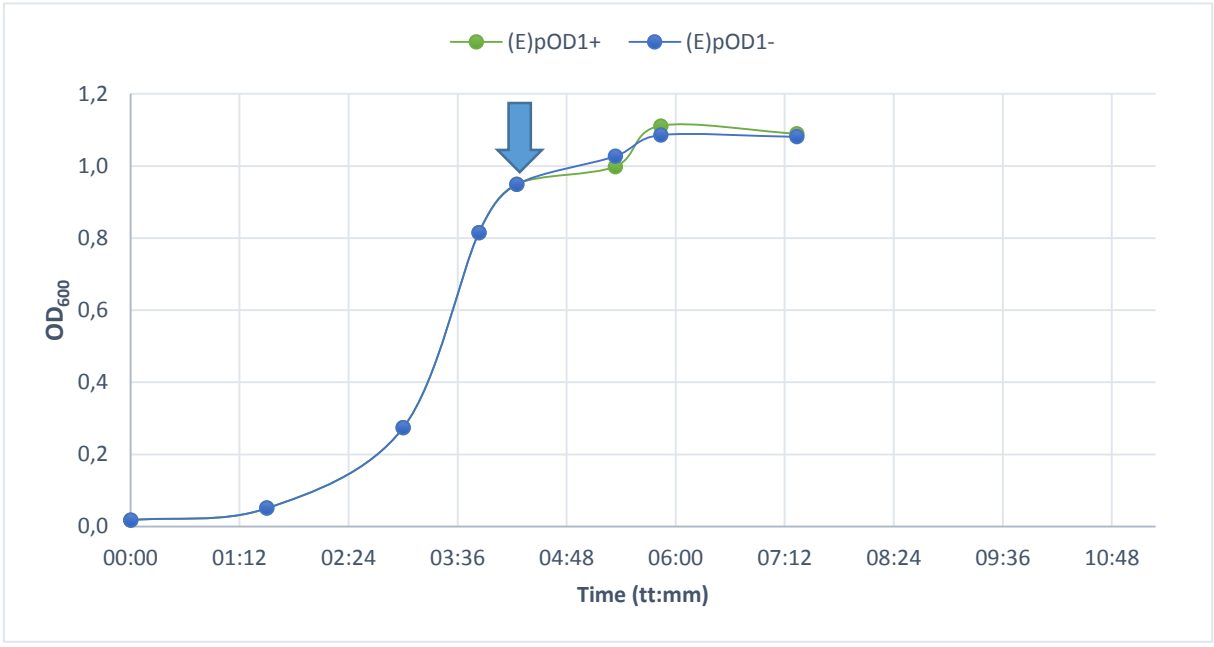

Figure 4.16: *Measured absorbance at OD_{600nm} over time for cultures of ER2566 cells (E) transformed with pOD1, incubated at 37 °C. After induction with IPTG at 04:20, marked by the arrow, the culture was incubated at 16 °C. The points represent time and value of absorbance measured, while the line between the point serve as a guide for the eye.*

A decrease in growth rate can be seen for all cultures in figures 4.11-4.16 after induction by IPTG. Since 37 °C is the optimal growth temperature for *E. coli*, this was an expected result when the incubation temperature was changed from 37 to 16 °C. A decrease in OD_{600nm}-values was only observed for the ER2655 cultures (figure 4.15 and 4.16). For the ER2566 culture transformed with pOD1, the absorbance started to decrease already 3 hours after induction, also for the non-induced control. ER2566 has successfully been used to produce recombinant protein using T7 expression system and the pET21 vector in the past (Wu, 2000). Based on the data obtained in this experiment, it is not a successful host strain for the production of SLP01 (or SelW) under the conditions used. Changing one or more production parameters, such as temperature, medium used or concentration of IPTG added, might give improved results.

No decrease in absorbance was observed for any of the BL21 strain cultures within the 5-7 hours where absorbance was measured after induction. Although the value seemed to stabilize at OD_{600} values of ~1.3 for the SLP01. Incubation after induction should have continued until the absorbance values stabilized. The reason for this not been done was the desire to obtain frequent absorbance measurements of the cultures, to avoid death of cells giving a decrease in OD_{600nm} and a lower yield of target protein. However, O/N incubation at 16 °C after induction is not uncommon for recombinant protein production using the T7 promoter system, and considering the low yield of soluble target proteins obtained, it would have been a natural next step for this study.

To verify production of target protein, samples from pellet and supernatant from both induced and non-induced cultures of the three candidates and the pOD1 control were analysed by SDS-PAGE, results shown in figure 4.17 and 4.18 below.

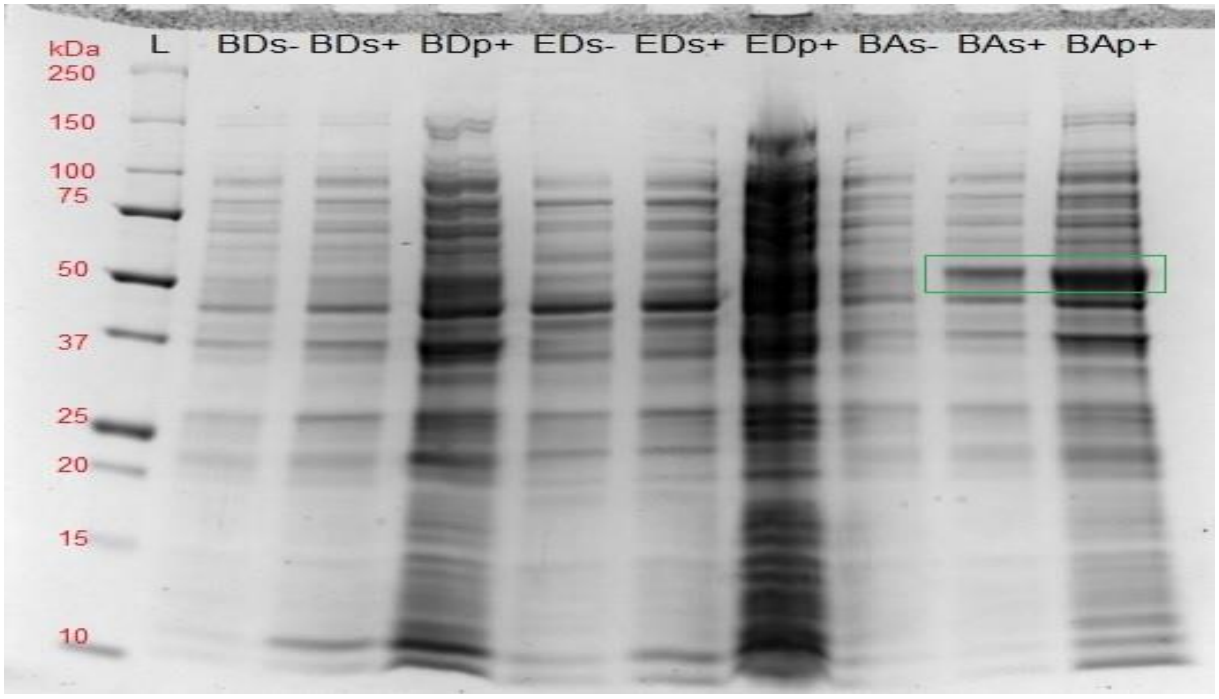

Figure 4.17: *SDS-PAGE of pellet (p) and supernatant (s) of induced (+) and non-induced (-) BL21 DE3 codon plus (B) and ER2566 (E) cells transformed with control plasmid pOD1 (D), or BL21 DE3 cells transformed with pET21_APT01(A). Incubation temperature after IPTG induction was 16 °C. Biomarker from BioRad was used (L), with sizes in kDa shown in red. The green box marks a band that assumingly represent the APT01 protein.*

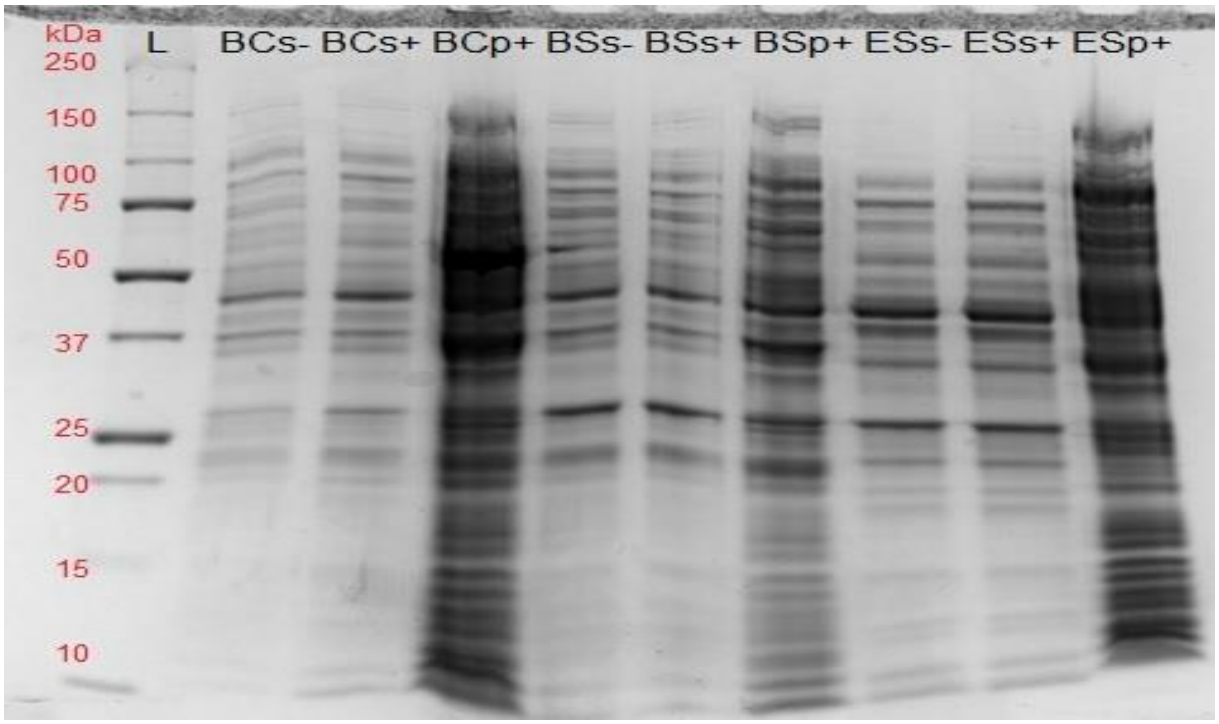

Figure 4.18: *SDS-PAGE of pellet (p) and supernatant (s) of induced (+) and non-induced (-) BL21 DE3 codon plus (B) and ER2566 (E) cells transformed with control plasmid pOD1 (D) or pET21_SLP01, or BL21 DE3 cells transformed with pET21_CPT01 (C) or pET21_SLP01 (S). Incubation temperature after IPTG induction was 16 °C. Biomarker from BioRad was used (L), with sizes in kDa shown in red.*

The analysis of pellet samples in figure 4.17 and 4.18 are difficult to interpret due to too much protein present. To get clearer results, new samples were prepared from pellet and supernatant from the induced cultures, where the centrifugation time was increased to remove contamination, and a smaller volume of the pellet samples was applied on the gel. The results from this gel can be seen in figure 4.19.

