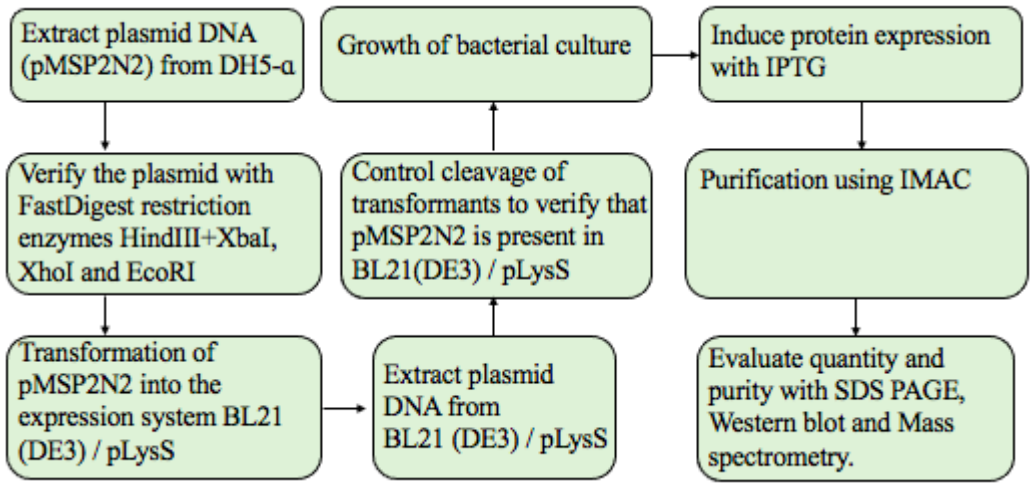
Overexpression and purification of MSP2N2 for nanodisc assembly and reconstitution of TRPA1

Isabelle Bäckström, is0802ba-s@student.lu.se
Molecular biology, MOBY01, Spring 2021
Supervisor: Urban Johanson

1. Abstract

Membrane proteins are important building blocks for several biochemical processes. The difficulty with membrane proteins is that they tend to be lowly expressed and can be difficult to manage. Therefore, membrane proteins are studied to a lesser extent than soluble proteins. Since membrane proteins are intrinsically hydrophobic, their purification and structural determination critically rely on solubilization by detergents or other alternative procedures, such as nanodiscs. The purpose of this project is to overexpress and purify a membrane scaffold protein (MSP2N2) to allow assembly of nanodiscs, which in turn can help determine the structure of a membrane protein, namely a Transient Receptor Potential cation channel A1 (TRPA1) from *Hylobius abietis*. The initial results presented in this study shows the difficulty of producing MSP2N2 in terms of contaminants and degradation. These findings provide development opportunities for how to successfully isolate MSP2N2 for further use.



Graphical abstract showing an overview of the experiments performed to overexpress and purify MSP2N2.

1.1 Introduction

The importance of membrane proteins is extensive for all forms of life as they play a key-role in the cell's energy metabolism, but also in signaling and communication with the surrounding environment (Luckey, 2014). The problem with membrane proteins is that they are difficult to handle, they are very unstable when moved from their native environment and often result in low protein yields. This is why membrane proteins are studied less than soluble proteins and why the number of known structures of membrane proteins is still relatively small. The protein structure database is dominated by structures solved by X-ray, but one problem with X-ray technology is that it requires the protein to be crystallized, which can be especially difficult to accomplish for membrane proteins. Therefore, alternative methods such as Cryogenic Electron Microscopy (Cryo EM), which does not require crystallization and in addition has proven to be good for large membrane proteins, may provide an attractive option (Birch et al., 2020).

One system that has dominated the structural determination and functional study of membrane proteins includes detergent micelles and mixed detergent/lipid micelles that creates an environment similar to the natural lipid bilayer (Ritchie et al., 2009). The use of detergents has been a fairly effective way of solubilizing membrane proteins, but there are some crucial limitations. When detergents are used, there is a risk that protein-protein interactions as well as protein-lipid interactions are lost, which can destabilize the membrane protein. Some lipids are also crucial for the function of the membrane protein, and if these are removed the protein can be inactivated. To try to solve some of the problems that arise from the use of detergents, another system has been developed, namely nanodiscs based on membrane scaffold proteins (MSPs) (Birch et al., 2020). MSP nanodiscs are derived from high-density lipoproteins (HDLs) surrounded by two belts of Apolipoprotein A-1 (Apo-A1), which can be customized to create disks of various sized diameters and have different epitope tags to help with different purification strategies (Ritchie et al., 2009).

One objective of the current research on membrane proteins at the division of Biochemistry and Structural Biology at Lund University, is to determine the structure of a relatively large integral membrane protein called Transient Receptor Potential cation channel A1 (TRPA1) from *Hylobius abietis*. This is important to be able to develop repellents against the insect that cause damage to spruce plants as they eat the bark of the plants, which could be an alternative to toxic sprays and a complement to the physical protections of, for example, wax used today. In addition, it is interesting from a basic research perspective to compare how different variants of TRPA1 work as it could be important for the development of future drugs. The structure of the orthologous human TRPA1 (hTRPA1) has previously been determined with cryo EM, using nanodiscs for solubilization (Suo et al., 2020). Therefore, we would like to evaluate whether this could be a successful strategy also regarding the *H.abietis* orthologue (HaTRPA1).

To enable reconstitution trials of HaTRPA1 in nanodiscs, an isolated MSP is a crucial constituent in addition to the lipids which are purchased from commercial suppliers. A particular variant, MSP2N2, has been used in the structural determination of hTRPA (Grinkova et al., 2010; Suo et al., 2020). The purpose of my bachelor project is to be a part of the process by overexpressing and purifying MSP2N2 to allow assembly of nanodiscs and the reconstitution of HaTRPA1 provided by a PhD student in the research group, and also to participate in structural determination of HaTRPA1. However, the latter may not be realistic from a time perspective so the main focus of my study will be to overexpress and purify MSP2N2 for further use.

2. Material and methods (For detailed information, see Appendix 1.)

2.1 Plasmid preparation for purification of plasmid DNA and transformation of pMSP2N2 from DH5-α to the expression system BL21 (DE3) / pLysS.

The plasmid is delivered in DH5-α cells, plasmid preparation and transformation is needed to transfer the plasmid to the expression system BL21 (DE3) / pLysS. Two falcon tubes containing Luria Broth (LB) and Kanamycin was inoculated with the same colony taken from a Luria agar (LA) plate consisting of an *Escherichia coli (E-coli)* glycerol stock with the plasmid pMSP2N2 to make starter cultures. Plasmid DNA was extracted from the starting cultures using the "Plasmid miniprep kit" protocol associated with GeneJET Plasmid Miniprep Kit from Thermo Fisher Scientific (Sweden). DNA concentration was checked using nanodrop (see table 10). To verify that the plasmid was intact, three FastDigest restriction enzyme digests (HindIII+XbaI, XhoI and EcoRI) were prepared in accordance with the manufacturer's recommendations (ThermoFisher Scientific, Sweden). The three digestions as well as the uncleaved plasmid were checked on a gel (see figure 1).

The transformation of pMSP2N2 into the expression system BL21(DE3) / pLysS was done in accordance with standard protocols from the research group. The transformation was verified using 6 LA plates containing Kanamycin. Two new starting cultures were prepared with colonies from the transformation. The plasmid DNA was extracted using the "Plasmid miniprep kit" protocol associated with GeneJET Plasmid Miniprep Kit from Thermo Fisher Scientific (Sweden) and the concentration was checked using nanodrop (see table 12). Control cleavage of transformants using FastDigest EcoRI and FastDigest XhoI was performed in accordance with the manufacturer's recommendations (ThermoFisher Scientific, Sweden) and checked on a gel to verify that pMSP2N2 was present in BL21(DE3) / pLysS (see figure 2).

2.2 MSP2N2 expression

Growth of cells is done in a fermenter. Four hours from start, optical density (OD) is inspected, then the OD is checked every hour (see table 13). For each sample taken to check OD, 1mL is saved to run a western blot. After 4h and 10min, protein expression was induced by the addition of IPTG. After another two hours, the cells were harvested and the generated cell pellet was weighted and stored. Western blot was performed on samples from the fermenter in accordance with standard protocols in the research group (see figure 3). A Mini-Protean TGX Precast gel 10%, 10-well comb, 30 µl/well (Bio-Rad, sweden) was used, as well as the Trans-Blot Turbo Transfer system (Bio-Rad, Sweden). The antibodies used were 6xHis monoclonal clontech 631212 from mouse and Anti-mouse lgG (whole molec.) Alkaline phosphatase Sigma-Aldrich A9316 from goat.

