Overexpression of the plant aquaporin *At*SIP2;1 in *Escherichia coli*

Hiep Nguyen

Supervisor: Urban Johanson

Version 2: 2020/06/02

A B S T R A C T

In molecular biology and biochemistry, it is key to have large amounts of the isolated protein at hand when performing functional and structural studies. Unfortunately, the levels of an individual membrane protein in native tissues is often exceptionally low, making it challenging to purify a sufficient amount. A relatively simple and inexpensive way to get large amounts of a protein is to overexpress in *Escherichia coli* and then extracted it. However, this heterologous expression has also many problems as well as challenges. In this project I employed *E. coli* in an attempt to express, the plant aquaporin *At*Sip2;1, which is a water channel normally residing in the endoplasmic reticulum of *Arabidopsis thaliana*. Due to the limited time of the project I have not been able to test if the protein can be expressed in the *E. coli* expression system.

Key words: Aquaporin, AtSIP2;1, heterologous overexpression, protein purification, membrane protein, expression system, *Escherichia coli*.

Introduction

In molecular biology and biochemistry, it is important to have sufficient quantities of protein when performing functional and structural studies. One way to get protein is to purify it from natural sources, directly from their native tissues. However, when it comes to membrane proteins, they are neither abundantly expressed nor found in large amounts in native tissues, which why it is generally not advantageous to purify them from natural sources. (Nordén et al. 2011; Yin et al. 2007). Furthermore, the use of protein from natural sources cannot provide an opportunity of genetically modifying proteins to facilitate their extraction and detection (Schlegel et al. 2010). Therefore, membrane proteins are often obtained by overexpression in other cells / organisms and subsequent purification of the protein. (Wagner et al. 2007).

There are several different expression systems used to express a protein such as the bacteria, yeast, insect, or mammalian expression systems (Yin et al, 2007). Which expression system to use depends on which protein you want to express, how much time you have to express, how easy the expression system should be or type of post-translational modifications and several factor (Yin et al. 2007). Expression of heterologous protein in micro-organism has been a major advantage in biotechnology. Using these systems, proteins can be produced from relatively simple and inexpensive culture media compared to other more expensive expression and extraction methods (Yin et all. 2007). The bacterium *Escherichia coli* is one of the most popular vehicles for overexpression of membrane proteins (Wagner et al, 2007; Drew et al, 2002). The main advantages of *E. coli* expression system are the low cost, that it is relatively simple to manage, no posttranslational modifications, and a short generation time as well as short delay of expression (Sarramegna et al. 2007).

However, the *E. coli* expression system is not always best, in addition to many benefits, this expression has many problems as well as challenges. *E. coli* membrane composition and plasma environment is different from eukaryotic cells, and this can affect production and the correct folding of membrane protein. In addition, the expression of membrane protein can reduce the cell growth, in some cases, it can even be toxic to host cells (Sarramegna et al. 2007).

In this project, AtSIP2;1, a member of the smallest and least studied plant aquaporin sub-family, was targeted in an attempt to express it in *E. coli*. Aquaporins are water channels that are located in the membranes of the cells,

where in they facilitated the transport of water and several other small molecules, including glycerol, urea, ammonia, and CO2 across the membrane. In plants, aquaporins play an important role in plant growth and water transport (Ishikawa et al. 2005).

Materials and methods

Host strains, plasmids, and culture medium. The different bacterial strains and plasmids used in this experiment are listed in Table 2. The gene encoding *AtSIP*2;1s was provide in the *pXβG-ev1* plasmid, cloned in a *BglII* site and found in the *E. coli* strain *LS20* (030127 *XL-1 Blue MRF*). The *pET23c* vector purified from the *E. coli* strain *BL21*(*DE3*) *pLysS* was chosen as cloning vector. *E. coli XL-1 Blue* was used as a cloning strain, while *BL21(*DE3) *pLysS* was used to expression of protein. *XL-1Blue* competent cells used at transformation were from Agilent Technologies. For expression inhouse prepared competent cells *E. coli* strain *BL21 (DE3)* was used at the transformation. All cultures of *E. coli* strains were grown overnight, in LB medium with *ampicillin* (100 µg/ml) in 50 ml falcon tube at 37ºC in a shaker at 200 -250 rpm. For agar plates, 1.5% agar was added to the medium.