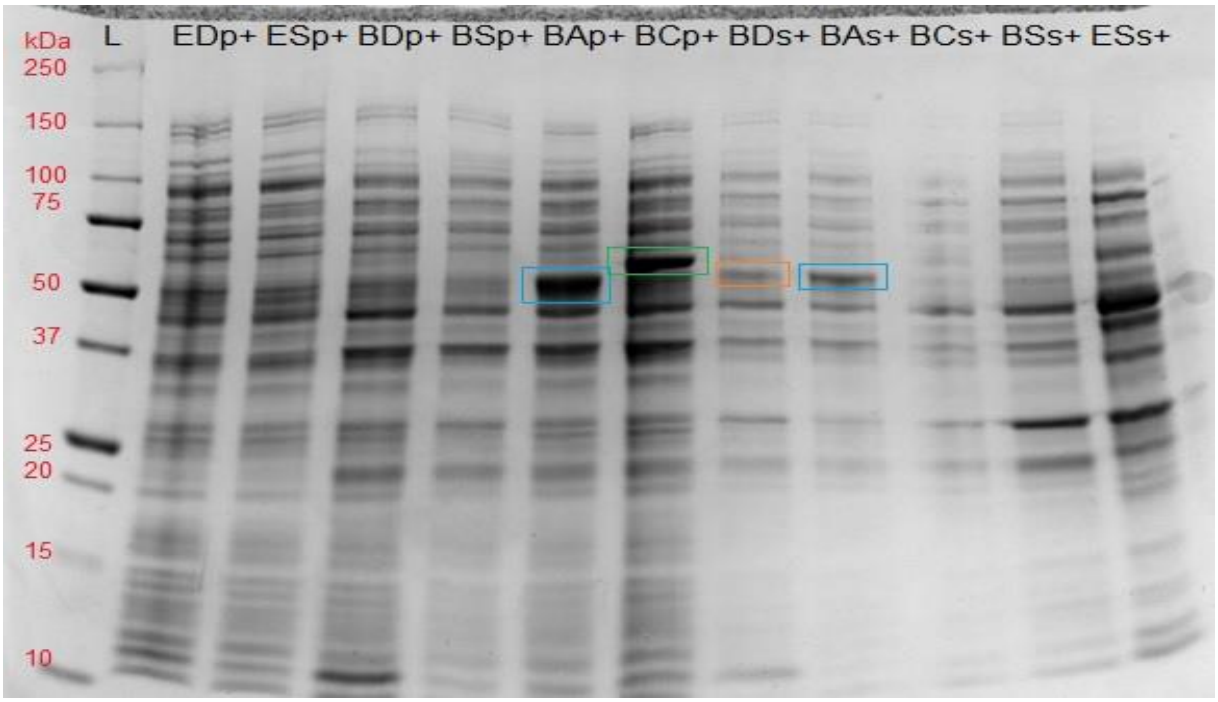

Figure 4.19: *SDS-PAGE of pellet (p) and supernatant (s) of induced (+)BL21 DE3 codon plus (B) and ER2566 (E) cells transformed with control plasmid pOD1 (D) or pET21_SLP01, or BL21 DE3 cells transformed with pET21_APT01 (A), pET21_CPT01 (C) or pET21_SLP01 (S). Biomarker from BioRad was used (L), with sizes in kilo Dalton shown in red. The blue and green boxes marks a band that assumingly represent the APT01 and CPT01 protein, respectively. The orange box marks a band equal to that of ATP01 indicating contamination between wells.*

Again, it can be observed that pET21_APT01 transformed cells produced the APT01 protein after induction with IPTG, and that most of the protein was present as an insoluble form in the pellet (marked with a green box in figure 4.17 and blue boxes in figure 4.19). A distinct band from a probable CPT01 production can also be observed at about 57 kDa for the induced pellet sample of pET21_CPT01 carrying cell culture (marked with a green box in figure 4.19), as in the 37 °C incubation results from figure 4.10. But this time, no broader band can be seen in figure 4.19 for the non-induced supernatant sample of CPT01, reinforcing the theory that the respective sample in figure 4.8 was contaminated. A band matching the size of that, assumingly, produced by the APT01 protein, can be observed in the supernatant sample of the induced

control (pOD1), marked with an orange box in figure 4.19. The control plasmid pOD1 does not carry the APT01 gene, but *selW,* encoding a selenoprotein of about 9.8 kDa size. Therefore, the presence of a distinct band of this size in the control sample cannot be used for the verification of a potential APT01 production. However, this band is not visible for the same sample in figure 4.9, which suggests that this sample in figure 4.19 was contaminated from the neighbouring well containing the APT01 protein. No bands indicating a production of the SLP01 protein can be observed, for either the *E. coli* BL21 (DE3) or the ER2566 cell cultures in figures 4.18 and 4.19.

4.3.3 Heat denaturation series

One of the assumed features of the chosen protein candidates were that, if they were successfully produced, they would be stable at high temperatures due to adaptations to the environment in the deep subsurface oil reservoirs. The APT01 peptidase seemed to be the only candidate that was produced in soluble form (figure 4.17 and 4.19), and was therefore the only candidate tested for heat stability. Two heat denaturation series were performed on samples from the supernatant of induced pET21_APT01 cultures (section 4.3.2). Samples were kept on heat blocks with different temperatures for different duration. Figure 4.13 shows the SDS-PAGE results of the heat denaturation.

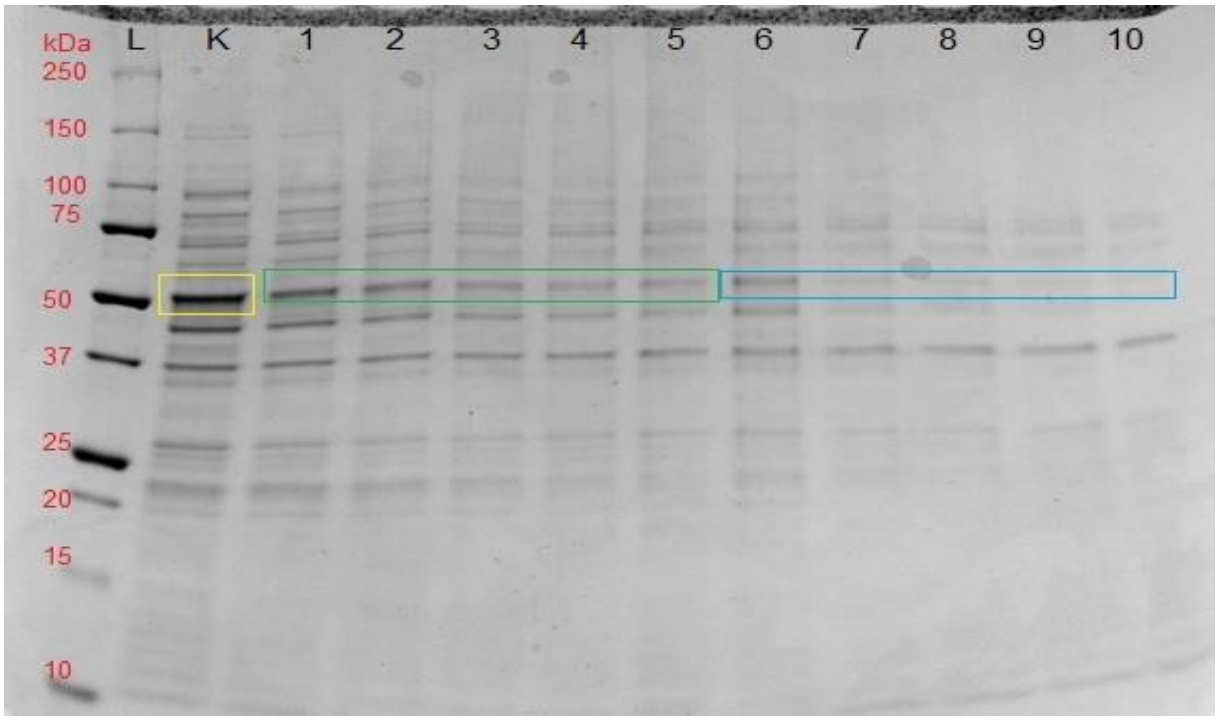

Figure 4.20: *SDS-PAGE of 11 samples from supernatant of cell cultures of* E. coli *BL21 cells transformed with pET21_APT01, induced with IPTG and incubated at 16 °C. Sample 1-5, where ATP01 is marked with a green box, was subjected to 65 °C for 5, 15, 30, 60 and 120 minutes, respectively. Sample 6-10, ATP01 marked with a blue box, was subjected to 80 °C in the same order. 'K' represents a sample of supernatant not subjected to any heat prior to SDS preparation (APT01 marked with a yellow box). Biomarker from BioRad was used (L), with sizes in kDa shown in red.*

Several bands can be observed for all samples subjected to 65 °C, while the amount of visible bands observed for samples 7-10 drastically diminishes. Only one clearly distinct band is present for samples subjected to 80 °C, and this seem to represent a protein with the size of ~40 kDa. This band might be produced by one of the

components in the DnaK chaperone system, consisting of the heat shock proteins DnaK, DnaJ and GrpE (Schönfeld, 1995).