2.3 MSP2N2 purification with Immobilized metal affinity chromatography (IMAC) (Attempt 1)

Purification of MSP2N2 is done with Immobilized metal affinity chromatography (IMAC) using a nickel column. The eluate was collected in fractions (E0-E9) and protein concentration in each fraction was checked using nanodrop (see table 14). Flow through of the lysate, flow through of the third wash and flow through of eluate 0-5 from the purification was used to charge to a gel in accordance with standard protocols in the research group, which was then stained with coomassie (see figure 4). The gel used was a Mini-Protean TGX Precast gel 10%, 10 well comb, 30µl/well (Bio-Rad, Sweden). Size exclusion chromatography (SEC) was performed on E1 and E2 in accordance with standard protocols in the research group using the NGC 10 Medium Pressure Chromatography system (Bio-Rad, Sweden). The column used was a SuperDex 200 10/300GL (GE healthcare, Sweden). E1 and E2 were pooled to be concentrated and filtered before the SEC, which was done using an Amicon Ultra-4 Centrifugal filter Unit (30 kDa) (Sigma-Aldrich, Sweden) and a 1ml syringe with a 0.45 µm filter. After concentration and filtration, protein concentration was checked using nanodrop (see table 15). The sample loaded to the column (E1+E2) was of 100 µl.

2.4 MSP2N2 purification trial number 2

Purification of MSP2N2 is done as described above, but in comparison with trial 1, the purification was performed with 5 times as much lysate and the lysate is allowed to bind overnight instead. Protein concentration of the eluate can be seen in table 16. A western blot and a gel stained with coomassie was done in accordance with standard protocols in the research group on the flow through of the lysate, the third wash and E1-E3 (see figure 5 and 6). Of E1 and E2, two fractions of each were made where one fraction of each was boiled before being charged to the wells. Gels used were Mini-Protean TGX Precast gels 10%, 10-well comb, 30 µl/well (Bio-Rad, Sweden).

E1 and E2 from the second purification were pooled and concentrated using “Vivaspin 15R centrifugal concentrator” (Sigma-Aldrich) and “Spin-X 0.45 µm Costar1" (Sigma-Aldrich, Sweden). Protein concentration on what went through the Vivaspin filter versus what did not go through the filter was checked using nanodrop (see table 17). What did not pass through the filter (E1+E2) was spin filtered using “Spin-X 0.45 µm Costar1. “Spin-X 0.45 µm Costar1” was used twice and what went through the filter both times was pooled and protein concentration was checked using nanodrop (see table 18). What did not pass through the Spin-X filter during the first attempt was lost, but what did not pass through the second time was collected and protein concentration was checked using nanodrop (see table 18) and 20 µl was saved and stored. SEC was conducted as described above on the sample containing flow through from the Spin-X (100 µl).

A western blot as well as a gel stained with coomassie was performed in accordance with standard protocols in the research group on flow through of the lysate, E0 and E3 from the elution, flow through from the concentration with Vivaspin, flow through from the Spin-X column, the sample that did not go through the Spin-X filter and fraction 29 and 30 from the SEC (see figure 7 and 8). The gel used was a Mini-Protean TGX Precast gel 10%, 12 well comb, 20µl/well (Bio-Rad, Sweden).

For the sake of comparison, a western blot was performed in accordance with standard protocols in the research group on samples before induction with IPTG, after induction with IPTG, cell lysate (after cultivation in the fermenter) and E1+E2 from the second purification (what did not go through the Spin-X filter) (see figure 9). The gel used was a Mini-Protean TGX Precast gel 10%, 10-well comb, 30µl/well (Bio-Rad).

Three bands (see figure 7) are sent away for mass spectrometry performed by Katja Bernfur (in house service) (Appendix 2).

3. Results

Plasmid DNA of MSP2N2 in DH5-α was extracted from two starting cultures and the purified DNA products were checked using nanodrop to obtain the DNA concentrations, which can be seen in table 10.

Table 10. DNA concentration of purified plasmid DNA (pMSP2N2 in DH5-α) using Nanodrop.

Component	DNA concentration
Purified plasmid DNA (from tube 1)	60 ng/µl
Purified plasmid DNA (from tube 2)	55 ng/µl

Restriction enzymes HindIII+XbaI, XhoI and EcoRI cleaved the purified DNA (MSP2N2 in DH5-α), making different DNA fragments. Cleaved products were run on a gel and the agarose gel electrophoresis in figure 1 enabled the identification of the desired cleaved products around 6000 bp (uncleaved plasmid), 5202 bp + 1218 bp (HinDIII and XbaI), 5666 bp + 540 bp + 214 bp (XhoI) and 5880 bp + 540 bp (EcoRI).

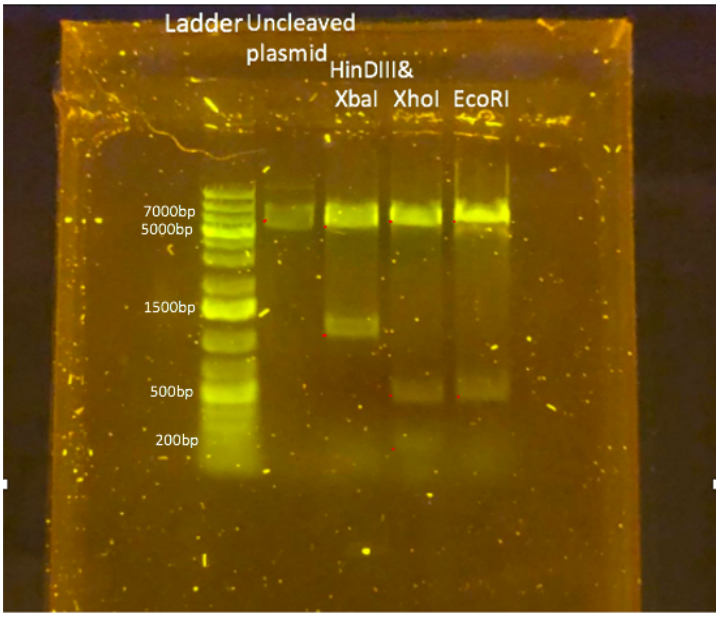


Figure 1. Cleaved plasmid (pMSP2N2 in DH5-α) with FastDigest HinDIII&XbaI (well 3), FastDigest XhoI (well 4) and FastDigest EcoRI (well 5). Well 2 from the left shows uncleaved plasmid. Red dots show desired length/products: 6420 bp (uncleaved plasmid), 5202 bp + 1218 bp (HinDIII and XbaI), 5666 bp + 540bp + 214 bp (XhoI) and 5880 bp + 540 bp (EcoRI). GeneRuler 1 kb plus DNA ladder was used.

The purified pMSP2N2 was transformed into the expression strain BL21 (DE3) / pLysS using LA plates containing Kanamycin. Colonies were then counted which can be seen in table 11. The negative controls show no colonies, 1 (100 µl) and 2 (100 µl) show fewer colonies than 1 (remaining) and 2 (remaining). 1 (remaining), however, has quite drastically fewer colonies than 2 (remaining).

Table 11. Number of colonies on transformed plates. (Transformation of pMSP2N2 from DH5-α to the expression system BL21 (DE3) / pLysS).

Plate	Number of colonies
0 (100 µl)	0
1 (100 µl)	1
2 (100 µl)	1
0 (remaining)	0
1 (remaining)	9
2 (remaining)	20

pMSP2N2 was isolated from the expression system BL21 (DE3) / pLysS in order to verify that it contained the right plasmid DNA. The purified DNA products were checked using nanodrop to obtain the DNA concentrations, which can be seen in table 12. Restriction enzymes FastDigest EcoRI and FastDigest XhoI cleaved the purified DNA, making different DNA strands. Cleaved products were run on a gel and the agarose gel electrophoresis in figure 2 enabled the identification of the desired cleaved products at 5666 bp + 540 bp + 214 bp (XhoI) and 5880 bp + 540 bp (EcoRI).

Table 12. DNA concentration of purified plasmid DNA (pMSP2N2 in BL21(DE3) / pLysS) using Nanodrop.

