Can Sterol Carrier Protein-2 function as a solubility tag in *E. coli*?

Amanda Lundén

Examinator, Jordi Altimiras
Tutor, Johan Edqvist
Can Sterol Carrier Protein-2 function as a solubility tag in E.coli?

Amanda Lundén

Expressing foreign proteins in E.coli is a major challenge because they often tend to develop into unsolvable and inactive proteins. They aggregate into so called inclusion bodies which prevent expression of the protein. This problem might be avoided by fusing the gene of the foreign protein with a soluble protein called solubility tags, which function is to enhance the solubility of the foreign protein. This report investigates whether Sterol Carrier Protein-2 (SCP-2) could function as a solubility tag. The experiment was carried out by fusing SCP-2 to two recombinant proteins, Green fluorescent protein (GFP) and a form of chloroamphenicol acetyl transferase (CATΔ9). The gene fusion was then inserted into a pET-15 vector and transformed into the E.coli strain BL21(DE3) to be expressed. The results obtained from Western blot and PageBlue staining indicates that SCP-2 does not enhance the solubility of GFP or CATΔ9 since neither of them was expressed. Furthermore, previous studies have shown that GFP can in fact be expressed using maltose binding protein (MBP) as a solubility tag. Unfortunately, no success has been made regarding CATΔ9. In conclusion, regarding the results from this report, SCP-2 does not function as a solubility tag. However, further studies should be carried out on SCP-2 with more experiments before rejecting the possibility to use SCP-2 as a solubility tag.
# Content

1. **Abstract** ........................................................................................................... 3
2. **Introduction** ........................................................................................................ 3
3. **Material & methods** ............................................................................................. 5
   3.1 *E. coli* strains ..................................................................................................... 5
   3.2 pET-15b vector ................................................................................................... 6
   3.3 Genes .................................................................................................................. 7
   3.4 Transformation and DNA manipulation ............................................................. 7
   3.5 Protein expression .............................................................................................. 8
      3.5.1 Total cell protein fraction .............................................................................. 9
      3.5.2 Soluble cytoplasmic protein fraction ............................................................ 9
      3.5.3 Western blot .................................................................................................. 9
      3.5.4 PageBlue staining ....................................................................................... 10
4. **Results** .................................................................................................................. 10
   4.1 Construction of fusion genes .............................................................................. 10
   4.2 Expression of fusion proteins ............................................................................ 13
5. **Discussion** ............................................................................................................ 15
   5.1 Societal & ethical considerations ....................................................................... 17
6. **Acknowledgement** .............................................................................................. 18
7. **References** .......................................................................................................... 18
1 Abstract

Expressing foreign proteins in E.coli is a major challenge because they often tend to develop into insoluble and inactive proteins. They aggregate and form inclusion bodies which lead to a prevented protein expression. This problem might be avoided by fusing the gene of the foreign protein with a soluble protein called a solubility tag which function is to enhance the solubility of the foreign protein. This report investigates whether Sterol Carrier Protein-2 (SCP-2) could function as a solubility tag. The experiment was carried out by fusing SCP-2 to two recombinant proteins, Green fluorescent protein (GFP) and a form of chloroamphenicol acetyl transferase (CATΔ9). The gene fusions were then inserted into a pET-15b vector and transformed into the E.coli strain BL21(DE3) to be expressed. The results obtained from Western blot and PageBlue staining indicates that the gene fusion with SCP-2 might not be expressed. This means that it is not possible to say if SCP-2 does function as a solubility tag together with GFP and CATΔ9. In conclusion, regarding the results from this report, it is hard to state if SCP-2 function as a solubility tag and should therefore not be excluded. Further studies should be carried out on SCP-2 with more experiments before rejecting the possibility to use SCP-2 as a solubility tag.

2 Introduction

The production of recombinant proteins in E.coli often results in biological inactive and insoluble proteins (Kapust & Waugh 1999). This is due to the fact that the foreign proteins aggregate and form insoluble inclusion bodies which lead to a prevented protein expression (Costa et al. 2014). This is regarded as a substantial problem which, in some cases, can be overcome by fusing the gene encoding the recombinant protein of interest with a soluble protein, called a solubility tag (Bell et al. 2013). The function of the solubility tag is to enhance the folding which consequently leads to a higher solubility of the recombinant protein. This is because the protein does not aggregate and form inclusion bodies when folded properly. (Kapust & Waugh 1999).