A band assumingly representing presence of ATP01 is visible in all of the samples subjected to 65 °C (sample 1-5), marked with a green box in figure 4.13. The band seems to become more diffuse as the duration of heat exposure increases, indicating that the amount of the APT01 protein denatured was increased when subjected to 65 °C for a longer time. However, not much difference can be seen for the APT01-band of sample 3-5, suggesting that there were some proteins present that were stable at 65 °C. For the 80 °C heat series, hardly any APT01 can be observed in the blue box for sample 7-10 in figure 4.13. From these results, it would seem like some the soluble APT01 produced by the *E.* coli BL21 cells are stable at 65 °C, but not at 80 °C.

Plausible reasons for the obtained target protein of APT01 not being heat stable, can be incorrect post-translational modifications of the enzyme, lack of necessary co-factors or heat-stability chaperones. The structures of heat stable proteins do not necessarily differ much from those of less thermostable proteins (Kumar, 2000). Any factor influencing on the structure of a protein, such as pH, co-factors and pressure (discussed further below), can therefore affect the thermostability.

Heat denaturation can be used to purify thermostable target enzymes produced by recombinant gene technology, as most of the other proteins produced by mesophiles, such as *E. coli,* will denature at temperatures above 40 °C. However, as can be seen in figure 4.20, there are still several bands representing other proteins present in the sample after heat denaturation at 65 °C, and purification using 80 °C will not be successful since the target enzyme does not seem to be stable at this temperature. A possible next step for this study could be to test for heat stability for temperatures between 65-80 °C, and to what degree that purifies the samples.

4.3.4 Protein production in 1L cultures

There are several ways of increasing the probability for production of soluble proteins from recombinant expression. Lowering the incubation temperature did not improve the yield of the target proteins (section 4.3.2). A different theory for unsuccessful recombinant protein production is that the protein of interest is produced at such a low concentration in the cell that it makes it difficult to detect presence and activity without purification of the protein. As the His-tag modification was unsuccessful and the protein obtained was not heat stable, purification of the proteins of interest would require alternative approaches, which there were not time for at this stage. Instead, an up-scaling of the cultivation was performed as a last attempt to produce the peptidase candidates in soluble form.

The overnight inoculum was added to 1 litre of medium, and IPTG was added for induction when absorbance measures equalled ~3. Due to a limited number of Erlenmeyer flasks large enough for 1 litre cultures, the cultures were not divided to give a non-induced control. The cells transformed with pOD1 was therefore the only control sample for comparisons, for this round of protein production. As there had been no visible production of SLP01 previously, this candidate was not included in the up-scaled cultivation (also due to limited equipment). After induction, the cultures were incubated at 16 °C for five hours. As for the previous rounds, a frequent measure of absorbance was desired, so the time for incubation was determined by previous cultivation results.

As an attempt to time the experiment so that the three cultures reach an OD value of about 3, and hence could be induced at the same time, the pOD1-culture was inoculated an hour earlier than the two others (APT01 and CPT01). This decision was based on the results from earlier in this study showing that the BL21 cells transformed with pOD1 had a slower growth compared to those transformed with pET21_APT01 or pEt21_CPT01 (figure 4.12-4.14 and appendix A.3). However, for this experiment the pOD1-containing cells grew rapidly reaching OD of 3 much earlier than the other two. The flask was therefore taken out of the shaking incubator at time 05:00, and kept at room temperature (without shaking) until the other two cultures reached the desired cell density. Induction was performed at the same time for all three cultures (at time 04:45 for APT and CPT, and 05:45 for pOD). After IPTG was added, all three cultures

where incubated at 16 °C in a shaking incubator. Measured OD values are displayed as graphs in figures 4.21, 4.22 and 4.23.

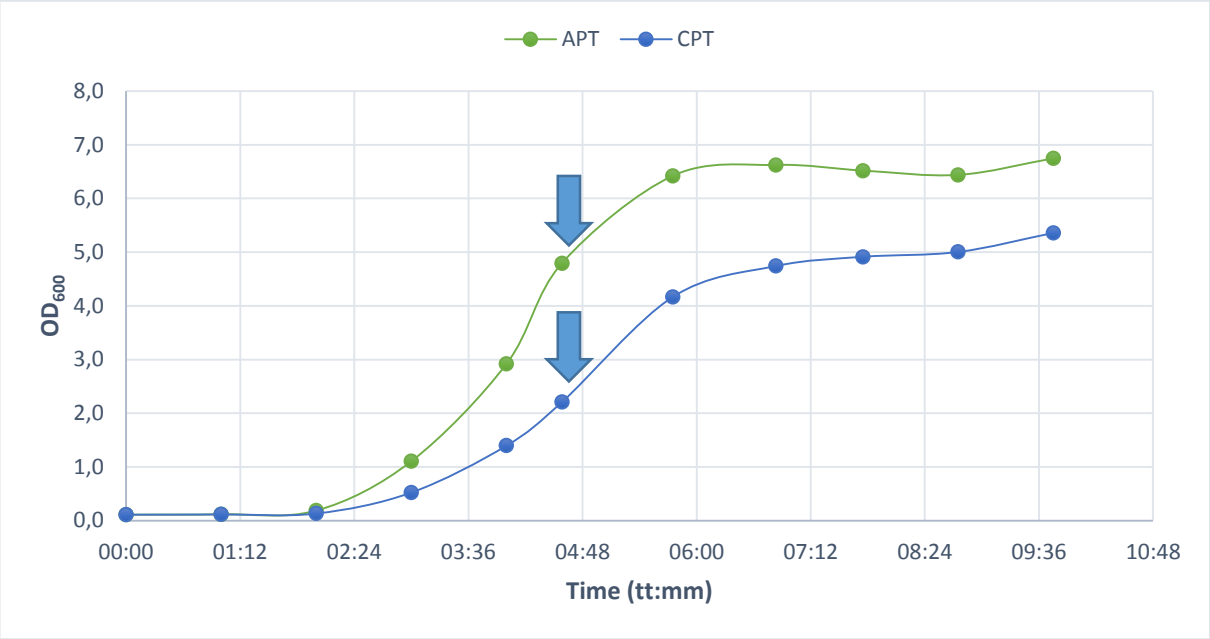

Figure 4.21: *Measured absorbance at OD_{600nm} over time for cultures of BL21 cells transformed with pET21_APT01 (APT) or pET21_CPT01 (CPT), incubated at 37 °C. After induction with IPTG at 04:45, marked by the arrow, the cultures were incubated at 16 °C. The points represent time and value of absorbance measured, while the line between the point serve as a guide for the eye.*

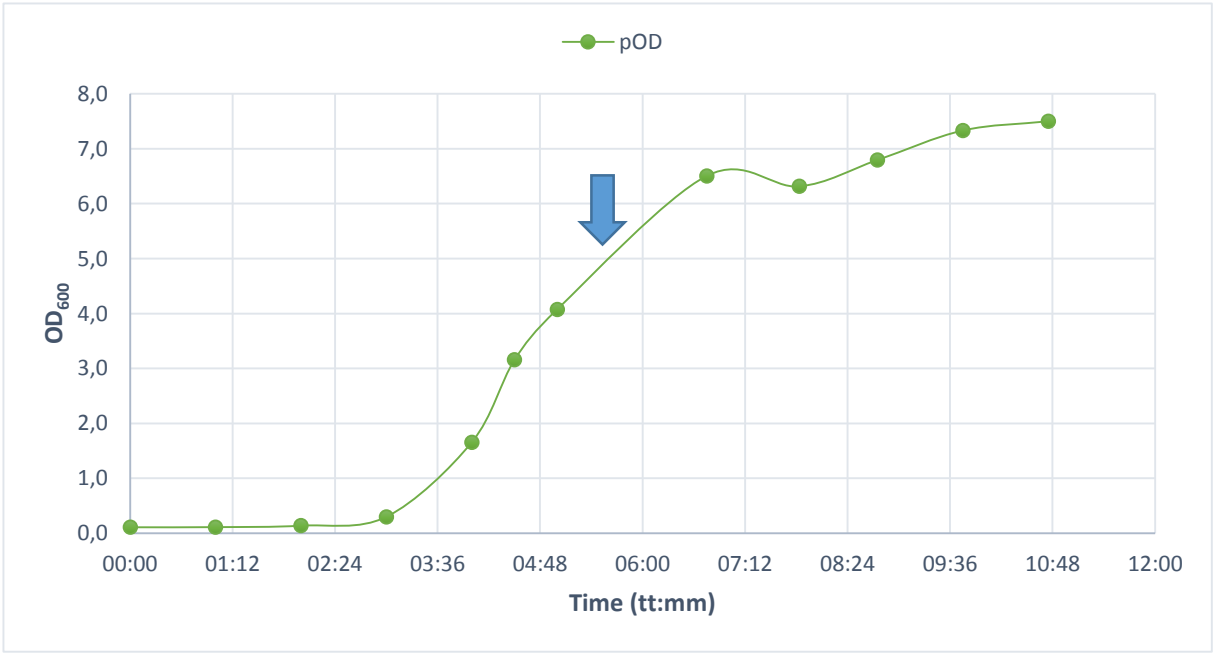

Figure 4.22: *Measured absorbance at OD_{600nm} over time for cultures of BL21 cells transformed with pOD1 (pOD), incubated at 37 °C. After induction with IPTG at 05:45, marked by the arrow, the culture was incubated at 16 °C. The points represent time and value of absorbance measured, while the line between the point serve as a guide for the eye.*

The growth rate seemed to decrease, as expected (section 4.3.2), for all three parallels after the incubation temperature was changed from 37 to 16 °C. A steady, continuous increase in the absorbance values can be observed for the CPT01-culture from time 06:50-09:45, while the cells transformed with pET21_APT01 seem to stabilize at an OD-value of about 6.5, before increasing at time 09:45.

As previous, both samples from the pellet and supernatant of the lysed cells were analysed using SDS-PAGE. The results from the separation of the proteins produced can be seen in figure 4.23.