Standard molecular biological techniques All *Phusion High-Fidelity-*/ *DreamTaq DNA polymerase*, restriction enzymes and *T4 DNA ligase* were from Thermo Fisher Scientific and used according to the manufacturer's recommendation. Plasmid purification done with Thermo Fisher Scientific *Plasmid DNA Miniprep Kits* or *QIAGEN Plasmid Midi Kits* and used as recommended by the manufacturer.

Primers design Oligonucleotide primers in Table 2 were designed using the Primer3Plus website (https://primer3plus.com/cgibin/dev/primer3plus.cgi). The primers were chosen with a length between 15-22 bp, so that the melting temperature (Tm) is 52ºC - 65ºC. The primers annealing temperature (Ta) were similar for both primers (5ºC difference) and hairpins, cross-dimers were avoided. The start codon was retained on forward primer and a BamHI restriction site was added to the 5′ end. In the reverse primer, a HindIII restriction site was added and the stop codon omitted to allow in frame fusion with the vector encoded His6-tag to the C-terminus. The restriction enzyme sites for BamHI and HindIII were used to create sticky ends of the amplified AtSIP2;1 gene which would allow an efficient ligation into the plasmid vector pET23c.

Amplification of *AtSIP2;1* gene *pXβG-ev1* plasmid with *AtSIP2;1* gene from strain *LS20* (030127) *XL-1 Blue MRF* was prepared using *Thermo Fisher Scientific Plasmid DNA Miniprep Kits*. The plasmid was used as a template to amplify the coding sequence of the *AtSIP2;1* gene using the *Phusion High-Fidelity DNA polymerase* together with the forward and reverse primers. In this experiment a generic Touchdown PCR program is used. Gradually decreasing the annealing temperature until amplification eventually occur. Different temperature parameters are described in Table 1.

Table 1: Touchdown PCR program.

Phase A	**Step**	**Temperature**	**Time**
1	Denature	94ºC	4 min
2	Denature	94ºC	1 min
3	Anneal	64ºC (-0.5ºC each step) *	1 min
4	Elongate	72ºC	30 s**
Repeat steps 2-4 (29 times)			
Phase B	**Step**	**Temperature**	**Time**
5	Denature	94ºC	1 min
6	Anneal	50ºC	1 min
7	Elongate	72ºC	30 s
Repeat steps 5-7 (9 times)			
Phase C	**Step**	**Temperature**	**Time**
8	Elongate	72ºC	4 min
9	Storage	4ºC	forever

*Every time the cycle is repeated, the annealing temperature decreased 0.5 ºC/cycle.

**An elongation (synthesis) step of 30 s allows Phusion DNA polymerase to synthesizes 1 kb.

Preparation of cloning vector The PCR products were cloned into the *pET23c* plasmid between the *BamHI* and *HindIII* sites, resulting in an extension with the vector encoded the His6-tag to the C-terminus of the protein sequence, and then transformed into the *E. coli XL 1-Blue* cells. The cloning plasmid *pET23c* was extracted from an overnight culture with the *QIAGEN Plasmid Midi Kit*, and then cleaved with restriction enzyme (*Ban*HI and *Hind*III).

Ligation *of AtSIP2;1* into *pET23c* Enzyme T4 DNA ligase was used to ligate PCR fragments into plasmid *pET23c* between restriction site *Bam*HI and *Hin*dIII to construct the plasmid pET23c-AtSIP2;1. Different rations/proportion (1:1; 3:1; 5:1 and 0:1) of PCR fragment and plasmid were used at ligation to be able to compare results between different plates and test how good restriction enzyme work to ensure the results.