In addition, according to Waugh (2005), there is a general perception regarding the problem with inclusion bodies, that it can be avoided by a different method rather than using solubility tags. Using a eukaryotic host instead, such as yeast or insects cells, is thought to solve the problem with inclusion bodies. In a recent study carried out by Dalton & Barton (2014), they compared prokaryotic and eukaryotic expression systems and came up with some advantages and disadvantages regarding eukaryotic expression systems. Yeast for instance; have provided high yields of
recombinant proteins at a fairly low cost and high time efficiency. Unfortunately, yeast produces proteases which can lead to protein degradation and thus renders them problematic (Dalton & Barton, 2014). Until further studies on eukaryotic expression systems have been carried out and a more efficient expression system has been found, *E. coli* is the most proven host system for producing recombinant protein (Waugh, 2005). Therefore, the solubility problem in this host need to be overcome, and one way to achieve this is to discover new solubility tags.

New solubility tags are frequently emerging and a lot of studies are made on the subject. Some solubility tags which have been investigated are the Fh8 solubility tag and the most common glutathione S-transferase (GST) and maltose binding protein (MBP) (Costa et al. 2014, Kapust & Waugh 1999). These are not universally applicable and therefore new solubility tags need to be discovered (Cantu-Bustos et al. 2016). In a recent study by Cantu-Bustos et al. (2016) they investigated whether the protein CusF (Cu sensitive free protein) could function as a solubility tag. They tested several solubility tags, including MBP and CusF on the green fluorescent protein in *E. coli* and concluded that CusF showed the highest fluorescent activity. CusF have the ability to bind copper ions, and using this ability together with a Western blot, purified recombinant protein can be produced (Cantu-Bustos et al. 2016). Furthermore, CusF showed to be an excellent solubility tag for *in vitro* protein synthesis. It is also a very small protein (9.9 kDa) compared to MBP (~40 kDa) which is a property that is useful since it allows for higher protein yield after tag removal and protein purification (Cantu-Bustos et al. 2016). This is in contrast to large solubility tags which often requires more metabolic energy during production of recombinant proteins than small ones and so produce a lesser yield of the protein (Waugh 2005).

Sterol Carrier Protein-2 (SCP-2) is also a small protein (13.2 kDa in humans) which improves the transfer of lipids between membranes in mammals and other organism. They have a tunnel in which, for example, cholesterol and fatty acids bind to (Edqvist et al. 2004). Moreover, SCP-2 is often expressed as a natural fusion protein in eukaryotic organisms (Blomqvist & Edqvist 2006). This means that it is a natural occurring protein in eukaryotic organisms which gene can merge to another gene, creating one protein. It is also able to be produced in a soluble form at a high concentration in *E. coli* (Kapust & Waugh 1999). Regarding these facts, it would be interesting to study if SCP-2 is able to function as a solubility tag, since it has useful properties. In order to do so, SCP-2 can be fused to insoluble recombinant proteins, for example, green fluorescent protein (GFP) and a form of chloroamphenicol acetyl
transferase (CATΔ9), because they are not able to exist in a soluble form in *E.coli* (Kapust & Waugh 1999).

GFP and CATΔ9 have both been fused to a solubility tag before and CATΔ9 was without success while GFP fused to the solubility tag maltose binding protein and CusF were expressed in a soluble form in *E.coli* (Kapust & Waugh 1999, Cantu-Bustos et al. 2016).

Since no solubility tag is universally applicable and does not possess all useful properties, a frequently used method is to combine a solubility tag with a purification tag and so derive maximum benefit from the tags. The solubility tag enhances the solubility and improves the yield while the purification tag facilitates the proteins purification. The His6-tag is the most common tag used for protein purification, mainly since it binds to immobilized transitions metals such as nickel (Esposito & Chatterjee 2006, Waugh 2005). In this study, His6 will function as a purification tag and SCP-2 will enhance the folding of GFP and CATΔ9 in *E.coli*. The gene fusion created when SCP-2 is being fused to GFP or CATΔ9 will be inserted into a vector called pET-15b. This vector has a hexahistidine sequence (His6) which function as the binding site for the primary antibody used for detecting the purified GFP and CATΔ9 proteins.