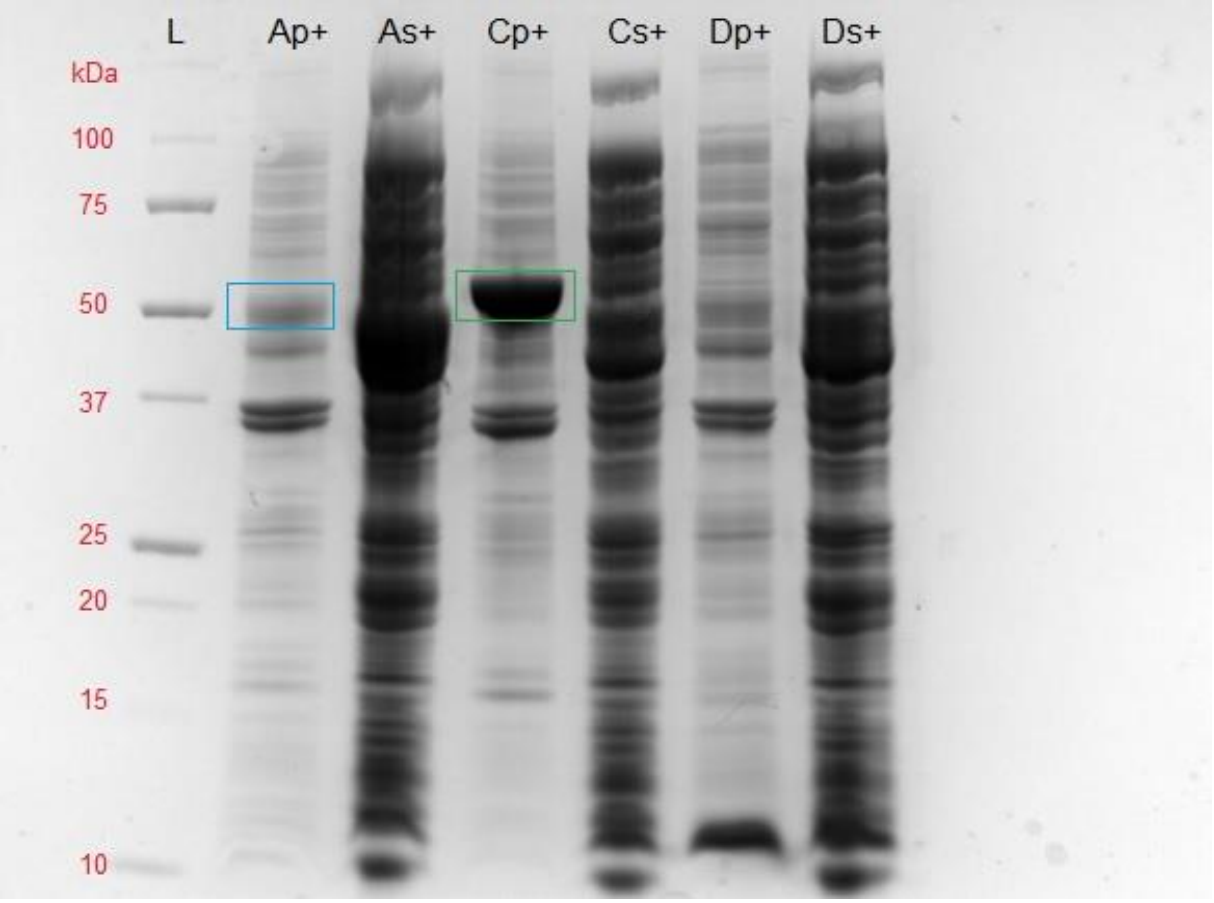

Figure 4.23: *SDS-PAGE of pellet (p) and supernatant (s) of induced (+) BL21 DE3 codon plus cells transformed with control plasmid pOD1 (D), pET21_APT01 (A), or pET21_CPT01 (C). Incubation temperature after IPTG induction was 16 °C. Biomarker from BioRad was used (L), with sizes in kilo Dalton shown in red. The blue and green box marks the band assumingly representing the APT01 and CPT01 protein, respectively.*

After this round of recombinant protein production, the amount produced of the protein of interest in insoluble form seemed to be much more for CPT01 compared to APT01, indicated by the width and intensity of the bands marked with the green and blue box, respectively. The supernatant samples from the upscaled cultivation contained a lot of protein and resulted in many bands on SDS-PAGE gel, causing the bands to become undefined and difficult to interpret. Figure 4.23 shows an image of the gel where the supernatant samples were centrifuged for a prolonged time and a low volume were applied on the gel. Still, it is hard to distinguish between the different bands for the supernatant samples. However, there does not seem to be a band for APT01 or CPT01 from the supernatant samples, indicating that what was produced of these proteins remained insoluble.

Compared to previous results (figure 4.9, 4.17 and 4.19), the amount of APT01 relative to the other proteins in the same sample appear to be much less. Reasons for this can be that conditions for 100 ml cultivation is favoured over that of 1 L, for this candidate. As mentioned above, three times the concentration of tryptone and yeast extract was used in the 1 L culture medium, to compensate for demand for a higher OD value. During incubation the intensity of the shaking was adjusted to 180 rpm for the 1 L cultures, compared to 225 rpm for the 50/100 ml cultures. This, along with the size of the Erlenmeyer flasks and how much of the total volume of the flask is filled with medium, affects the formation of foam and accessibility of oxygen to the cells. Apparently, these conditions were favoured for production of CPT01 in 1 L cultures, and APT01 in 50/100 ml cultures. However, the 1 L conditions did not seem to favour soluble production of these proteins, for either of the two candidates.

Due to the continuous increase in growth rate, a longer incubation period could have been used to obtain a higher yield of protein, and perhaps also in soluble form as necessary post-translational modifications of the enzymes might have needed more time to be performed.

4.3.5 Activity assay using skim milk agar

Single CFUs from transformed BL21 cells were streaked on milk agar plates containing ampicillin and IPTG to test if active peptidases were secreted from the cells. Proteinase K was also applied on the plate, as a positive control. No activity, except for the positive control, could be observed on these plates after 24 and 48 hours incubation at 37 °C, 45 °C or 60 °C.

All induced supernatant samples from the protein production using 50 ml and 1 litre cultures and 16 °C incubation temperature (after induction), were tested on MA plates, along with the induced control vector (pOD1) as a negative control and the positive control proteinase K. After both 24 hours and 48 hours incubation at 37 °C, only the positive control produced a clearing on the plate, as seen in figure 4.24. This suggests that no active peptidase was produced in soluble form for the cells transformed with pET21_APT01, pET21_CPT01 or pET21_SLP01.

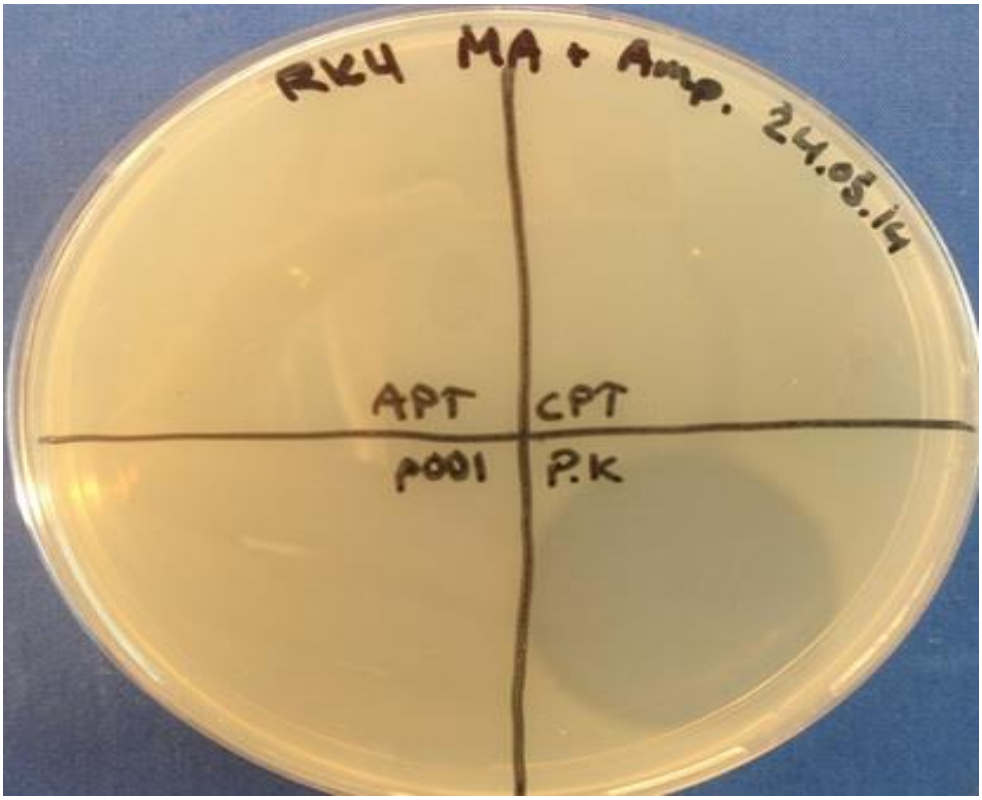

Figure 4.24: *MA plate with drops of induced supernatant samples from the protein production attempt that used cultures of 1 litre and incubation temperature was set to 16 °C. 'APT',' CPT' and 'pOD1' represents the quadrant where sample was added from the cell cultures transformed with pET21_APT01, pET21_CPT01 and pOD1, respectively. 'P.K' marks the quadrant where the positive control was added.*

To observe activity on the MA plates was not expected for the samples for CPT01 or SLP01, as no production of the proteins had been identified in the supernatant samples from the induced cultures. The hopeful candidate for this activity assay was the supernatant sample containing soluble APT01, seen in figure 4.17 and 4.19. However, no 'halo' was made by the APT01 from this sample. Possible explanations for this are many, as it is not known whether this enzyme requires co-factors or chaperones not present in *E. coli* in order to function as intended. The activity of this enzyme may also be dependent on high temperature and/or pressure.

The *in situ* temperatures of the two oil wells sampled to make the metagenomic DNA library used for this project, are 86 and 83 °C for well I and II, respectively (Lewin, 2013). In other words, the putative enzymes identified and produced have adapted to this temperature, and the optimum temperature for activity is probably around 80 °C. A possible solution for the negative result on the MA plates can be that the activity at temperatures far below the enzymes optimum temperature is so low that this activity assay in not sensitive enough to detect the hydrolysis performed by the target proteins.