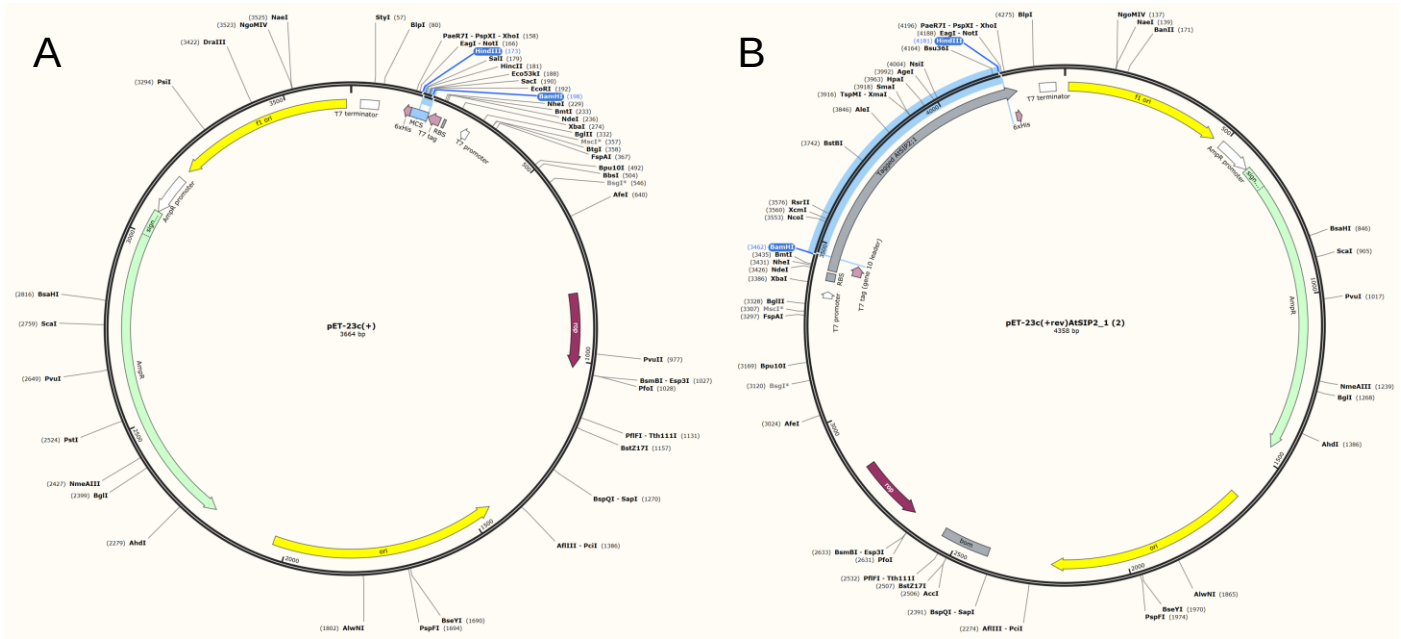

Figure 1: Plasmid maps. A. pET23c, B. pET23c--AtSIP2;1. The AtSIP2;1 gene (light blue) has inserted into pET23c plasmid between the *Ban*HI and *Hin*dIII restriction sites (highlighted inblue), both plasmids have a T7 promoter, His6-tag, and ampicillin resistance. In B the sequence is reversed and complemented relative A. The plasmid maps are created with SnapGene.

Table 2: Bacteria, plasmids and primers used in the project.

Strains /primers	Properties/sequence
E. coli LS20 XL-1 Blue MRF	Contain *pXβG-ev1-AtSIP2.1 plasmid*
E. coli XL1-Blue	Cloning strain
E. coli BL21 (DE3)	Contain *pET23c* plasmid and *pLysS* plasmid Expression strain
E. coli BL21(DE3)	Expression strain
AtSIP2;1 Forward	**5´**-GATGGATCCGTATGGGACGAATCGGTTTAGTAG-**3´**
AtSIP2;1 Reverse	**5´**-ATCAAGCTTCTCAGATTTGGCCTTAGGTTTC-**3´**

Amount of insert PCR product and vector DNA can be calculated with the equation:

$$Insert\ mass = ratio \cdot \frac{insert\ length\ (bp)}{vector\ length\ (bp)} \cdot vector\ mass\ (ng)$$

Their ratio is molar ratios of insert/ vector; vector length is 3663 bp, insert length is 714 bp and the vector mass is 100 ng.