So, the aim of this project is to investigate whether *Arabidopsis* SCP-2 is useful as a solubility tag in the *E.coli* strain BL21(DE3). To study this, SCP-2 together with GFP or CATΔ9 will be inserted into a pET-15b vector and expressed in BL21(DE3). Western blot and PageBlue staining are to be used as detection methods for purified GFP and CATΔ9 proteins.

3 Material & methods

3.1 *E.coli* strains

In this report two different *E.coli* strains were used, DH5α and BL21(DE3). DH5α is not a pathogen since it has been developed for laboratory use only. It has multiple mutations which enables high-efficiency transformations. The mutations are: ΔlacZ ΔM15Δ(lacZYA-argF) U169 recA1 endA1 hsdR17(rK-mK+) supE44 thi-1 gyrA96 relA1 (Taylor et al. 1993).

BL21(DE3) is constructed especially for expression of recombinant proteins. The mutations are: lon-11, Δ(ompT-nfrA)885, Δ(galM-ybhJ)884, LAMDE3, DE46, mal+([K-12]), (LamS), hsdS10 (Studier & Moffatt 1986).
3.2 pET-15b vector

The pET-15b vector (Fig.1) was isolated from the *E.coli* strain DH5α which had been grown overnight in tubes containing LB-medium and ampicilin (37 °C at 200 rpm). The isolation was done using a GeneJET™ Plasmid Miniprep Kit from Fermentas. The vector was cleaved with enzymes BamHI (Biolabs 10U/µl) and NdeI (Promega 10U/µl) with an addition of buffer D (Promega, 10x) and water. After cleavage of the vector the sample reactions were incubated at 37 °C followed by 65 °C in order to inactivate the enzymes. One sample reaction underwent a PCR-purification using a GeneJET™ PCR Purification Kit from Fermentas, followed by a gel electrophoresis on both reaction samples.

![Fig. 1. The pET-15b vector with important sites, BamH1 and Nde1(black arrows). Figure from Addgene](https://www.addgene.org/vector-database/2543/ (accessed 18 may 2016))
3.3 Genes

The genes used in this report were extracted from the *E.coli* containing plasmids pGEX SCPAt (SCP-2) (Edqvist et al. 2004) and pMDC123 (CATΔ9) (Carbonell et al. 2014). The third gene used, which encodes GFP, came from isolated psmRS-GFP plasmid DNA (Davis & Vierstra 1996), and was handed to us by J. Edqvist.

3.4 Transformation and DNA manipulation

The plasmids containing genes encoding SCP-2 and CATΔ9 were grown overnight in tubes containing LB-medium and ampicillin (37 °C at 200 rpm) (Thermo Scientific). They were then isolated as described before regarding the pET-15b-vector. All the isolated plasmids containing genes were then amplified by PCR using primers (Thermo Scientific, table 1) and master mix containing 5 µl 10X Dream Taq buffer, 1 µl 10 mM dNTP mix, 0.4 µl Dream Taq DNA polymerase, 0.5 µl of the plasmids and water to a total volume of 50 µl. The following program for the PCR was used: 95 °C s for 3 minutes, a cycle of 95 °C for 30 s, 55 °C for 1 minute and 72 °C for 1 minute. This cycle were repeated 35 times. The last step was in 72 °C for 10 minutes. This was followed by a gel electrophoresis.

*Table 1. List of primers (Thermo Scientific), forward (F) and reverse (R), used for amplifying SCP-2, GFP and CATΔ9.*

<table>
<thead>
<tr>
<th>Gene</th>
<th>Primer sequence</th>
</tr>
</thead>
<tbody>
<tr>
<td>SCP-2 NdeI/EcoRI</td>
<td>5’-ACCAGACATATGATGGCGAATACCCAACTCAATCCGA-3’</td>
</tr>
<tr>
<td>SCP-2 NdeI/HindIII</td>
<td>5’-TAGGACAGAATTCACATTCACTTTTGAGGTCTAGGGAAG-3’</td>
</tr>
<tr>
<td>GFP</td>
<td>5’-CAGTCAAAGGTCCTTTCAACTTTGAGGTCTAGGGAAG-3’</td>
</tr>
<tr>
<td>CATΔ9</td>
<td>5’-GGTAACAAAGCCTAATGAGATGAAAGGAGAGAATCTTT-3’</td>
</tr>
<tr>
<td></td>
<td>5’-GATACACGGATCCTTTACTGTTGTAATCCATGCA-3’</td>
</tr>
</tbody>
</table>