Component	DNA concentration
Purified plasmid DNA (from tube 1)	39 ng/µl
Purified plasmid DNA (from tube 2)	45 ng/µl

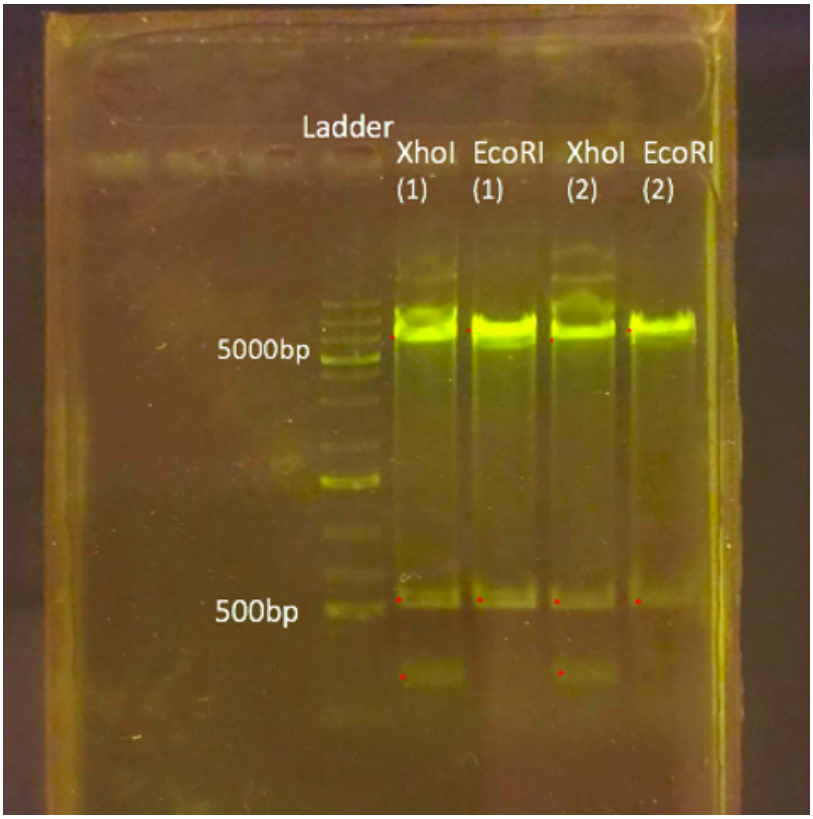


Figure 2. Cleaved plasmid (BL21 (DE3) / pLysS) with FastDigest XhoI (well 2 and 4, from the left) and FastDigest EcoRI (well 3 and 5) . XhoI (1) corresponds to purified plasmid DNA from tube 1 and XhoI (2) purified plasmid DNA from tube 2, the same applies to EcoRI. Red dots show desired length/products: 5666 bp + 540 bp + 214 bp (XhoI) and 5880 bp + 540 bp (EcoRI). Generuler 1 kb plus DNA ladder was used.

A first attempt to cultivate cells prior to expression of MSP2N2 was made according to a 5 day protocol based on Korz et al. (1995) and generated a total amount of 167 g of cells. However, this was done with pMSP2N2 in the wrong host (DH5-α) and therefore no further work was done on this pellet. Bacterial growth of *E-coli* (BL21 (DE3) / pLysS) with MSP2N2 was then done in a fermenter according to Ritchie et al. (2009) (for one day) and 4h after inoculation with starting culture, OD was checked, which can be seen in table 13. When the OD was 3.75, it was time to induce protein expression with IPTG, and when the OD was 19, the cells were harvested. The cell pellet weighed in at 84 g.

Table 13. Optical density (OD) shown from the time the fermenter is inoculated with the starter culture.

Time from start/inoculation	OD
4h	3.75
5h	8.75
6h	19

Western blot was performed on samples taken from the fermenter when the OD was checked 4 hours and 5 hours from inoculation with starting culture to ensure that the protein expression induced with IPTG worked (see figure 3). Two weak bands corresponding to the protein size of MSP2N2 (46 kDa) can be seen in well 4 and 5, which corresponds to 50 min after the culture was induced with IPTG. Two even weaker bands corresponding to the protein size of MSP2N2 can also be spotted in well 2 and 3 from the left, which was before the culture was induced with IPTG. Two thicker bands can be seen in both well 4 and 5 around 100 kDa and around 150 kDa (50 min after induction with IPTG) and two weaker bands (but thicker than the bands around 45 kDa) can be seen around 15 kDa and 30 kDa in all wells, thus both before and after induction with IPTG.

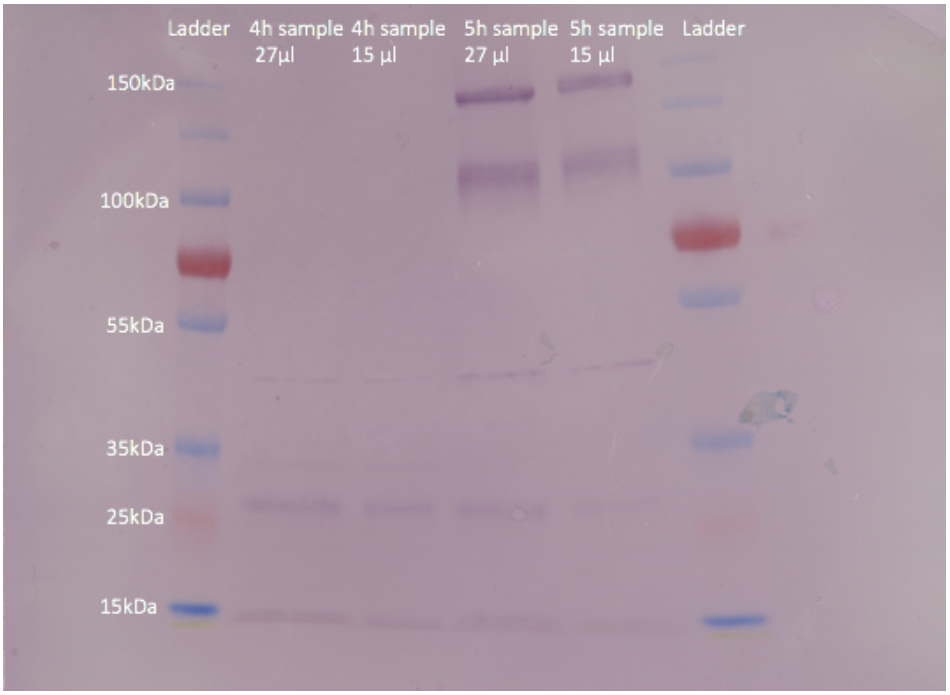


Figure 3. Western blot. Well 2 and 3 from the left contain samples from 4h after the fermenter was inoculated with starting culture (before induction with IPTG), well 4 and 5 contain samples from 5h after start (50 min after induction with IPTG). Well 4 and 5 show two weak bands corresponding to the size of MSP2N2 (46 kDa). Two even weaker bands that also correspond to the size of MSP2N2 (46 kDa) can be seen in well 2 and 3. Two thicker bands can be seen in both well 4 and 5 around 100 kDa and around 150 kDa, and two weaker bands (but thicker than the bands around 45 kDa) can be seen around 15 kDa and 30 kDa in all wells. Thermo Scientific PageRuler Plus Prestained Protein Ladder, 10 to 250 kDa was used.

Purification of MSP2N2 after the cells had been harvested was done using a nickel column (Ion Metal Affinity Chromatography (IMAC)). MSP2N2 was eluted with 40 mM Tris/HCl, 0.3 M NaCl, 0.4 M Imidazole and collected as 10 fractions (E0-E9). Protein concentration of the different fractions was checked using nanodrop, which can be seen in table 14. The protein concentration increased between E0 and E2 and was highest in E2 before becoming lower again. E1+E2 gave together 2.5 mg/2 ml protein and since 10 ml lysate was loaded on the column (2.5 mg/10 ml) and we had 200 ml lysate from the beginning, the expected protein yield is $2.5 \times 20 = 50$mg and given that we started with 84 g of cells, an expected yield would be 50 mg/84 g = 0.60 mg protein per g of cells.

Table 14. Protein concentration on the eluate from the nickel column purification (IMAC) (attempt 1) using nanodrop.