2 µl (2 units) of T4 DNA ligation enzyme was used in all reaction. After addition, all the reaction was incubated at room temperature for 1 to 2 hours.

Transformation of bacteria Following an incubation on ice for 1 hour, the AtSIP2;1-pET23c plasmid was transformed into *E. coli XL-1 Blue* by a heat shock, induced by an incubation for 45 seconds in a 42ºC water bath. Six LB agar-plates containing ampicillin 100 µg/ml were prepared for growing of transformant bacteria. Three of them were 1:1; 3:1 and 5:1 ratios reaction; a religation control (0:1) with only cut pET23c plasmid to control if the T4 DNA ligase could close the plasmid, i.e. if only one restriction enzyme had cut efficiently; a positive control with only uncut plasmid to be able to monitor the transformation efficiency; and a negative control with only water to control sterility at work. All agar plates were incubated overnight at 37ºC.

Screening for insert with PCR 50 colonies in overnight culture from the 5:1 ratio divide into 10 pools to ran PCR to check the correct insertion of AtSIP2;1. The pools were heated at 94ºC, for 20 minutes to lyse the bacteria and then used as template in PCR as described above. Compared to the PCR in the cloning step the total volume of reaction was half (25µl instead of 50µl) and *DreamTaq DNA polymerase* use instead of *Phusion High-Fidelity DNA polymerase.* The colonies that possible contain the correct insert were grown overnight in LB medium with ampicillin.

Plasmid restriction cleavage analysis The plasmids were extracted from the overnight culture with Thermo Fisher Scientific *Plasmid DNA Miniprep Kit*, and then cleaved with restriction enzymes (*Bam*HI and *Hin*dIII) to verify the correct insertion.

Results and discussion

Purification and restriction enzyme control of *pXβG-ev1-AtSIP2.1* plasmid The purification of plasmid *pXβG-ev1-AtSIP2;1* was successful (Nano-drop test 49.2 ng/µl and 13.6 ng/µl). The plasmid was extracted from two overnight cultures of *E. coli LS20*, and then cleaved with *SmaI* and *BglII* restriction enzyme to confirm that is was the correct plasmid. Agarose gel electrophoresis image (fig. 2) shows the correct plasmid construction on the lane 5, with three strong bands, where of a two form a double band located approximately between 4000, and 3500 bp, and one corresponding to a much smaller fragment of 750 bp. The largest band is an indication of partially cut plasmid, the two smaller bands corresponds to the vector *pXPG-ev1* and the *AtSIP2;1* gene. The size of them corresponds well expected fragments where linearized *pXPG-ev1-AtSIP2;1* plasmid has a size of 3919 bp, *Bgl*II cleaved *pXPG-ev1-AtSIP2;1* plasmid has a size of 3205 and *AtSIP2;1* gene has a size of 714 bp. *Sma*I cuts once within the gene and once in the vector backbone, resulting in the fragments of 463 and 3456 bp, which are seen in lane 4.

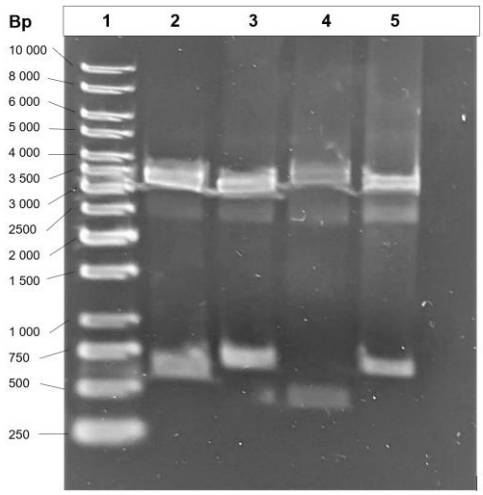

Figure 2: Purification and restriction enzyme control of *pXβG-ev1-AtSIP2;1* plasmid. Agarose gel electrophoresis picture on *pXβG-ev1-AtSIP2.1* plasmid. Lane 1 shown the DNA ladder, lanes 2 and 3 shown the sample from another student; lane 4 and 5 shown the *pXβG-ev1-AtSIP2.1* plasmid sample treated with *Sma*I and *Bgl*II, respectively.