The remaining amount of PCR-product, after the gel electrophoresis, was purified, followed by cloning of SCP-2/EcoRI, SCP-2/HindIII, GFP and CATΔ9 into a pGEM®-T vector according to instructions of pGEM®-T Vector System 1 from Promega.
After cloning, the vector containing SCP-2/EcoRI, SCP-2/HindII, GFP and CATΔ9 was transformed to *E. coli* DH5α. Transformed *E. coli* cells was selected for growth on agar plates containing ampicilin, x-gal and IPTG, also according to pGEM®-T Vector System 1 from Promega.

The vectors with the inserted SCP-2/EcoRI, SCP-2/HindIII, GFP and CATΔ9 were isolated as described before, followed by cleavage of the isolated vector with restriction enzymes (table 2). An additional restriction cleavage was carried out, followed by a PCR. Thereafter ethanol purification was done on both SCP-2 and a gel electrophoresis, before the ligation.

**Table 2. Reagents used for cleavage with restriction enzymes of SCP-2, GFP and CATΔ9.**

<table>
<thead>
<tr>
<th>Gene</th>
<th>Buffer</th>
<th>Enzyme N-terminal</th>
<th>Enzyme C-terminal</th>
</tr>
</thead>
<tbody>
<tr>
<td>SCP-2 EcoRI/pGEM-T</td>
<td>Buffer D</td>
<td>Nde1 (Promega)</td>
<td>EcoRI (Fermentas)</td>
</tr>
<tr>
<td>SCP-2 HindIII/pGEM-T</td>
<td>Buffer 2</td>
<td>Nde1 (Promega)</td>
<td>HindIII (Biolabs)</td>
</tr>
<tr>
<td>GFP/pGEM-T</td>
<td>Buffer EcoRI</td>
<td>BamH1 (Fermentas)</td>
<td>EcoRI (Fermentas)</td>
</tr>
<tr>
<td>CATΔ9/pGEM-T</td>
<td>Buffer 3</td>
<td>BamH1 (Fermentas)</td>
<td>HindIII (Biolabs)</td>
</tr>
</tbody>
</table>

Two separate ligations of the fragments SCP-2, GFP and CATΔ9 into the pET-15b vector were performed according to the T4 DNA Ligase protocol from Thermo Scientific. One ligation contained SCP-2 Nde1/EcoRI and GFP and another ligation contained SCP-2 Nde1/HindIII and CATΔ9.

### 3.5 Protein expression

The ligations were transformed to an *E.coli* strain (BL21(DE3)) which was capable of expressing the vector containing inserted genes. As before, single colonies were inoculated into tubes containing LB-medium and ampicilin and was grown overnight (37 °C at 200 rpm). A small amount of the overnight cultures, one for each ligation, was diluted and transferred to E-flask containing LB-medium and ampicilin. Optical density (OD$_{600}$) of the cultures was then measured followed by incubation for 3 hours (37 °C at 200 rpm) after which optical density was measured again. The two cultures were divided into four E-flasks, two for each ligation, where one was a control containing only the culture, LB-medium and ampicilin and the other which, except for LB-medium, ampicilin and culture, also contained 200 μl IPTG. These four E-flasks were then incubated a second time for 3 hours (37 °C at 200 rpm).
3.5.1 Total cell protein fraction

To obtain total cell protein fraction, a small amount of the four E-flasks were transferred to Eppendorf-tubes and centrifuged for 10 min (4 °C at 10 000 rpm). After centrifugation, the supernatant was discarded and the pellet was re-suspended in 90 μl cold sodium phosphate buffer (pH 7.4, 20 mM) with an addition of 30 μl NuPAGE® LDS Sample Buffer (4x) and 10 μl 0.5 M DDT. The samples were incubated for 10 min at 70 °C and cooled down to room temperature before being loaded onto two protein gels, one for PageBlue staining and one for western blot.