	Concentration (mg/ml)
E0	0.125
E1	0.812
E2	1.729
E3	0.130
E4	0.071
E5	0.071
E6	0.068
E7	0.065
E8	0.109
E9	0.052

After the purification attempt of MSP2N2 using IMAC, a gel was run which was then stained with coomassie on samples from flow through of the lysate as well as flow through of the third wash and eluate fractions 0-5, which can be seen in figure 4. The wells corresponding to wash 3, E0 and E3-E5 are empty. Both E1 and E2 turn out to be quite complex samples. In E1 we can see bands around 15 kDa, 25 kDa, 70 kDa and 100 kDa and the bands further up are more clearly visible than the bands further down. In E2 we can see a band corresponding to the size of MSP2N2 (46 kDa), but this band does not dominate but rather the bands that, just like in E1, can be seen around 15 kDa, 25 kDa, 70 kDa and 100 kDa.

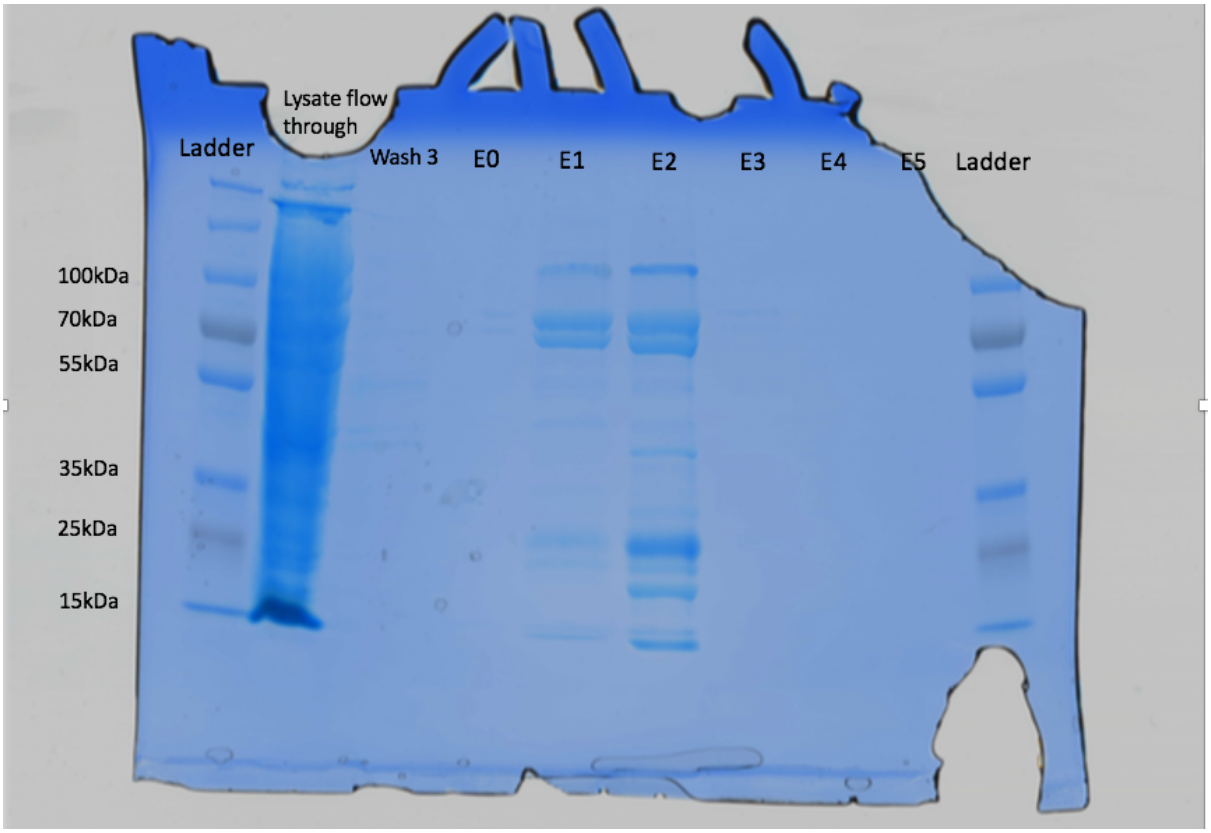


Figure 4. Coomassie stained gel from the first purification attempt of MSP2N2. Well 2 (from the left) contains flow through of the lysate, well 3 contains flow through from wash 3 and well 4-9 contains fractions 1-5 from the elution. Thermo Scientific PageRuler Plus Prestained Protein Ladder, 10 to 250 kDa was used.

Given that the two eluate fractions from the first purification attempt of MSP2N2 that were expected to contain our protein and that were run on a gel turned out to be so complex, SEC was performed. Before the SEC, E1 and E2 were pooled and concentrated using an Amicon Ultra-4 Centrifugal filter Unit (30 kDa cutoff) and filtered using a 1 ml syringe with 0.45 µm filter. After concentration and filtration, protein concentration was checked using nanodrop which can be seen in table 15. The yield of protein received from the concentration step was 0.5 mg/0.5 ml. The results from the SEC were not informative, air bubbles drifted the baseline and no clear peaks could be distinguished (data not shown).

Table 15. Protein concentration from the concentration and filtration of E1+E2 (first purification attempt) with Amicon Ultra-4 Centrifugal filter Unit (30 kDa cutoff) and a 1 ml syringe with a 0.45 µm filter.

	Concentration (mg/ml)
E1+E2 after concentration step	1.006
Concentrated sample after 0.45µm filter	0.149

Given that concentration and filtration with Amicon Ultra-4 Centrifugal filter resulted in so much loss of sample after purification using IMAC and that SEC gave no clear results, the experiment was repeated from the purification of MSP2N2, this time starting from 50 ml of lysate. Eluate was again collected in fractions and protein concentration was checked in each fraction using nanodrop, which can be seen in table 16. The protein concentration increased between E0 and E2 and was at its highest in E2 before becoming lower again. E1+E2 together gave 7.0 mg/2 ml, which is not 5x more protein as was expected.

Table 16. Protein concentration on the eluate from nickel column purification number 2.

	Concentration (mg/ml)
E0	0.428
E1	3.209
E2	3.810
E3	0.380
E4	0.140
E5	0.099
E6	0.095
E7	0.069
E8	0.052
E9	0.042

After the second purification attempt of MSP2N2 using IMAC, a gel stained with coomassie and an western blot was run on samples from the flow through of the lysate, the third wash and eluate fractions 0-3, which can be seen in figure 5 and 6. What you can see is that both the gel and the membrane have become overloaded so it is difficult to distinguish any results, and that both the gel and the membrane are quite complex. In figure 5 in all wells except the one corresponding to the flow through of the lysate, we can see bands around the size of MSP2N2 (46 kDa). We can also see several dominant bands both above (around 70 kDa) and below (around 30 kDa), especially in well 5-8 from the right. Approximately the same result can be distinguished in figure 6. The heated eluate fractions (E1-E2) resulted in loss of thick bands above 70 kDa in figure 5, but instead the bands got stuck higher up in the well. The heated eluate fractions showed no visible difference in figure 6.

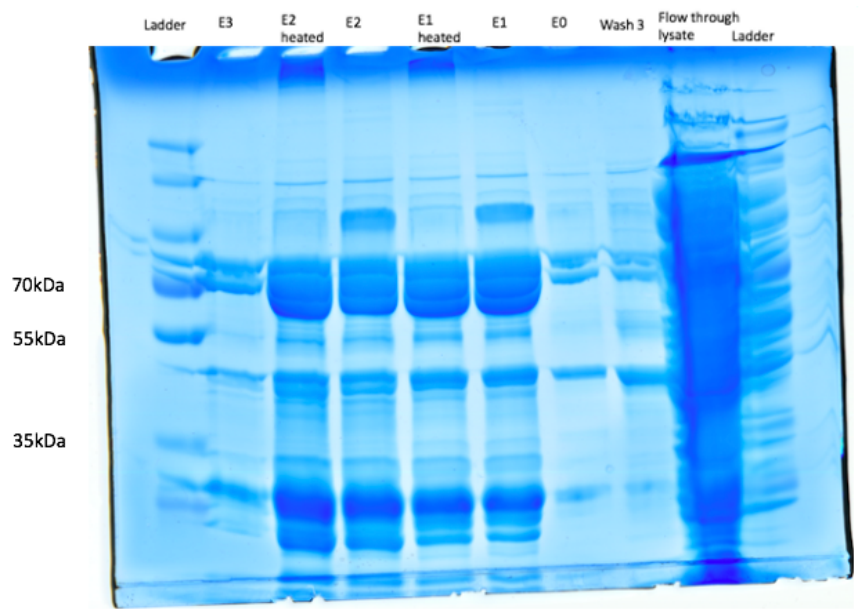


Figure 5. Coomassie stained gel from the second purification attempt of MSP2N2. Well 2 (from the right) contains flow through the lysate, well 3 contains flow through from wash 3 and well 4-9 contains fractions 1-3 from the elution, whereby one E1 and one E2 are heated (well 6 and 8). Thermo Scientific PageRuler Plus Prestained Protein Ladder, 10 to 250 kDa was used.