Purification and restriction enzyme control of *pET23c* plasmid Results from agarose gel electrophoresis (fig. 3) indicated the extraction of pET23c plasmid from *E. coli BL21* (*DE3*) were possible successful. The undigested plasmid reaction shows four different bands which agrees with what to expect. This strain of *E. coli BL21* (*DE3*) contains two different plasmids (*pET23c* plasmid, 3664 bp and *pLysS* plasmid, 4886 bp) so on the gel the purified plasmid preparation will result in four different bands (two relaxed and two supercoiled plasmids). Single digested plasmid (lane 3 and 4) migrated slightly slower than double digested plasmid (lane 5) which do not really correspond to the expectation since the plasmid cleaved with two restriction enzymes is not significantly smaller than single cleaved plasmid (only 25 bp difference). However, the *pLysS* plasmid has a single *Hin*dIII site and two *Bam*HI sites, hence the former should result in a fragment of 4886 bp, and the latter in fragments of 4244 and 642 bp. The double digest would cut off 346 bp from the 4244 bp fragment, shortening it to 3898 bp. Presumably, the larger fragments are not resolved from each other in the electrophoresis and the smaller are too weak to be detected on the gel. At the same time, the chance of relaxed plasmids plasmid is greater when cleaved with 2 restriction enzymes than with only one. The concentration of plasmid in reactions (Nano-drop test average of different samples is around 10 ng/µl) is too low and the DNA ladder is too weak, leading to unclear gel image which makes it difficult to interpret results.

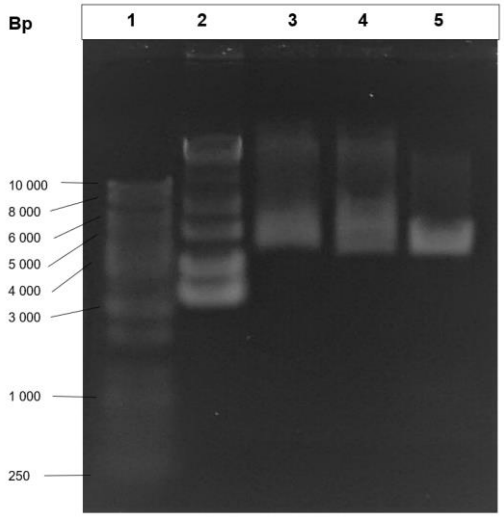

Figure 3: Purification and restriction enzyme control of *pET23c* plasmid. Agarose gel picture on *pET23c* plasmid cleaved with restriction different enzyme. Lane 1 shown the DNA ladder, lanes 2 shown the uncut plasmid which purifier from *E. coli BL21 (DE3*); 3 and 4 containing restriction digestion by simple restriction enzyme, *Bam*HI and *Hin*dIII, respectively. Lane 5 shown the which dubble digest by *Bam*HI and *Hin*dIII.

PCR product analysis Results from agarose gel electrophoresis (fig. 4) shown the amplification of *AtSIP2;1* gene from plasmid *pXβG-ev1-AtSIP2;1* in *E. coli LS20* was successfully. The gel picture shows that the PCR product has an estimated size of around 750 bp, which corresponds well to the *AtSIP2;1* gene size of 714 bp. Besides that, there were no bands present in the negative control, which indicates the PCR reaction were not contaminated with impurities.