3.5.2 Soluble cytoplasmic protein fraction

To obtain soluble cytoplasmic fraction, the rest of the content of the E-flask was transferred to 50 ml tubes and centrifuged for 10 min (22 °C at 4000 rpm). As described before, the supernatant was discarded and the pellet was re-suspended in 2 ml of cold sodium phosphate buffer (ph 7.4, 20 mM). To lyse the cell membrane, the samples were sonicated using a Banson Digital Sonifier™ with a 10 s interval of ultrasound pulse and a 30 s interval with pause for a total of 1 min. From the sonicated samples, 1.5 ml was transferred to Eppendorf-tubes and centrifuged for 10 min (22 °C at 14 000 rpm). A small amount of the supernatant was transferred to new Eppendorf-tubes in which 30 μl NuPAGE® LDS Sample Buffer (4x) and 10 μl 0.5 M DDT was added. The samples were then incubated for 10 min at 70 °C and cooled down to room temperature before being loaded onto two protein gels, one for PageBlue staining and one for western blot.

3.5.3 Western blot

A Western blot was performed using antibodies to detect proteins of interest. The primary antibody used was Pierce™ 6x His Epitope Tag Antibody and the secondary antibody was Pierce™ Antibody Horseradish Peroxidase, both from Thermo Scientific. The antibodies were diluted 10000:1 with 20 ml PBST and 2 μl antibody. The ladder used, was MagicMark™ XP Western Protein Standard in 20 to 220 kDa from ThermoFisher Scientific.
3.5.4 PageBlue staining

The staining was performed according to instructions from PageBlue™ Protein Staining Solution from Thermo Scientific. The ladder used was PageRuler™ Plus Prestained Protein Ladder in 10 to 250 kDa from ThermoFischer Scientific.

4 Results

4.1 Construction of fusion genes

The pET-15b-vector was isolated from DH5α and to prepare for ligation with genes of interest the pET-15b-vector was cleaved with enzymes BamH1 and Nde1. To verify a proper cleavage of the vector and to remove a smaller fragment created after cleavage, one sample reaction underwent a PCR-purification, followed by a gel electrophoresis on both samples. The results from gel the electrophoresis shows that the purified pET-15b vector (fig. 3. blue arrow) have a clearer band on the gel compared to the non-purified pET-15b vector (green arrow). The size for pET-15b vector on the gel is about 6000 bp as seen from the ladder.

Fig.3. Bands of purified pET-15b vector (blue arrow) and non-purified pET-15b vector (green arrow). Ladder used came from (GeneRuler™ Kb DNA Ladder).
The plasmids containing the genes encoding SCP-2/EcoRI, SCP-2/HindIII, GFP and CATΔ9 were amplified by PCR followed by PCR-purification. To improve the efficiency of the ligation to the pET-15b-vector in later stages, the plasmids containing the genes were cloned into a pGEM-T-vector. After cloning, the vector was transformed onto agar plates (Fig. 4) and single white colonies from the transformation were inoculated into tubes containing LB-medium and ampicillin. The results from the transformation after cloning into pGEM-T vector were successful since white colonies were present on all plates (Fig. 4). However, the quantity of colonies on the plates differed. SCP-2/EcoRI had the most colonies and GFP had the least colonies.

Fig. 4. Agar plates containing ampicillin, x-gal and IPTG with SCP-2/EcoRI, SCP-2/HindIII, GFP and CATΔ9 cloned into a pGEM-T vector. White colonies are present on all plates, but differ in quantity. SCP-2/EcoRI (upper left), SCP-2/HindIII (upper right), GFP (bottom left) and CATΔ9 (bottom right).
To confirm a proper insert of the plasmids into the pGEM-T-vector, a restriction cleavage was carried out, followed by a gel electrophoresis (fig.5). The results from the gel electrophoresis on the genes SCP-2/EcoRI, SCP-2/HindIII, GFP and CATΔ9 demonstrate several genes which have been successfully cloned and cleaved by restriction enzymes (fig.5). These are the ones showing two distinct bands where the upper band is the pGEM-T vector and the lower band is the isolated gene encoding SCP-2, GFP or CATΔ9. The size of the pGEM-T vector on the gel is at about 3000 bp as seen from the ladder. The size of the isolated genes is 400 bp for SCP-2 and 750 bp for GFP and CATΔ9 respectively which can be seen from the same ladder.