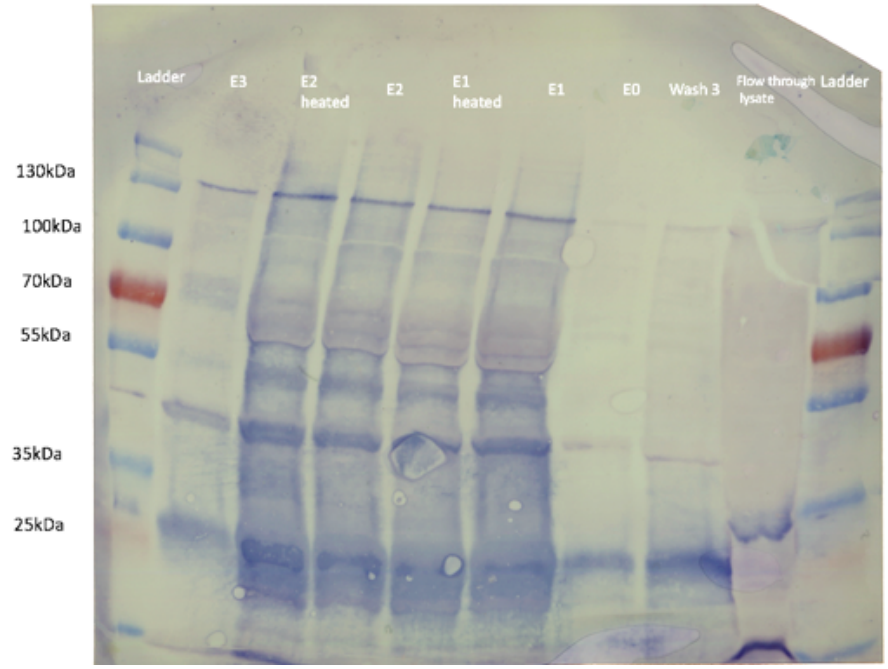


Figure 6. Western blot from the second purification attempt of MSP2N2. Well 2 (from the right) contains flow through the lysate, well 3 contains flow through from wash 3 and well 4-9 contains fractions 1-3 from the elution, whereby one E1 and one E2 are heated (well 6 and 8). Thermo Scientific PageRuler Plus Prestained Protein Ladder, 10 to 250 kDa was used.

Another SEC was performed to try to find the size of the different aggregates we could see on western blot/coomassie stained gel. Before the SEC, E1+E2 from the second purification was concentrated and filtered using “Vivaspin 15R centrifugal concentrator” (30 kDa cutoff) and “Spin-X 0.45 µm Costar1". Protein concentration on what went through the filters versus what did not go through the filters was checked using nanodrop (see table 17 and 18). The expectation was to get 10 mg/ml protein after concentration, as seen in table 18, we only received 3.4 mg/ml. Again, the results from the SEC were very vague with a drifting baseline and air bubbles (no data shown). What could be discerned was a small peak where the absorbance line corresponding to 280 nm went over the absorbance line for 260 nm and therefore fractions A29 and A30 from the SEC are later run on a gel and western blot.

Table 17. Protein concentration on E1+E2 and flow through from concentration/filtration using Vivaspin 15R centrifugal concentrator (second purification).

	Protein concentration mg/ml
E1+E2	3.826
Flow through	0.021

Table 18. Protein concentration on E1+E2 and flow through from concentration/filtration using Spin-X 0.45µm Costar1 (second purification).

	Protein Concentration mg/ml
E1+E2	3.361
Flow through	1.902

A second gel was run on flow through of the lysate, E0 and E3 from the elution, flow through from the concentration with Vivaspin, flow through from the Spin-X column, the sample that did not go through the Spin-X filter and fraction 29 and 30 from the SEC, which can be seen in figure 7. In well 2 and 3 from the right that contains eluate fractions 0 and 3 there are bands corresponding to the size of MSP2N2 (46 kDa), but also bands around 15 kDa, 25 kDa and 70 kDa. In well 4 from the right where flow through of the lysate is found, we can see a band around the size of MSP2N2 (46 kDa). Well 8, 9 and 10 from the right are empty (flow through of E1+E2 from the Vivaspin column, fraction A29 and A30 from the SEC). Well 6 and 7 from the right containing flow through of E1+E2 from the Spin-X column and E1+E2 that did not go through the Spin-X column has the most concentrated bands around 30 kDa, but bands can also be seen around 15 kDa, 70 kDa and 100 kDa. Possibly a very weak band can be spotted around 45 kDa which corresponds to the size of MSP2N2 (46 kDa) in well 7 from the right.

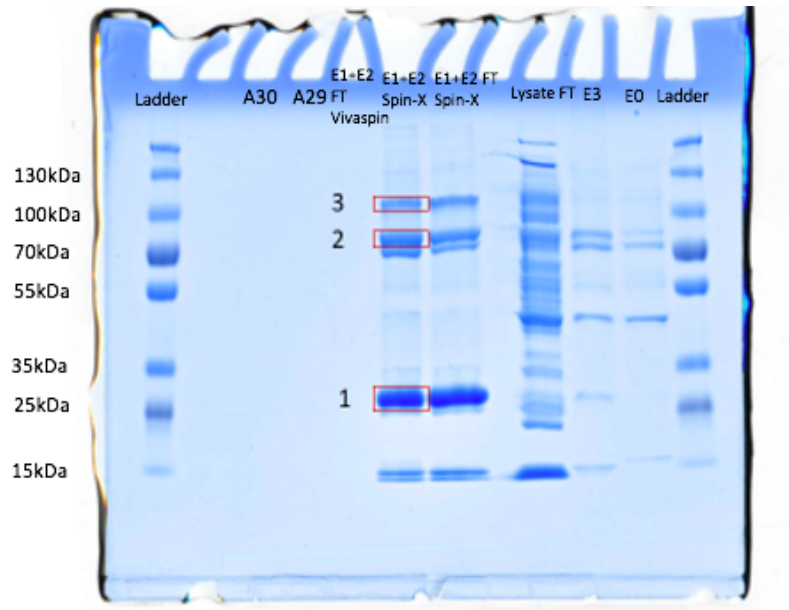


Figure 7. Coomassie stained gel (second purification). Well 2 and 3 (from the right) contain eluate fraction 0 and 3, well 4 contains flow through of the lysate, well 6 contains flow through of E1+E2 from the Spin-X column, well 7 contains sample of E1+E2 that did not pass through the filter (Spin-X), well 8 contains flow through of E1+E2 from the Vivaspin column, well 9 contains fraction A29 from SEC and well 10 contains fraction A30 from SEC. Thermo Scientific PageRuler Plus Prestained Protein Ladder, 10 to 250k Da was used. The red marked bands 1-3 were later used for mass spectrometry.

Since degradation of MSP2N2 could not be excluded from the previous gel, a western blot was also performed (more sensitive method) on samples E0, E3, flow through of the lysate (from IMAC), E1+E2 (that did not go through the Spin-X column) and fraction A29 +A30 from the SEC, which can be seen in figure 8. In well 2 and 3 from the right that contains eluate fractions 0 and 3 has bands corresponding to the size of MSP2N2 (46 kDa). Both E0 and E3 show bands around 25 kDa and in E3 we can also see a band around 15 kDa. Well 4 containing flow through of the lysate is quite clean and no band corresponding to MSP2N2 can be seen. In well 5 containing E1+E2 that did not pass through the Spin-X filter has a band corresponding to the size of MSP2N2 (46 kDa), but the sample is quite complex and we can see bands around 15 kDa, 30 kDa, 55 kDa, 70 kDa and 100 kDa. Well 7 and 8 containing fractions A29 and A30 from the SEC are empty.