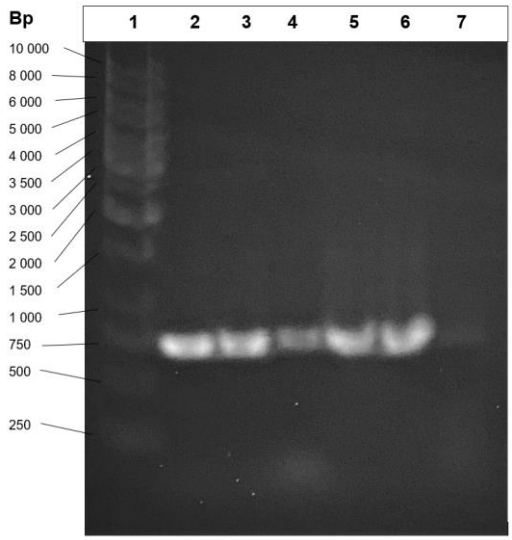

Figure 4: Agarose gel electrophoresis of the PCR product. Lane 1 shown the DNA ladder, lanes 2, 3 and 4 containing PCR samples from different another groups. Lane 5 and 6 containing the PCR product of AtSIP2,1, while the lane 7 containing the PCR negative control samples.

Detection of *pET23c-AtSIP2;1* plasmid by PCR and restriction enzyme cleave The transformation of *E. coli XL 1-Blue* as such was successful. But on account of the high number of colonies may likely to contain false positive colonies without insert it is difficult to find the correct insert plasmid. This result may be due to poor restriction enzyme activity, that have not cleaved the *pET23c* plasmid at both sites so the *AtSIP2;1* gene fragment cannot ligate into plasmid and the uncut or religated plasmid are transformed into the *E. coli* and give high background with colonies carrying the empty vector. Colony PCR together with restriction enzymes control was therefore used to facilitate the finding of clones with the correct insert *AtSIP2;1* gene. Colony PCR indicated possible correct insert in the reaction on lanes 2, 3, 4,5, 7, 9 and 11 in the agarose gel (fig. 5). The blurry DNA ladder makes it difficult to ensure the size of the PCR product. However, the positive control in lane 1, with the *pXβG-ev1-AtSIP2;1* template generates a product of similar size, so the result appears quite reliable.

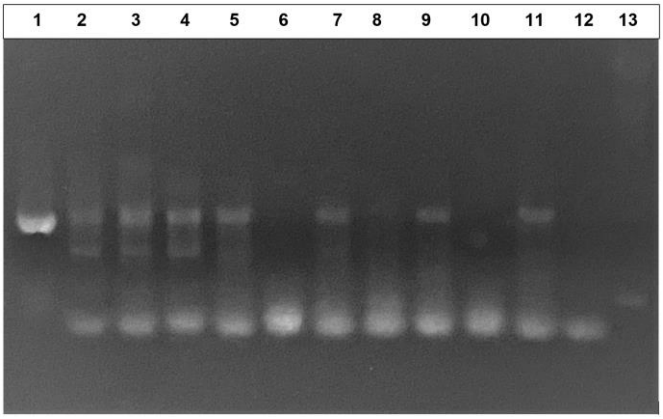

Figure 5: Detection of *pET23c-AtSIP2;1* plasmid by PCR. Agarose gel electrophoresis of the colony PCR product. Lane 13 shown the DNA ladder, lanes 12 shown the negative control, lane 1 is a positive control with *pXβG-ev1-AtSIP2;1* plasmid as template, lane 2-11containing from 50 different colonies collected in 10 pools.

Verification of *pET23c-AtSIP2;1* plasmid by restriction enzyme Results from agarose gel electrophoresis (fig. 6) and Nano-drop test (around 10 ng/µl) show that the plasmid extraction from four overnight cultures of four potential positive clones was not so successful. Hence, it is difficult to evaluate the results. From the gel image, colony in the lane 9 may contain correct insert of *AtSIP2;1* gene in plasmid *pET23c*. But because of the low concentration it is not possible to conclude that it is correct. On a side note, despite the smeary appearance in the upper part, the overloaded control digestion of the plasmid preparation of *pET23c* containing *pLysS* in lane 2 display bands that are likely to correspond to the expected fragments of 642 and 346 bp in the lower part.