To obtain fragments from the pGEM-T-vector of SCP-2/EcoRI, SCP-2/HindIII, GFP and CATΔ9 for the ligation with the pET-15b vector, the pGEM-T-vector was isolated and cleaved with restriction enzyme followed by PCR and ethanol purification on both SCP-2 fragments. To confirm a proper cleavage, a gel electrophoresis was carried out. The isolated and cleaved genes encoding SCP-2/EcoRI, SCP-2/HindIII, GFP and CATΔ9 from the pGEM-T-vector were ligated into the pET-15b-vector. Two ligations were carried out containing SCP-2/GFP and SCP-2/CATΔ9. The ligations were transformed to an E.coli strain (BL21(DE3)) which was capable of expressing the pET-15b-vector with inserted genes. For BL21(DE3) to be able to express recombinant protein, the T7-promotor with a following lac-operator on the pET-15b-vector requires an inducer. In this study Isopropyl β-D-1-thiogalactopyranoside (IPTG) were used as an inducer.
To confirm a complete ligation, a PCR with primers for GFP and CATΔ9 was performed as previously described followed by a gel electrophoresis. The ligation between pET-15b vector and SCP-2/GFP respectively SCP-2/CATΔ9 was shown to be successful as indicated by gel electrophoresis on the PCR products with primers for GFP and CATΔ9 (fig. 6). Two distinct bands are present where the upper band is the pET-15b vector at about 5700 bp and the lower band show the inserted GFP and CATΔ9 at 750 bp as indicated by the ladder (GeneRuler™ Kb DNA Ladder). White arrows indicate the samples chosen for transformation with *E. coli* BL21 (DE3). SCP-2 is not present on the gel as only primers for GFP and CATΔ9 were used in the PCR.

![Fig.6. Bands of SCP-2/GFP and SCP-2/CATΔ9 inserted into pET-15b vector after PCR. The upper bands indicate the pET-15b vector at 5700 bp (blue arrow) and the lower bands indicate the inserts GFP and CATΔ9 at 750 bp (green arrow). White arrows indicate the bands chosen for transformation with BL21 (DE3).](image)

### 4.2 Expression of fusion proteins

To confirm that soluble fusion proteins were expressed from the pET-15b vectors carrying the SCP-2 gene fusions, a Western blot was performed. The results from the Western blot were inconclusive since no light emission was detected with the CCD- camera (Las-4000 Mini, Fujifilm, Tokyo). There were no visible bands showing total cell protein fraction or soluble cytoplasmic protein fraction, the only visible thing on the image was the ladder (fig.7).
To further investigate and identify if the two vectors with inserted SCP-2 gene had expressed functional GFP and CATΔ9 proteins, PageBlue staining was carried out. To confirm that GFP and CATΔ9 had been expressed, some of the samples contained IPTG and some functioned as controls without IPTG. The results from the PageBlue staining showed bands from soluble cytoplasmic protein fraction and total cell protein fraction. All samples except one control from the soluble cytoplasmic fraction are visible (Fig. 8), but the bands however, are not distinct from the control which means that no proteins have been expressed.

Fig. 7. Light emission from Western blot of total cell protein fraction (T) and soluble cytoplasmic protein fraction (S). C=control

Fig. 8. Bands comprised of soluble cytoplasmic protein fraction (S) and total cell protein fraction (T) from the PageBlue staining. C=control.
5 Discussion

The aim of this report was to investigate whether the protein SCP-2 could function as a solubility tag and express functional proteins in *E. coli*. In this case, GFP and CATΔ9 were used as they are not naturally present in the *E. coli* strain used.

To be able to investigate if SCP-2 functions as solubility tag, a vector with inserted SCP-2 fused with GFP or CATΔ9 had to be constructed in order for the proteins to be transformed into the *E. coli* strain BL21(DE3). The vector used was pET-15b. Once inside the *E. coli*, a western blot and PageBlue staining were performed to investigate the expression of the proteins.