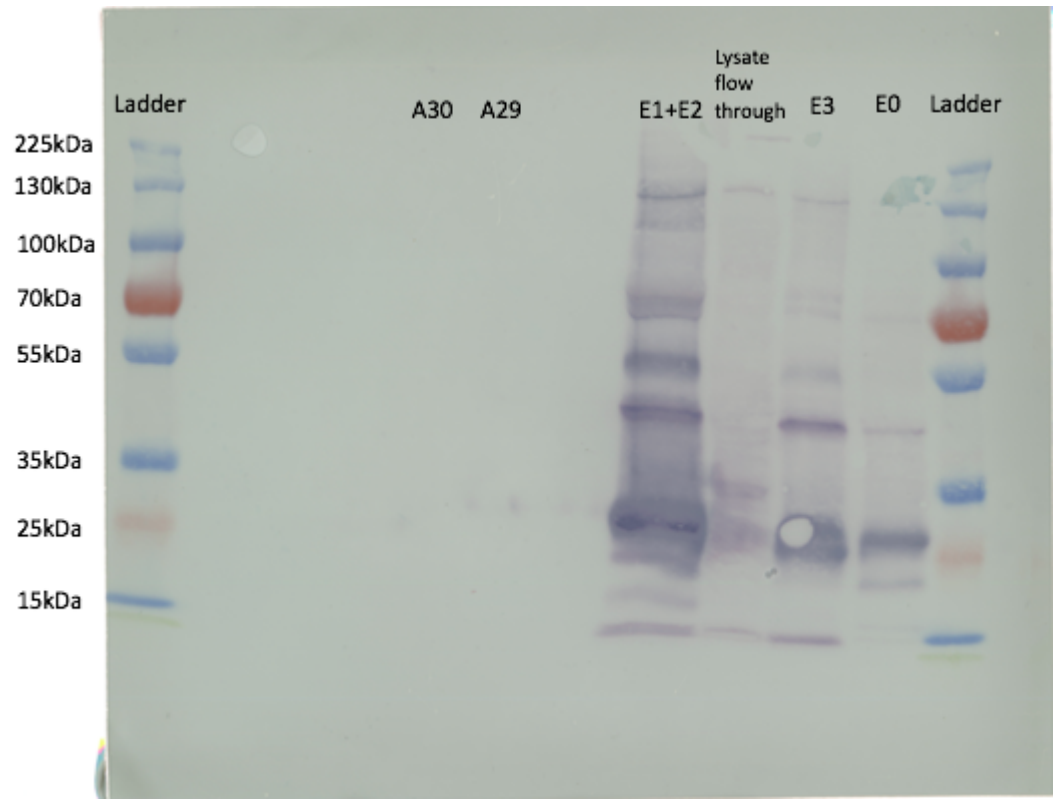


Figure 8. Western blot (second purification). Well 2 and 3 (from the right) contain eluate fraction 0 and 3, well 4 contains flow through of the lysate, well 5 contains E1+E2 that did not go through the Spin-X filter, well 7 and 8 contains fraction A39 and A30 from SEC. Thermo Scientific PageRuler Plus Prestained Protein Ladder, 10 to 250 kDa was used.

For the sake of comparison, a western blot was performed on samples before induction with IPTG, after induction with IPTG, cell lysate (after cultivation in the fermenter) and E1+E2 from the second purification (what did not go through the Spin-X filter), which can be seen in figure 9. The third well from the right containing a sample from the cultivation in the fermenter before induction with IPTG shows two bands, one at 15 kDa and one at 30 kDa. Well 5 containing a sample from the cultivation in the fermenter after induction with IPTG also shows bands at 15 kDa and 30 kDa, as well as a very weak band around 45 kDa which corresponds to the size of the MSP2N2 protein (46 kDa) and two bands at 100 kDa and 130 kDa. The cell lysate in well 7 has a band corresponding to the size of MSP2N2 (46 kDa) as well as bands at 15 kDa, 30 kDa, 100 kDa and 130 kDa. Well 9 containing E1+E2 from the second purification (what did not go through the Spin-X filter) is a rather complex sample with many bands, also here we can distinguish one band by the size of MSP2N2 (46 kDa), but also several bands both above and below 45kDa.

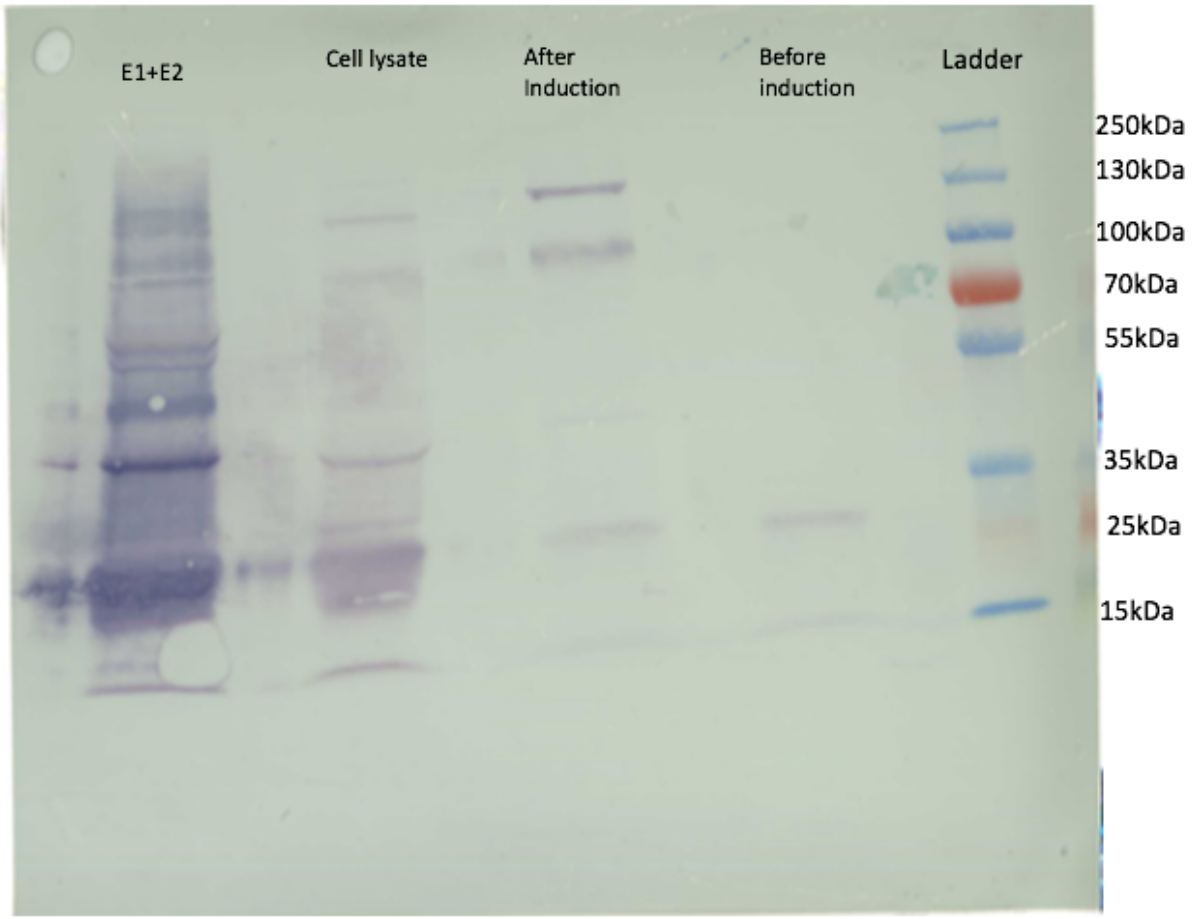


Figure 9. Western blot for comparison on samples before induction with IPTG (well 3 from the right), after induction with IPTG (well 5), cell lysate after cultivation in the fermenter (well 7) and E1+E2 from the second purification (what did not go through the Spin-X filter) (well 9). Thermo Scientific PageRuler Plus Prestained Protein Ladder, 10 to 250 kDa was used.

Mass spectrometry was performed on three bands from one of the gels stained with coomassie (see figure 7) because these bands have dominated on several occasions. The result of the analysis of the three gel bands by mass spectrometry can be seen in Appendix 2. No peptides belonging to the protein sequence of MSP2N2 (see figure 10) was found. Three other proteins were identified in the gel bands based on MS and MSMS data, namely: FKBP-type peptidyl-prolyl cis-trans isomerase SlyD OS, Bifunctional polymyxin resistance protein ArnA OS and RNA-binding protein Hfq (Appendix 2).