Reasons why the production of the target enzymes in a soluble form was not successful are numerous. The most probable reason is that, as a host organism expressing the target gene in a completely different environment than that of a deep subseafloor oil reservoir, *E. coli* and the surrounding environment does not contain the correct factors for successful production of the recombinant protein. Although the type of organisms who's DNA have been expressed successfully in *E. coli* is diverse, heterologous expression is still one of the main challenges within metagenomic research (Handelsman, 2004, Streit, 2004). The presence of components needed for post-translational modification of recombinant proteins are difficult to predict, and even more complex than the expression mechanisms. A study done by Gabor and colleagues (2004) showed that of the genes from 32 prokaryotic genomes, only 40 % could be detected when expressed in *E. coli*. *E. coli* is commonly used for heterologous protein production because of its strengths as a well-studied, cheap and easy-to-grow organism, however development of novel experimental systems for expression of metagenomic genes are needed.

The DNA sequences selected as gene candidates in this study are putative genes. Although, measures were made to increase the chance of identifying a functional gene, that might not have been successful. The sequences identified can represent so-called pseudogenes, sequences that have been a functional gene once, but are no longer expressed in the cell, often due to accumulation of mutations. Most pseudogenes have gene-like features, such as promoters, and can be perceived as a functional gene when just looking at the DNA sequence.

It has been shown that high pressure has a significant effect on both stability and activity of various proteins, including several peptidases (Eisenmenger, 2009). For instance, the activity of the serine peptidase α-Chymotrypsin was 30 times higher at 50 °C and 360 MPa, compared to at 20 °C and 0.1 MPa. While the stability of the same enzyme was increased, from less than five minutes half-life to being active for 30 minutes, when the pressure was increased from 0.1-180 MPa at 52.5 °C (Mozhaev, 1996). Also, the reaction rate of a thermostable peptidase, thermolysin (from *Bacillus thermoproteolyticus*), increased by more than 13 times when subjected to pressure of 100 MPa (Fukuda, 1984). Considering that the putative genes used for this project originate from organisms living in an environment with pressure values of 250 bar (25 MPa), it is plausible that both their activity and stability has decreased when being expressed at atmospheric pressure (1 bar), which would make it more difficult to produce the target peptidases in soluble, functional form.

4.4 Future work

Based on the results obtained from this study, further work should investigate the use of different expression systems in order to get production of more soluble target protein from the candidate genes chosen for functional testing and characterization. Using a host organism with similar codon usage is one optimisation possibility. The use of thermophiles as host organisms can be applicable, as recent studies have shown that inducible gene expression systems in such strains have been used successfully (Suzuki, 2013).

The possibility of solubilizing and refolding of the target protein observed in the pellet samples should also be investigated. Solubilization is normally done using chemical denaturants. However, there is no universal method for refolding proteins, and the recovering yield is usually low. Continuous development of new methods, such as oxidative column chromatography, will improve the number of proteins that can be refolded (Middelberg, 2002).

5. CONCLUSION

In this study, three putative genes for novel peptidases were identified from a metagenomic DNA library constructed from total DNA of samples from two deep subsea oil reservoirs on the Norwegian Continental Shelf. Each of these sequences showed high degree of homology to a known peptidase of sub-classification subtilisin-like serine peptidases, aminopeptidases or carboxypeptidases. Hence, given the abbreviations SLP01, APT01 and CPT01.

The T7 promoter expression system was used for protein production of the recombinant proteins in *Escherichia coli* strains BL21 (DE3) CodonPlus-RIPL and ER2566, using the plasmid vector pET21. Several attempts for protein production were performed, using culture volumes of 100 ml or 1 litre and temperatures 16 or 37 °C for incubation after induction of target gene expression with IPTG. Desired protein production was verified using SDS-PAGE analysis. Visible target protein was observed for the APT01 and CPT01 candidate produced by the BL21 strain, although in an insoluble form. Some soluble APT01 was observed from the 100 ml culture of BL21 cells, incubated at 16 °C. No visible production of the target protein for the SLP01 candidate could be observed for any of the production attempts.

Heat denaturation series was performed on the sample containing soluble APT01, at temperatures of 65 and 80 °C for a duration of 5, 15, 30, 60 and 120 minutes. The results showed APT01 to be thermostable at 65 °C, but none at 80 °C.

An activity assay, using skim milk agar (MA) plates, was performed on all samples from the protein production. No activity was observed for any of the three candidates.

REFERENCES

ALTSCHUL, S. F., MADDEN, T. L., SCHÄFFER, A. A., ZHANG, J., ZHANG, Z., MILLER, W. AND LIPMAN, D. J. 1997. Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. *Nucleic Acids Research,* 25**,** 3389–3402.

ARNAU, J., LAURITZEN, C., PETERSEN, G. E. AND PEDERSEN, J. 2006. Current strategies for the use of aYnity tags and tag removal for the purification of recombinant proteins. *Protein Expression and Purification,* 48**,** 1-13.

BANEYX, F. 1999. Recombinant protein expression in Escherichia coli. *Current Opinion in Biotechnology,* 10**,** 411–421.

BIDDLE, J. F., FITZ-GIBBON, S., SCHUSTER, S.C., BRENCHLEY, J.E. AND HOUSE, C.H. 2008. Metagenomic signatures of the Peru Margin subseafloor biosphere show a genetically distinct environment. *PNAS,* 105**,** 10583–10588.

BJÖRKMANGROUP. 2001. *TCA protein precipitation protocol* [Online]. California, USA. Available: http://www.its.caltech.edu/~bjorker%20/Protocols/TCA_ppt_protocol.pdf [Accessed April 25th 2014].

BROWN, N. J. 2003. *Biological markers and diagnostic test for angiotensin converting enzyme inhibitor and vasopeptidase inhibitor-associated angioedema.* USA patent application US20030180828 A1.

CENTER, D.-F. A. H. C. 2004-2014. *Vector: pEt21a* [Online]. Harvard Medical School. Available: http://plasmid.med.harvard.edu/PLASMID/GetVectorDetail.do?vectorid=316 [Accessed June 19th 2014].

CLARK, D. P. 2010. Recombinant DNA Technology. *In:* CLARK, D. P. (ed.) *Molecular biology.* London, UK: Elsevier.

COHEN, S. N., CHANG, A. C. Y., BOYER, H. W. AND HELLING, R. B. 1973. Construction of Biologically Functional Bacterial Plasmids In Vitro. *Proc. Nat. Acad. Sci.,* 70**,** 3240-3244.

DIPIPPO, J. L., NESBØ, C.L., DAHLE, H., DOOLITTLE, W.F., BIRKLAND, N.-K. AND NOLL, K.M. 2009. Kosmotoga olearia gen. nov., sp. nov., a thermophilic, anaerobic heterotroph isolated from an oil production fluid. *International Journal of Systematic and Evolutionary Microbiology,* 59**,** 2991–3000.

EISENMENGER, M. J. A. R.-D.-C., J I. 2009. High pressure enhancement of enzymes: A review. *Enzyme and Microbial Technology,* 45**,** 331-347.

FOOPHOW, T., TANAKA, T., ANGKAWIDJAJA, C., KOGA, Y., TAKANO, K. AND KANAYA, S. 2010. Crystal Structure of a Subtilisin Homologue, Tk-SP, from Thermococcus kodakaraensis: Requirement of a C-terminal β-Jelly Roll Domain for Hyperstability. *Journal of Molecular Biology,* 400**,** 865-877.

FUKUDA, M. A. K., S. 1984. Pressure dependence of thermolysin catalysis. *Eur. J. Biochem.,* 142**,** 565-570.

GABOR, E. M., ALKEMA, W. B. L. AND JANSSEN, D. B. 2004. Quantifying the accessibility of the metagenome by random expression cloning techniques. *Environmental Microbiology,* 6**,** 879-886.

HANDELSMAN, J. 2004. Metagenomics: Application of Genomics to Uncultured Microorganisms. *Microbiology and Molecular Biology Reviews,* 68**,** 17.

HICKS, M. A. A. P., K.L.J. 2014. Bioprospecting in the Genomic Age. *Advances in Applied Microbiology,* 87**,** 11-146.

JIANG, X., OOHIRA, K., IWASAKI, Y., NAKANO, H., ICHIHARA, S. AND YAMANE, T. 2001. Reduction of Protein Degradation by Use of Protease-Deficient Mutants in Cell-Free Protein Synthesis System of Escherichia coli. *J. Biosci. Bioeng.,* 93**,** 151-156.

KALE, A., PIJNING, T., SONKE, T., DIJKSTRA, B. W. AND THUNNISSEN, A.-M. W.H. 2010. Crystal Structure of the Leucine Aminopeptidase from Pseudomonas putida Reveals the Molecular Basis for its Enantioselectivity and Broad Substrate Specificity. *Journal of Molecular Biology,* 398**,** 703-714.

KIRK, O., BORCHERT, T. V. AND FUGLSANG, C. C. 2002. Industrial enzyme applications. *Current Opinion in Biotechnology,* 13**,** 345-351.

KOTLAR, H. K., LEWIN, A., JOHANSEN, J., THRONE-HOLST, M., HAVERKAMP, T., MARKUSSEN, S., WINNBERG, A., RINGROSE, P., AAKVIK, T., RYENG, E., JAKOBSEN, K., DRABLØS, F. AND VALLA, S. 2011. High coverage sequencing of DNA from microorganisms living in an oil reservoir 2.5 kilometres subsurface. *Environmental Microbiology Reports,* 3**,** 674–681.

KUMAR, S., TSAI, C.-J. AND NUSSINOV, R. 2000. Factors enhancing protein thermostability. *Protein Engineering,* 13**,** 179-191.

LANG, K. S., ANDERSON, J.M., SCHWARZ, S., WILLIAMSON, L., HANDELSMAN, J. AND SINGER, R.S. 2010. Novel Florfenicol and Chloramphenicol Resistance Gene Discovered in Alaskan Soil by Using Functional Metagenomics. *Applied and Environmental Microbiology,* 76**,** 5321–5326.

LEE, M. M., ISAZA, C. E., WHITE, J. D., CHEN, R. P.-Y., LIANG, G. F.-C., HE, H. T.-F., CHAN, S. I. AND CHAN, M. K. 2009. Insight into the substrate length restriction of M32 carboxypeptidases: Characterization of two distinct subfamilies. *Proteins: Structure, Function, and Bioinformatics,* 77**,** 647–657.