The results from the western blot (fig. 7) might indicate that SCP-2 have not been expressed or inserted into the vector properly and therefore not enhanced the folding of GFP and CATΔ9. No light emission from the CCD-camera was detected and the only thing visible from the exposure was the ladder.

The same results were obtained from the PageBlue staining (fig.8) which also imply that no proteins had been expressed. In this case, regarding the soluble cytoplasmic protein fraction and the total cell protein fraction, the gel from the staining does not show any bands from the IPTG samples that are different from the control without IPTG. If SCP-2 had worked as proposed and the proteins had been expressed there would have been bands on the gel not similar to the control. Since there were no distinct bands in either the total cell protein fraction or the soluble cytoplasmic protein fraction it means that no proteins have been expressed. The reason for using IPTG is that it function as an inducer for the T7-promoter in the *E. coli* strain BL21(DE3). Without IPTG the strain are not able to express heterologous proteins such as GFP and CATΔ9 (Chaudhary & Lee 2015).

However, apart from the western blot and PageBlue staining, the results regarding the pET-15b vector from early in the experiment show two bands on the gel. These are of expected size at about 6000 bp indicating a functional vector (Fig.3). The results also indicate that the purified pET-15b vector (fig. 3. blue arrow) shows a clearer band on the gel, which lead to the choice of using the purified vector for the ligation in coming steps. The genes SCP-2, GFP and CATΔ9 which were isolated, amplified and cloned into pGEM-T vector also showed promising results as two distinct bands composed of the pGEM-T vector and the isolated genes were present on the gel at the expected size when compared to the ladder.
This means that the stages up to the point of insertion into the pET-15b vector and the analysis had worked.

The use of a pGEM-T vector improves the efficiency of the ligation with pET-15b vector in later stages by having an overhang of single 3’-terminal thymidine at both ends of the vector. The overhang makes the PCR products more prone for ligation with the vector and also reduces the background of non-recombinant transformants (Zhao et al. 2009).

Moreover, the ligation of the pET-15b vector with SCP-2 fused to GFP or CATΔ9 seems to be successful as well (Fig.6). The results show two distinct bands composed of the pET-15b vector and GFP or CATΔ9 at the expected size when compared to the ladder (Fig.6). The two bands indicated by white arrows (Fig. 6) are the samples which were chosen for transformation to the E.coli strain BL21 (DE3). The reason for choosing these two bands was because they had high DNA concentrations which were obtained from a nanodrop prior to the PCR and gel electrophoresis. Furthermore, even though the ligation seemed successful, SCP-2 is not visible on the gel (Fig.6), because the PCR performed prior to the gel electrophoresis used primers only for GFP and CATΔ9. Therefore, it is not possible to verify that SCP-2 have been inserted into the pET-15b vector correctly. To be certain of a successful ligation, further tests should be performed, for example, additional PCRs where primers for SCP-2 are used as well.

Very few studies, if any at all, have been performed to investigate whether SCP-2 is able to function as a solubility tag in E.coli. Some studies have been made on the same subject but on different proteins. In a former study carried out by Kapust & Waugh (1999), they investigated if maltose binding protein (MBP) could function as a solubility tag. They used GFP and CATΔ9 as recombinant proteins and concluded that MBP was able to help express a functional GFP protein in E.coli. This was not the case for CATΔ9 which were not expressed. When comparing the results obtained from Kapust & Waugh (1999) with the results from this report, they differ as GFP was not expressed. They do come to the same results regarding CATΔ9 which were not expressed in any study.
Moreover, in the introduction, the use of eukaryotic expression systems was mentioned and that they might become preferable over *E.coli* expression system. There is still lacking enough information and systematic studies regarding yeast and insect cells; however, it could be a way of the future. Therefore, other expression systems apart from *E.coli* should be taken under consideration if they improves and thus have a better outcome regarding recombinant protein expression.

In conclusion, because the Western blot and the PageBlue staining did not show any distinct bands, the protein fusions might not have been expressed properly. Therefore it is hard to state if SCP-2 does function as a solubility tag together with GFP and CATΔ9, and should therefore not be excluded. Because of shortage of time during the experiments, further studies should be carried out on SCP-2 with more experiments before rejecting the possibility to use SCP-2 as a solubility tag.