MGHHHHHHHHDYDIPTTENLYFQGSTFSKLREQLGPVTQEFWDNLEKETEGLRQEMSKDLEE
VKAKVQPYLDDFQKKWQEEMELYRQKVEPLRAELQEGARQKLHELQEKLSPLGEEMRDRA
RAHVDALRTHLAPYSDELRQRLAARLEALKENGGARLAEYHAKATEHLSTLSEKAKPALEDL
RQGLLPVLESFKVSFLSALEEYTKKLNTQGTPVTQEFWDNLEKETEGLRQEMSKDLEEVKAK
VQPYLDDFQKKWQEEMELYRQKVEPLRAELQEGARQKLHELQEKLSPLGEEMRDRARAHV
DALRTHLAPYSDELRQRLAARLEALKENGGARLAEYHAKATEHLSTLSEKAKPALEDLRQGL
LPVLESFKVSFLSALEEYTKKLNTQ

Figure 10. Protein sequence of MSP2N2. Red marking corresponds to the recognition sequence of tobacco etch virus (TEV).

4. Discussion

If we start by evaluating the cleavage of the purified DNA (pMSP2N2 in DH5-α), then we can see in figure 1 that it gave the desired lengths of the cleaved plasmid, which means that the purified DNA is most likely the pMSP2N2 as aimed for. This made us choose to proceed with the transformation of MSP2N2 into the expression system BL21 (DE3) / pLysS. If we take a look at the results from our transformation plates (see table 11) then we can see that we have no colonies on the negative controls and it indicates that the work has been performed relatively sterile. That we have colonies on the rest of the plates indicates that we have competent cells, however, they do not seem to be very competent because in that case we should have received many more colonies per nanogram of cells, but since we only wanted to change the host, it is acceptable. The reason that the number of colonies differs quite drastically between 1 (remaining) and 2 (remaining) probably has to do with an previously unintentional cleavage reaction, were 1 µl of FastDigest enzyme EcoRI was accidentally added to purified plasmid DNA tube 1, which may have started a degradation before it was inactivated. The next cleavage with FastDigest EcoRI and FastDigest XhoI was to verify that the plasmid was intact after transformation into expression system BL21 (DE3) / pLysS. In figure 2, we can see the desired lengths of the cleaved products, so just like in the previous cleavage we could assume that the correct plasmid was present.

When culturing our cells in the fermenter, samples were taken 4 hours (before induction of protein expression with IPTG) and 5 hours from inoculation with starting culture (50 minutes after induction of protein expression with IPTG) and run on a western blot (see figure 3). We can see weak bands corresponding to the size of MSP2N2 (46 kDa) both before and after induction, which could mean that the promoter of the strain has not kept tight and thus induced expression even before IPTG was added. However, the most dominant bands are the high molecular weight bands (above 100 kDa and just below 150 kDa) that we can see after induction with IPTG rather than the bands that actually corresponds to MSP2N2 in size (46kDa), one thought is whether these high molecular weight bands could be multimers of MSP2N2, which could also explain why the bands corresponding to the size of MSP2N2 are so weak. The bands we can see at 15 kDa and around 30 kDa are available both before and after induction with IPTG, which could therefore indicate that they correspond to different proteins from the *E-coli* strain that was already present.

Eluate 2 from the first purification (where most protein should be present according to the nanodrop (see table 14), shows a band corresponding to the the size of MSP2N2 (46 kDa) in figure 4, however, it is quite weak and there are low and high molecular bands in addition to this that are difficult to explain. The idea that the high molecular weight bands could be a multimer of MSP2N2 remains and for that reason we later also tried to heat up E1 and E2 to see if we could break intermolecular bonds and get rid of the high molecular bands (see figure 5 and 6), which did not work, the only thing we could see was that some bands disappeared but they got stuck further up in the wells instead. If we compare figure 4 and figure 3, we can see corresponding bands around 15 kDa and 25 kDa also in figure 3 and then

already before induction with IPTG, which strengthens the fact that these bands probably do not belong to our protein but could rather be impurities that are copurified with the desired protein.

To see if one could get any clarity in all these low and high molecular bands, SEC was performed and before the SEC, pooled E1+E2 containing 2.5 mg/2 ml (see table 14) was concentrated using Amicon Ultra-4 Centrifugal filter Unit (30 kDa cutoff) and a 1 ml syringe with a 0.45 μm filter (since we can only load 0.5 ml to the SEC). In table 15 we can see that the concentration resulted in a lot of loss since we expected 2.5 mg/0.5 ml but instead only got 0.5 mg/0.5 ml (E1+E2 after concentration step). What on the other hand went through the 0.45μm filter had a protein concentration that was measured up to 0.149 mg/ml (0.149 mg/ml x 0.5 ml = 0.07 mg), which clearly indicates a loss. One reason for the loss of protein could be that instead of using a syringe filter, a Spin-X centrifuge Tube filter should have been used in order to not lose as much volume since you never have the opportunity to push all of the sample through the filter manually. In the end there was only 100 μl of the sample left (E1+E2) that could be used for the SEC as most of the sample was probably stuck in the filter which made it difficult to handle during injection. The fact that there was only so little of the sample left was probably also the reason why it was not possible to distinguish any direct results from the chromatogram (hence the data not shown).

Since the preparation before the SEC resulted in so much loss of the sample to work with, the purification with nickel column was performed again. According to Ritchie et al. (2009) page 221 it is mentioned that for reconstitution of another MSP protein, namely MSP1E3, approximately 200 μM (i.e 0.0002 mol/l) of MSP1E3 stock solution is needed. If we apply this to MSP2N2 with its molecular weight of 46 kDa (46000 g/mol), this means that we would need at least 9.1 mg/ml of MSP2N2, say that we then strive for 0.2 ml MSP2N2 stock, then we would need approximately 2 mg of protein. Previous purification with 10 ml of cell lysate gave 2.5 mg protein, so to ensure that a sufficient amount was obtained, the experiment was scaled up 5 times with 50 ml cell lysate (expected yield of protein is then 12.5 mg/50 ml). As can be seen in table 16, the elution gave only 7.0 mg protein (E1+E2) and not 5x as much protein as expected. In a study by Shenkarev et al. (2018) they created lipid-protein nanodiscs using MSP2N2 to study a voltage gated potassium channel and got a final yield of MSP2N2 of 70 mg/l bacterial culture and if we expected 12.5 mg/50 ml it means that we should get 20mg protein per liter, which is a pretty big difference in quantity.

Figure 5 and 6 of eluate 1 and 2 from the second purification attempt show an overloaded gel/western blot which makes it difficult to evaluate the results, but what we can say is that E1 and E2 are complex samples that show many bands and that the bands that corresponds to the size of MSP2N2 (46 kDa) are quite weak. A possible explanation for the complexity could be degradation, that proteases have degraded the protein and therefore the band at the size of MSP2N2 is weaker and the more dominant bands further down is a product of degradation (the purification had protease inhibitors in the form of PMFS, but these only work to a certain extent) and that there may also be multimers of the degradation product. The next possible explanation could again be non-specific binding / that contaminants have

stuck to the column as there could be other proteins in the *E-coli* strain that also have histidine and are therefore selected for.

When eluate 1 and 2 from the second purification were pooled and concentrated for the next SEC, we can see from table 18 that we did not get the protein concentration we expected (10 mg/ml) using Vivaspin 15R centrifugal concentrator, instead we only got 3.361 mg/ml and when we used Spin-X 0.45µm Costar1we encountered problems were we got a precipitate of the E1+E2 sample that clogged the filter and stopped the sample from successfully getting through the filter. So again, we lost a lot of the sample which resulted in only 100 µl being loaded to the SEC and the only result that could be discerned was a very small peak (data not shown) where the absorbance line for 280 nm was above the absorbance line for 260 nm (proteins absorb more light at 280 nm), which was the reason why we later went on with fraction A29 and A30 to run on a gel/perform western blot, which, however did not show anything. A possible explanation for the fact that we get precipitation of our protein can perhaps be obtained if we look at the study conducted by Shenkarev et al. (2018) and product information sheet of MSP2N2 from Sigma-Aldrich (2017) where sodium cholate has been used to keep MSP2N2 stable in solution. In Shenkarev et al. (2018), cholate is used as a detergent to solubilize a potassium channel from *Pichia pastoris* (lipids and protein of interest) and in Sigma-Aldrich (2017) when formation of nanodiscs with MSP2N2 is done they also used sodium cholate to keep the protein stable in solution, which is interesting because we did not have the protein in sodium cholate at the end of the purification and then we got precipitation when the sample was prepared for the SEC.