LEWIN, A., JOHANSEN, J., WENTZEL, A., KOTLAR, H.K., DRABLØS F. AND VALLA, S. 2013. The microbial communities in two apparently physically separated deep subsurface oil reservoirs show extensive DNA sequence similarities. *Environmental Microbiology,* 16**,** 545-558.

LEWIN, A., WENTZEL, A. AND VALLA, S. 2012. Metagenomics of microbial life in extreme temperature environments. *Current Opinion in Biotechnology,* 24**,** 1-10.

LI, L.-L., MCCORKLE, S.R., MONCHY, S., TAGHAVI, S. AND VAN DER LELIE, D. 2009. Bioprospecting metagenomes: glycosyl hydrolases for converting biomass. *Biotechnology for Biofuels,* 2.

LOENEN, W. A. M., DRYDEN, D. T. F., RALEIGH, E. A., WILSON, G. G. AND MURRAYY, N. E. 2014. Highlights of the DNA cutters: a short history of the restriction enzymes. *Nucleic Acids Research,* 42**,** 3-19.

MIDDELBERG, A. P. J. 2002. Preparative protein refolding. *Trends in Biotechnology,* 20**,** 437-443.

MOSS, G. P. 1992. *Enzyme Nomenclature. Recommendations of the Nomenclature Committee of the International Union of Biochemistry and Molecular Biology.* [Online]. London, UK: NC-IUBMB. Available: http://www.chem.qmul.ac.uk/iubmb/enzyme/ [Accessed April 19th 2014].

MOZHAEV, V. V., LANGE, R., KUDRYASHOVA, E. V. AND BALNY, C. 1996. Application of High Hydrostatic Pressure for Increasing Activity and Stability of Enzymes. *Biotechnology and Bioengineering,* 52**,** 320-331.

PETROSINO, J. F., HIGHLANDER, S., LUNA, R. A., GIBBS, R. A. AND VERSALOVIC, J. 2009. Metagenomic Pyrosequencing and Microbial Identification. *Clin Chem. ,* 55**,** 856-866.

PINGOUD, A., WILSON, G. G. AND WENDE, W. 2014. Type II restriction endonucleases – a historical perspective and more. *Nucleic Acids Research,* Advance Access.

RAWLINGS, N. D., MORTON, F.R. AND BARRETT, A.J. 2007. An Introduction to Peptidases and the MEROPS Database. *Industrial Enzymes. Structure, Function and Applications.* The Netherlands: Springer

RONDON, M. R., AUGUST, P.R., BETTERMANN, A.D., BRADY, S.F., GROSSMAN, T.H., LILES, M.R., ET AL. 2000. Cloning the Soil Metagenome: a Strategy for Accessing the Genetic and Functional Diversity of Uncultured Microorganisms. *Applied and Environmental Microbiology,* 66**,** 2541–2547.

SABREE, Z. L., RONDON, M. R. AND HANDELSMAN, J. 2009. Metagenomics. *In:* SCHAECHTER, M. (ed.) *Encyclopedia of Microbiology.* University of Wisconsin-Madison, Madison, WI, USA: Elsevier.

SANCHO, J., SERRANO, L. AND FERSHT, A. R. 1992. Histidine Residues at the N- and C-Termini of a-Helices: Perturbed pKas and Protein Stability. *Biochemistry,* 31**,** 2253-2258.

SANZ, Y. 2007. Aminopeptidases. *In:* MACCABE, J. P. A. A. P. (ed.) *Industrial Enzymes. Structure, Function and Applications.* The Netherlands: Springer

SCHÖNFELD, H.-J., SCHMIDT, D., SCHRÖDER, H. AND BUKAU, B. 1995. The DnaK Chaperone System of Escherichia coli: Quaternary Structures and Interactions of the DnaK and GrpE Components. *J. Biol. Chem.,* 270**,** 2183-2189.

STREIT, W. R., DANIEL, R. AND JAEGER, K.-E. 2004. Prospecting for biocatalysts and drugs in the genomes of non-cultured microorganisms. *Current Opinion in Biotechnology,* 15**,** 6.

SUZUKI, H., YOSHIDA, K-I. AND OHSHIMAC, T. 2013. Polysaccharide-Degrading Thermophiles Generated by Heterologous Gene Expression in Geobacillus kaustophilus HTA426. *Applied and Environmental Microbiology,* 79**,** 5151-5158.

SØRENSEN, H. P. A. M., K. K. 2005. Soluble expression of recombinant proteins in the cytoplasm of Escherichia coli. *Microbial Cell Factories,* 4.

TABOR, S. 2001. Expression Using the T7 RNA Polymerase/Promoter System. *Curr. Protoc. Mol. Biol.,* 11**,** 16.2.1–16.2.11.

TAYLOR, R. G., WALKER, D. C. AND MCLNNES, R. R. 1993. E.coli host strains significantly affect the quality of small scale plasmid DNA preparations used for sequencing. *Nucleic Acids Research,* 21**,** 1677-1678.

TERPE, K. 2006. Overview of bacterial expression systems for heterologous protein production: from molecular and biochemical fundamentals to commercial systems. *Appl Microbiol Biotechnol,* 72**,** 211–222.

TORREZ, M., SCHULTEHENRICH, M. AND LIVESAY, D. R. 2003. Conferring Thermostability to Mesophilic Proteins through Optimized Electrostatic Surfaces. *Biophysical Journal,* 85**,** 2845-2853.

WU, C., SEITZ, P. K. AND FALZON, M. 2000. Single-column purification and bio-characterization of recombinant human parathyroid hormone-related protein (1–139). *Molecular and Cellular Endocrinology,* 170**,** 163-174.

YABUKI, Y., KUBOTA, K., KOJIMA, M., INOUE, H. AND TAKAHASHI, K. 2004. Identification of a glutamine residue essential for catalytic activity of aspergilloglutamic peptidase by site-directed mutagenesis. *FEBS Letters,* 569**,** 161–164.

APPENDIX A.1 – Gene sequences

CA▼TATG = *Nde*I restriction site, ATG = START codon

A▼GATCT = *Bgl*II restriction site

T▼GATCA = *Bcl*I restriction site, TGA = STOP codon

CATx6 = His-tag

TAA = STOP codon

C▼TCGAG = *Xho*I restriction site

CCAAT/CAAT, ATTGG = Protective bases

Aminopeptidase - APT01

CCAATCA▼TATGCGTATCGCCGTGGACTTTAGCCCGCCGCTGACCCTGAAAACCTCGTGTCTGGTGCT
GGGTGTGTGGCAAGATCGTACCAAAACGCAGCTGCTGAAAGACCTGGATGACAGCCTGCATGGTGCCC
TGCGTCTGTCAGTTGACTCGGGCGCATTTTCTGGCAAACAAGGTGAAACCCAGCTGTTCCAAGCTGGT
GCGGGTGTCCCGGCAGCACGTATTCTGCTGGTTGGCCTGGGTCCGTTTGAAGCAGCTGATGCAGACGT
TCTGCGTCGCGGTGCAGCAGAAGCAGCTCGTGTTCTGCAGAAACAACACGTCGGCGAAGCCGGTATGG
TGTGCCCGGAATCACCGGCACGTCTGAACCCGGCTCAGGTGGCACAAGCACTGGCTGAAGGCCTGCTG
CTGGCGTCGTATCGTTTCGATCGCTACCTGACGCAGAAACGTGAAGAACTGCCGCCGCTGCTGGAAAA
TGTGGTTATTCTGGCGGCCCAAGAACATGTGGAATCCCTGGCAGCTGCGGTTGATCGTGCACGCCTGA
TCTGCCGTGGTGTGTCACTGGCGCGCGACCTGGTTAACGAACCGGGCAATGTCAAATCGCCGGAATTT
CTGGCGGAACAGGCCCGTAAAATGGCCGCACGCCACAACATTGTGTGTACCGTTCTGGAACGCGATCA
GCTGGAAAAAGAAGGCTTCGGTGCACTGCTGGCAGTTGCACAAGGTAGCGCACGTCCGCCGCGTCTGA
TTGTCCTGGAATATCGTGGCGGTGAACCGCAGGAACGTCCGGTCGCGCTGGTTGGCAAAGGTGTCGTG
TTTGACAGCGGCGGTATCTCTCTGAAACCGGGCGAAAAAATGGATGAAATGAAAATGGACATGGCAGG
CGCTGCGGCCGTTATTGGTGCTATGGAAGCAGCTGCGGGTCTGCGTCTGCCGGTTAACCTGGTCGGCA
TCGTGCCGGCGGTTGAAAATCTGCCGAGCGCAACCGCTTATCGCCCGGGTGATATTATCACCAGTCTG
TCCGGCCAGACGATTGAAGTGCTGAATACCGATGCGGAAGGTCGTCTGATCCTGGCAGACGCACTGAC
GTACGCGGGTCGTATGGAACCGCGTACCGTGATTGATCTGGCCACCCTGACGGGCGCGTGCATTATCG
CCCTGGGTCATGAAGCAAGTGCTGTTTTCTCCAACCACGACGAACTGGCGCGTAATCTGATGCGCGCC
GGCGAAAGCTCTGGTGAACGCCTGTGGCAGCTGCCGCTGTGGGAAAACTATGGTAAACAACTGAAAAG
TGAAGTCGCCGATGTGAAAAATACGGGCGGTCGTCCGGCAGGTACCATTACGGCCGCATCCTTTCTGC
AGCGTTTCGCACCGGAAGGTCCGTGGGCTCATGTCGATATCGCAGGCACGGCTTGGGAAGACAAAGGT
ACCGCACTGTGTCCGAAAGGTGCAACCGGCGTTGGTGTCCGTATGCTGATCGATTACCTGGAACAGGT
GA ▼GATCTT▼GATCACATCATCATCATCATCATTAACC▼TCGAGATTGG (1540bp)