### 5.1 Societal & ethical considerations

The development and the usage of solubility tags could become a solution to overcome the main technical problem when producing biological active recombinant proteins in eukaryotic organism. *E.coli* is a heterologous expression system which provides useful properties for protein production such as the ability to produce large amount of protein in very little time for a low cost. *E.coli* was the first host used to produce recombinant insulin and a lot of progress has been made regarding the translation, transcription and the secretion of proteins produced in *E.coli*. However, inclusion bodies are still a major problem when producing recombinant protein and can in some cases lead to genetic diseases (Kapust & Waugh 1999, Gundersen 2010). Therefore, this problem needs to be avoided and using a solubility tag is preferable. Since solubility tags are not universally applicable the need for discovering new tags are substantial.

In this study the *E.coli* strain used were genetically modified and is therefore regarded as an ethical consideration. This bacteria however, was only used in the laboratory and dealt with carefully. Since we did not make the bacteria resistant for antibiotics it should not pose a threat of some kind.
6 Acknowledgement

I am grateful to Isak Mikkola for co-working with me on this project, and

to Johan Edqvist for his knowledge, guidance and help before, during and

after the experiments.

7 References

Bell MR, Engleka MJ, Malik A, Strickler JE (2013) To fuse or not to
fuse: what is your purpose? Protein science 22, 1466-1477
http://dx.doi.org/10.1002/pro.2356

Cantu-Bustos JE, Vargas-Cortez T, Morones-Ramirez JR, Balderas-
Renteria I, Galbraith DW, McEvoy MM, Zarate X (2016) Expression and
purification of recombinant proteins in *Escherichia coli* tagged with the
metal-binding protein CusF. Science Direct 121, 61-65
http://dx.doi.org/10.1016/j.pep.2016.01.007

Carbonell A, Takeda A, Fahlgren N, Johnson SC, Cuperus JT, Carrington
JC (2014) New generation of artificial MicroRNA and synthetic trans-
acting small interfering RNA vectors for efficient gene silencing in
*Arabidopsis*. Plant Physiology 165, 15-29
http://dx.doi.org/10.1104/pp.113.234989

Chaudhary AK, Lee EY (2015) Tightly regulated and high level
expression vector construction for *Escherichia coli* BL21 (DE3). Journal
of Industrial and Engineering Chemistry 31, 367-373
http://dx.doi.org/10.1016/j.jiec.2015.07.011

protein solubility, purification and immunogenicity in *Escherichia coli*:
the novel Fh8 system. Frontiers in Microbiology 5, 63
http://dx.doi.org/10.3389/fmicb.2014.00063

Dalton AC, Barton WA (2014) Over-expression of secreted proteins from
mammalian cell lines. Protein Science 23, 517-525
http://dx.doi.org/10.1002/pro.2439

Davis SJ, Vierstra RD (1996) Soluble derivatives of green fluorescent
protein (GFP) for use in *Arabidopsis thaliana*. Weeds World 3, 43-48

Edqvist J & Blomqvist K (2006) Fusion and fission, the evolution of
http://dx.doi.org/10.1007/s00239-005-0086-3
http://dx.doi.org/10.1074/jbc.M405099200

http://dx.doi.org/10.1016/j.copbio.2006.06.003

http://dx.doi.org/10.1111/j.1600-0404.2010.01382.x

Kapust RB, Waugh DS (1999) Escherichia coli maltose-binding protein is uncommonly effective at promoting the solubility of polypeptides to which it is fused. Protein Science 8, 1668-1674
http://dx.doi.org/10.1110/ps.8.8.1668

Studier FW, Moffatt BA (1986) Use of bacteriophage T7 RNA polymerase to direct selective high-level expression of cloned genes. Journal of Molecular Biology 189, 113-130
http://dx.doi.org/10.1016/0022-2836(86)90385-2

http://dx.doi.org/10.1016/S0958-1669(00)00199-3

Taylor RG, Walker DC, McInnes RR (1993) E.coli host strains significantly affect the quality of small scale plasmid DNA preparations used for sequencing. Nucleic Acid Research 11, 1677-1678


http://dx.doi.org/10.1016/j.tibtech.2005.03.012