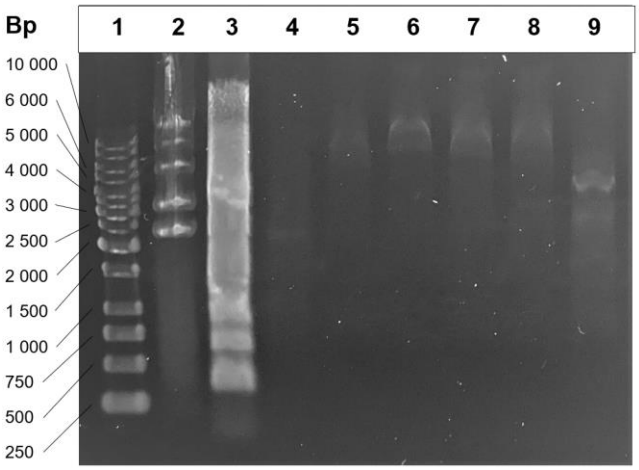

Figure 6: Detection of *pET23c-AtSIP2;1* plasmid by restriction enzyme. The gel picture of plasmids purified from four overnight cultures. Lane 1 shown the DNA ladder, lane 2 contain uncut *pET23c*, lane 3 *pET23c* digested by *Ban*HI and *Hin*dIII, lane 4 and 5 contain samples from another student, lane 6 – 9 contain plasmid purified from four overnight cultures.

Conclusion and Future work

Due to the limited time of the project I have not been able to examine if the plant aquaporin AtSIP2;1 can be efficiently expressed in *E. coli*. But I may have succeeded in subcloning the gene into the expression vector. In the future we will continue with the project that we have planned. We will run new colony PCRs and purify plasmid, following cleavage with restrictions enzyme to show that we have received the correct insert. Further, we will induce the expression of the protein and detect by any production by western blotting using an antibody against the His-tag. If this is successful it will open up for structural and functional studies of this largely uncharacterized membrane protein.

Reference

David Drew, Linda Fröderberg, Louise Baars, Jan-Willem L. de Gier: *Assembly and overexpression of membrane proteins in Escherichia coli*. Biochimica et Biophysica Acta 1610 (2003) 3–10

Fumiyoshi Ishikawa, Shinobu Suga, Tomohiro Uemura, Masa H. Sato, Masayoshi Maeshima: *Novel type aquaporin SIPs are mainly localized to the ER membrane and show cell-specific expression in Arabidopsis thaliana*. FEBS Letters 579 (2005) 5814– 5820

Kristina Nordén, Maria Agemark, Jonas ÅH Danielson, Erik Alexandersson, Per Kjellbom and Urban Johanson: *Increasing gene dosage greatly enhances recombinant expression of aquaporins in Pichia pastoris*. BMC Biotechnology (2011) 11:47

V. Sarramegna, F.Talmont, P. Demange and A. Milon: *Heterologous expression of G-protein-coupled receptors: comparison of expression systems from the standpoint of large-scale production and purification*. Cell. Mol. Life Sci. 60 (2003) 1529– 1546

Susan Schlegel, Mirjam Klepsch, Dimitra Gialama, David Wickström, Dirk Jan Slotboom2 and Jan-Willem de Gier: *Revolutionizing membrane protein overexpression in bacteria*. Microbial Biotechnology (2010) 3(4), 403– 411

Jiechao Yin, Guangxing Li, Xiaofeng Rena, Georg Herrler: *Select what you need: A comparative evaluation of the advantages and limitations of frequently used expression systems for foreign genes*. J Biotechnol (2007), 127(3):335-347

Samuel Wagner, Louise Baars, A. Jimmy Ytterberg, Anja Klussmeier, Claudia S. Wagner, Olof Nord, Per-Åke Nygren, Klaas J. van Wijk, and Jan-Willem de Gier: *Consequences of membrane protein over-expression in Escherichia coli*. Mol Cell Proteomics (2007) 6(9) 1527-1550