Carboxypeptidase - CPT01

CAATCA▼TATGGATGCTCTGGAAAAACTGAAAGCTCTGTCGAAAGAACTGGAACTGATTAACGCCGCC
GTCGCTACGATGGCCTGGGACCAACGCACCTATATGCCGCCGAAATCAGCAGGCACGCGTTCGGAAGC
TATTGGTTACCTGAGCACCATCGCATTCAAAAAATTCATCTCTGATGAAACGGGCGAAATTATCCGTG
AACTGGAAAAAGAAGATAACTTTAATCGCCTGGACGAAAATGAAAAAGCGATGGTGCGTATTGCCAAA
CGCGAATATGAAAAAGCAAAAGCTATCCCGCCGGAACTGTTTCAGAAATTCACCATTACGGCGAGCAA
ATCTGAAACCGTTTGGGAACAGGCGAAGAAAAACAACGATTTCAAAAGTTTCCAACCGTATCTGGAAG
AACTGCTGGAAATGCTGCGCGAAATGGCGGAACTGTATGGCTACAAAGAAAACCCGTATGATGCCCTG
CTGGACAAATACGAACCGGGTATCACCACGCGTAAACTGAAGAAAATTATCGAAACCCTGAAAGCGGA
ACTGATTCCGTTTCTGCGCGAAATTATCGAACAGAAAGGTAAAACGGATCCGTCAATCCTGTATGGCC
GTTTCGGTAAAAAAGCCCAAGAAAAACTGTCGATCCGCGCACTGAAAGCTATTGGCTACGATTTTGAA
GCAGGTCGTCTGGACGAAACCGTGCATCCGTTCACGATTAGTCTGGGCGCTGGTGATGTGCGCGTTAC
CACGAAATATGACCCGCACTTTCTGCAGCCGTCCCTGTACGGCACCTTCCATGAAGGCGGTCACGCGC
TGTATGAACAAGGTCTGCCGGAAGAATTTAAATATACCCCGATCTACGGCGCCGTTAGCCTGGGTATT
CATGAAAGTCAGTCCCGTATGATGGAAAATATGGTCGCACGCTCTTATGAATTCCTGAAATTTTTCTA
CCCGGAAATCAAAAAAGTGTTTCCGAAACAATTCGGCCGTGTTAGTCTGGATCGTTTCTATCGCGCAA
TCAATCATGTCGAACCGTCCCTGATTCGTATCGAAGCTGACGAAGTGACCTACAACTTTCACATTATG
CTGCGCTTCGAACTGGAAGAAGCGATGCTGAACAAATCTCTGGAAGTTAAAGACCTGCCGCGTGCGTG
GAACGAAAAAATCAAAGAATACCTGGGCATCGAACCGCAGAACGATGCAGAAGGTGTTCTGCAAGACG
TCCACTGGGCTAATGGCATGATTGGTTATTTTCCGAGCTACATGCTGGGCAACCTGTACGCGGCCCAG
CTGTTCGCGAAAGCCGAAGAAGAAATCCCGAAACTGCGTAAAAACATCGAAAAGGGTAACGTCGCGGT
GCTGATCGAATGGCTGCGTGAAAATATTCATCGCCACGGCAAAAAATATCTGCCGGAAGAACTGATTA
AAATCAGCACCGGTGAAGAACTGAACCCGGAATATTTTATTCGCTACATCAAAGAAAAATATACGAAA
ATTTACGAAATCA ▼GATCTT▼GATCACATCATCATCATCATCATTAAC▼TCGAGATTGG
(1551bp)

Subtilisin-like peptidase - SLP01

CCAATC▼ATATGCGTAAAGTTCTGGGTCTGCTGGTCGCGTTTCTGATGCTGGGTTTCGTTGTCGCTTC
GGTCGCCGCTCTGCCGTCCCCGGATACCAAACCGTATACCCAGCCGAAAAACTACGGCCTGCTGACGC
CGGGTCTGTTTCGTAAAGCACAGCGCATGGATTGGGAACAAGAAGTGAGCACCATTATCATGTTCGAC
ACGCCGCGTAATCAGCGCATCGCGCTGAAAATTCTGAAAGCTCTGGGCGCGGAAGTGAAATATCAATA
CGAAGTTATTCCGGCCATCGCAGTGAAAATGAAAGTTCGTGATCTGCTGGTTATTGCGGGCTTTCTGG
ACGCCACCAGCTCTGGTCGCAGTAAAGTCCAGATTCCGGGCATCCAGTTCATCCAAGAAGATTACAAA
GTCAAAGTGGCCGTTGAAACCGAAGGCCTGGACGAATCCGCGGCCCAAGTGATGGCAACCAACATGTG
GAATCTGGGTTATGATGGTTCAGGCATTACCATCGGCATTATCGATACGGGTATCGACGCGTCGCATC
CGGATCTGCAGGGTAAAGTTATTGGCTGGGTTGACTATGTCAACGGCCGTAGTTCCCCGTACGATGAC
AATGGTCATGGCACCCACGTGGCAAGCATTGCAGCTGGTACGGGTGCAGCCTCTAACGGTAAATATAA
AGGCATGGCCCCGGGTGCAAAACTGGTCGGCATCAAAGTGCTGGGTGCAGATGGTAGCGGCTCTATTA
GTGATATTATCGCAGGTGTTGACTGGGCTGTCAAAAACAAAGATAAATACGGCATCAAAGTTATTAAT
CTGAGCCTGGGCTCATCGCAGAGCTCTGATGGTACCGACTCCCTGTCACAAGCCGTGAACAATGCTTG
GGATGCGGGTCTGGTGGTTTGCGTTGCAGCTGGTAACTCTGGCCCGAATAAATATACCGTCGGTAGCC
CGGCGGCCGCATCTAAAGTGATTACCGTTGGCGCGGTCGATAAATACGACGTGATTACCGATTTTAGT
TCCCGTGGCCCGACGGCAGACAACCGTCTGAAACCGGAAGTCGTGGCTCCGGGTAATTGGATTATCGC
TGCGCGTGCGTCAGGTACCTCGATGGGTCAGCCGATCAATGATTATTACACGGCAGCACCGGGTACCA
GCATGGCAACGCCGCATGTTGCAGGTATTGCTGCACTGCTGCTGCAGGCACACCCGTCTTGGACCCCG
GATAAAGTTAAACGTGCACTGATCGAAACGGCTGACATTGTCAAACCGGATGAAATCGCCGACATTGC
GTATGGCGCTGGCCGCGTGAACGCGTACAAAGCCGCATATTACGATAATTATGCGAAACTGACCTTTA
CGGGCTACGTTGCCAACAAAGGTTCACAGACCCATCAATTTACGATCTCGGGTGCAGGTTTCGTGACC
GCAACGCTGTATTGGGATAATTCGGGTAGCGATATTGACCTGTATCTGTACGATCCGAACGGCAATCA
GGTCGACTATAGTTACACCGCGTATTACGGCTTCGAAAAAGTGGGTTATTACAACCCGGCTGCGGGTA
CCTGGACGATCAAAGTTGTCTCTTATAGTGGCTCCGCTAACTACCAGGTTAATGTGGTTAGCGATGGT
ACCCTGGGTCAACCGGGCGGTGGCGAACCGGCACCGGAACCGACCCCGGAACCGACGGTCGATGAAAA
AACCTTTACGGGCACCGTGCACCGTTATTACGATCGCAGTGACACGTTCACCATGACGGTGAATTCCG
GCGCGACCAAAATTACGGGTGATCTGACCTTTGACACGGGTTATCATGATCTGGACCTGTACCTGTAT
GATCCGAACAAAAATCTGGTGGACCGTAGTGAATCATCGAACTCCTATGAACACGTTGAATACACCAA
TCCGGCGCCGGGCACGTGGTATTTCCTGGTGTACGCATATTACACCTATGGCTACGCTAGCTATCAGC
TGGATGTCAAAGTGTATTACGGTA ▼GATCTT▼GATCACATCATCATCATCATCATTAAC▼TCGAGAT
TGG (2038bp)

APPENDIX A.2 – SDS-PAGE result

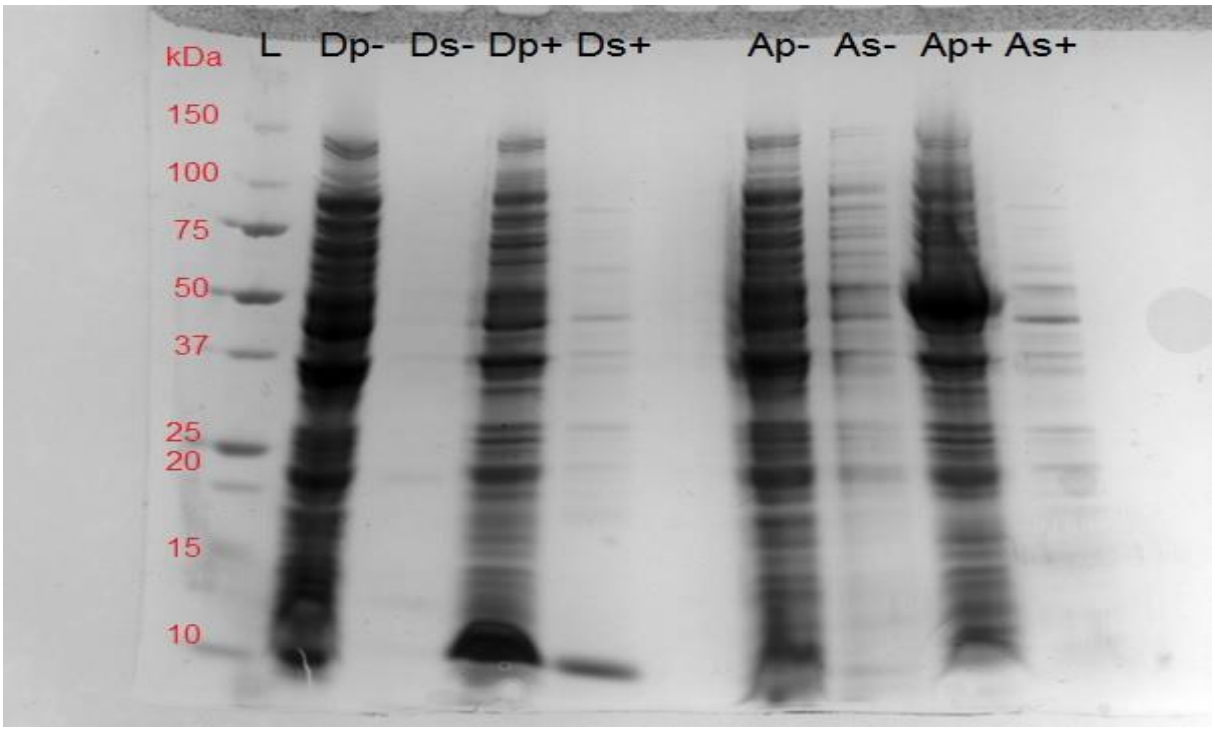

Figure A.2.1: *SDS-PAGE of pellet (p) and supernatant (s) of induced (+) and non-induced (-) BL21 DE3 codon plus cells transformed with control plasmid pOD1 (D), or pET21 plasmid carrying the gene for APT01 (A). Biomarker from BioRad was used (L), with sizes in kilo Dalton shown in red.*

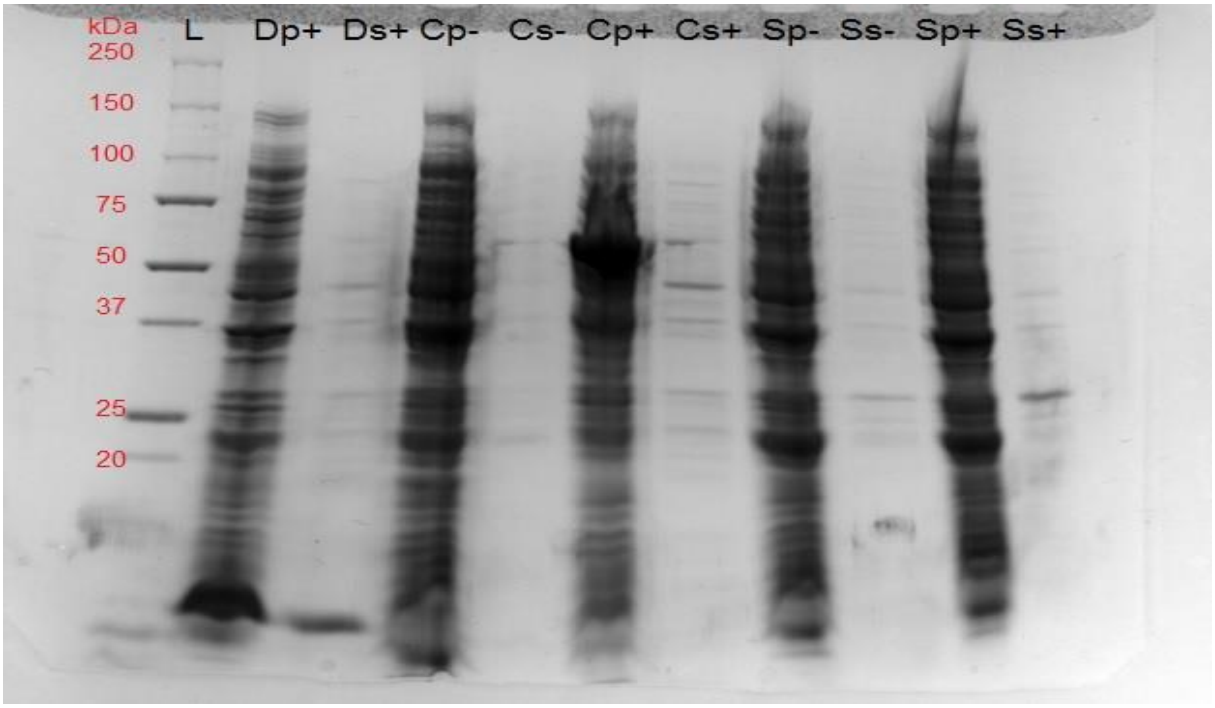

Figure A.2.2: *SDS-PAGE of pellet (p) and supernatant (s) of induced (+) and non-induced (-) BL21 DE3 codon plus cells transformed with control plasmid pOD1 (D), or pET21 plasmid carrying the gene for CPT01 (C) or SLP01 (S). Biomarker from BioRad was used (L), with sizes in kilo Dalton shown in red.*

APPENDIX A.3 – OD_{600nm} values

Table A.3.1: *Absorbance measurements for protein production at 37 °C, using 50/100 ml culture (section 4.3.1), for induced (+) and non-induced (-) BL21 cells transformed with pET21_APT01 or pET21_CPT01.*

Time (tt:mm)	APT+ (OD_{600nm})	APT- (OD_{600nm})	CPT+ (OD_{600nm})	CPT- (OD_{600nm})
00:00	0,025		0,023	
01:00	0,05		0,056	
03:10	0,873		0,91	
04:00	1,087	1,058	1,048	1,055
05:00	1,319	1,47	1,532	1,48
06:00	1,481	1,661	1,699	1,647
07:05	1,533	1,679	1,556	1,687
07:40	1,655	1,797	-	1,802
08:50	1,516	1,81	-	1,802
09:00	-	1,825	-	-
09:20	-	1,839	-	-
09:40	-	1,851	-	-

Table A.3.2: *Absorbance measurements for protein production at 37 °C, using 50/100 ml culture (section 4.3.1), for induced (+) and non-induced (-) BL21 cells transformed with pET21_SLP01 or pOD1.*

Time	SLP+ (OD_{600nm})	SLP- (OD_{600nm})	pOD+ (OD_{600nm})	pOD- (OD_{600nm})
00:00	0,017		0,021	
01:00	0,038		0,046	
03:10	0,554		0,708	
04:00	0,875		0,928	0,957
05:00	1,387	1,398	1,292	1,433
06:00	1,539	1,604	1,371	1,618
07:05	1,393	1,617	1,404	1,705
07:25	-	1,721	1,463	1,774
08:05	-	1,702	1,483	1,808
08:50	-	-	1,471	1,817
09:20	-	-	-	1,826

Table A.3.3: *Absorbance measurements for protein production at 16 °C, using 50/100 ml culture (section 4.3.2), for induced (+) and non-induced (-) BL21 cells transformed with pET21_APT01.*

Time	APT+ (OD_{600nm})	APT- (OD_{600nm})
00:00	0,043	
01:30	0,078	
03:00	0,581	
03:30	0,985	
04:50	1,011	1,020
05:50	1,047	1,093
06:50	1,163	1,140
07:50	1,189	1,170
08:50	1,215	1,231
09:50	1,271	1,291
10:50	1,326	1,335

Table A.3.4: *Absorbance measurements for protein production at 16 °C, using 50/100 ml culture (section 4.3.2), for induced (+) and non-induced (-) BL21 cells transformed with pET21_CPT01.*

Time	CPT+ (OD_{600nm})	CPT- (OD_{600nm})
00:00	0,023	
01:30	0,115	
03:00	0,775	
03:20	0,919	
04:30	1,045	0,999
05:30	1,112	1,080
06:30	1,142	1,160
07:30	1,196	1,190
08:30	1,217	1,202
09:30	1,247	1,242
10:30	1,314	1,272

Table A.3.5: *Absorbance measurements for protein production at 16 °C, using 50/100 ml culture (section 4.3.2), for induced (+) and non-induced (-) BL21 cells transformed with pET21_SLP01.*

Time	(B)SLP+ (OD_{600nm})	(B)SLP- (OD_{600nm})
00:00	0,034	
01:30	0,033	
03:00	0,131	
03:50	0,360	
05:20	1,137	
06:30	1,154	1,156
07:30	1,192	1,203
08:30	1,225	1,216
09:30	1,260	1,274
10:30	1,301	1,32
11:10	1,313	1,323

Table A.3.6: *Absorbance measurements for protein production at 16 °C, using 50/100 ml culture (section 4.3.2), for induced (+) and non-induced (-) BL21 cells transformed with pOD1.*

Time	(B)pOD1+ (OD_{600nm})	(B)pOD1- (OD_{600nm})
00:00	0,029	
01:30	0,031	
03:00	0,177	
03:30	0,541	
04:50	1,047	
06:50	1,154	1,133
07:50	1,166	1,181
08:50	1,212	1,211
09:50	1,244	1,263
10:50	1,304	1,308

Table A.3.7: *Absorbance measurements for protein production at 16 °C, using 50/100 ml culture (section 4.3.2), for induced (+) and non-induced (-) ER2566 cells transformed with pET21_SLP01.*

Time	(E)SLP+ (OD_{600nm})	(E)SLP- (OD_{600nm})
00:00	0,019	
01:30	0,078	
03:00	0,593	
03:30	0,926	
04:50	0,996	1,021
05:50	1,041	1,021
06:50	1,057	1,082
07:50	1,123	1,11
08:50	1,105	1,131

Table A.3.8: *Absorbance measurements for protein production at 16 °C, using 50/100 ml culture (section 4.3.2), for induced (+) and non-induced (-) ER2566 cells transformed with pOD1.*

Time	(E)pOD1+ (OD_{600nm})	(E)pOD1- (OD_{600nm})
00:00	0,018	
01:30	0,051	
03:00	0,274	
03:50	0,815	
04:15	0,949	
05:20	0,998	1,027
05:50	1,111	1,086
07:20	1,089	1,081

Table A.3.8: *Absorbance measurements for protein production at 16 °C, using 1 litre culture (section 4.3.4), for induced (+) and non-induced (-) BL21 cells transformed with pET21_APT01, pET21_CPT01 or pOD1.*

	OD_{600nm}		OD_{600nm}		
Time	**pOD1**	**Time**	**APT01**	**CPT01**	**IPTG induced (yes/no)**
00:00	0,1085	**00:00**	0,1112	0,1115	N
01:00	0,1111	**01:00**	0,1182	0,1158	N
02:00	0,1380	**02:00**	0,1868	0,1326	N
03:00	0,2992	**03:00**	1,1065	0,5235	N
04:00	1,6570	**03:30**	2,1800	-	N
04:30	3,1570	**04:00**	2,9190	1,3960	N
05:00	4,0760	**04:35**	4,7940	2,2080	N
06:45	6,5060	**05:45**	6,4220	4,1680	Y
07:50	6,3160	**06:50**	6,6240	4,7420	Y
08:45	6,7960	**07:45**	6,5180	4,9140	Y
09:45	7,3340	**08:45**	6,4380	5,0040	Y
10:45	7,5020	**09:45**	6,7460	5,3580	Y