In figure 7 we can see bands corresponding to the size of MSP2N2 (46 kDa) in E0, E3, E1+E2 from the Spin-X column and E1+E2 that did not pass through the Spin-X filter. The bands that dominates in E0, E3, E1+E2 from the Spin-X column and E1+E2 that did not pass through the Spin-X filter are the bands above and below the size of MSP2N2, and therefore we chose to use this gel and cut out three bands (see figure 7) and run mass spectrometry on them to find out what they are (Appendix 2). One thought is still that the thick bands at 30 kDa in well 6 and 7 (from the right) could be MSP2N2 but in a degraded form and therefore we wanted to evaluate this further by running another western blot (see figure 8), but it turned out that E1+E2 showed about the same thing on western as on the coomassie gel. What speaks against degradation is the induction pattern from the fermenter (see figure 9) because there we can see that these bands that could possibly be degradation already existed before protein expression was induced with IPTG, which instead strengthens the argument that speaks for background and that other proteins in the *E-coli* strain may have histidines and therefore they are enriched on the nickel column. However, we can see in figure 9 that the thick/concentrated bands around 30 kDa are strong already in the cell lysate, which indicates that if there were to be degradation then it probably happens already when you break the cells with sonication (which is not unreasonable because when you break down the cells, proteases are released) and not necessarily that it would occur during purification.

If we take a look at the results of analysis by mass spectrometry (Appendix 2) of the three gel bands marked in figure 7, no peptides belonging to MSP2N2 were detected. Three other proteins were identified using MS and MSMS data, namely FKBP-type peptidyl-prolyl cis-trans isomerase SlyD OS, Bifunctional polymyxin resistance protein ArnA OS and RNA-binding protein Hfq. The protein sequence with 37% coverage to gel band 1 (FKBP-type peptidyl-prolyl cis-trans isomerase SlyD OS) has a lot of histidines at the end so it is not strange that this protein would stick to the column. Even if it is not exactly our E-coli strain it is expected that this protein is included in our strain as well, and according to Kovermann et al. (2013), SlyD is a chaperone that naturally binds to nickel ions.

In the case of band 2 and 3 we can see that the protein which matches gel band 2 also has histidines and therefore could be selected for, but gel band 3 is harder to explain since the band should have a size of over 100 kDa while the protein that matches is only 10 kDa and it is not even the same organism. However, both band 1 and 2 could be explained with the help of Andersen et al. (2013) which highlights the problem of purifying a protein in an *E-coli* strain using a his-tag due to contamination of other proteins in the *E-coli* host which are also rich in histidines. It turns out that two such common contaminants are SlyD (match of gel band 1) and ArnA (match of gel band 2), and these compete for the binding sites of the nickel column. To solve this problem, an *E-coli* low-background-strain, also named LOBSTR which lowers the affinity for nickel in SlyD and ArnA can be used. Andersen et al (2013) also mention another contaminant, a RNA binding protein called Hsp15 which has many positively charged amino acids and is able to bind to SlyD due to SlyD being very negatively charged. This gave us a clue to gel band 3 since Hfq also is an RNA binding protein, the idea was that Hfq may be related to Hsp15 and therefore we checked if Hfq also had many positively charged amino acids to see if it could have interacted through electrostatic bonding with SlyD and become stuck to the IMAC. The total number of positively charged residues of Hspq (Uniprot, P0A6X3) was only 8 (Expasy ProtParam) compared to Hsp15 (Uniprot, P0ACG8) which had 30 (Expasy ProtParam), which makes the argument rather weak.

However, mass spectrometry is not directly a quantitative method, so even if the band has a certain size, there is a risk that it could be a double band from contamination. Mass spectrometry also has quite low coverage so that only certain peptides are detected, so if our protein of interest has peptides that fly poorly then a background of something else can dominate, so that we can not detect our protein does not have to mean that it is not there.

4.1 Conclusion and future approaches

In conclusion, this project shows that overexpressing and purification of MSP2N2 can be difficult in practice. Either MSP2N2 has been degraded, other proteins in the *E.coli* strain have dominated, the induction with IPTG has not worked or it has not been the right *E.coli* strain or plasmid at all. This is based on the fact that the bands corresponding to the size of MSP2N2 were generally quite weak and other low and high molecular bands dominated, and that one could actually see bands corresponding to MSP2N2 even before induction.

For future studies, one could cleave the plasmid with restriction enzymes several times as well as sequencing to further ensure that the plasmid was okay. Regarding induction, one could have been more critical of the fact that the band did not emerge stronger after the addition of IPTG, however, we thought that the protein could possibly have been shaped as a multimer and that it would explain the weakness of the band, but one could have compared with similar studies earlier, such as Shenkarev et al. (2018) where you could see in figure 4 that the band came out as its original size and not as a multimer. Because of this, one should try redoing the cultivation and induction with IPTG and then check more specifically with western blot around 45 kDa at several time points to see if any band appears there compared to before induction.

One could also note that the competent cells used for transformation (BL21 (DE3)/pLysS) came from previous teaching and we did not actually check if the strain had the pLysS part (from the mass spectrometry we could at least see that we had E.coli). Lysozym S breaks down T7 RNA polymerase as otherwise may leak from the promoter, if pLysS was not present, it could strengthen the argument that MSP2N2 was induced even before the addition of IPTG due to leakage from the promoter. For future experiments one should check if the strain has pLysS by smearing the strain on a plate with Chloramphenicol since the plasmid encoding LysS is resistant to chloramphenicol (Studier et al., 1990). If you repeat the experiment, it may also be worth pointing out that you do not have to run the the cultivation of cells in the fermenter for only one day, you could also run a fermentation for 5 days according to Korz et al. (1995) to get a higher cell density, a first attempt to grow the cells was made according to this protocol and then we got a total of 167 g of cells compared to when we followed Ritchie et al. (2009) and got 84 g (the first time, however, we cultured the plasmid in the wrong host, DH5-α).

To evaluate contaminants, one could have a transformed *E-coli* as a control, a corresponding strain that does not contain our construct. You could do this by running cultivation in two flask cultures (instead of growing in fermenter to save time) and then have one culture with IPTG and one without IPTG and compare which bands appear in each cell lysate. If you get these high and low molecular bands in the culture without IPTG, it would probably mean that we have proteins that for us are irrelevant in the E-coli strain. Another idea would be to try to cleave of the His-tag from our protein because in that case the protein would not stick to the nickel column that selects for histidine, but if there are contaminants that also have histidine bound they would still get stuck and then you could instead save what goes through the

column and use IMAC as a way to get rid of the contaminants. The His tag could be cleaved off using a highly specific cysteine protease called tobacco etch virus (TEV). The sequence TEV recognizes consists of 7 amino acids: E-X-X-Y-X-Q-S where X can be different amino acids. The protease will cleave between glutamine and serine (Terpe, 2003). Here the recognition site is ENLYFQG, see figure 10 (marked in red). When TEV cuts of the His-tag, the size of MSP2N2 will go from 46 kDa to 43 kDa, a difference of 3 kDa so considering that you might also be able to see the difference in size of the protein in a gel and if the thick band we have seen several times around 30 kDa would be a degradation product of MSP2N2, we would have seen the jump in size even more clearly here (3/45 versus 3/30).

Nor is it impossible that we have had contamination of a completely different *E.coli* strain, for example in the fermenter, if we take a look at the protocol which was followed for expression (Ritchie et al., 2009), it feels like the amount of added Kanamycin for the cultivation was quite low (25 mg) and if you would redo the experiment, it could be an idea to increase the concentration.

Regarding degradation, a solution could be to redo the whole purification process faster and under more controlled temperature conditions where you make sure that the samples are kept on ice at all times, because if degradation is suspected, it is assumed that it occurs faster at room temperature. One could also try other protease inhibitors if PMFS did not work. You could also evaluate how homogeneous the sample is and check the size distribution (if it is now the case that you tend to get aggregates of your protein) with the help of dynamic light scattering (DLS).

5. References

Andersen, K. R., Leksa, N.C., Schwartz, T. U. (2013) Optimized *E.coli* expression strain LOBSTR eliminates common contaminants from His-tag purification. *Proteins: Structure, Function, and Bioinformatics. Volume (81), 1857-1861.* DOI: 10.1002/prot.24364

Birch, J., Cheruvara, H., Gamage, N., Harrison, P. J., Lithgo, R., and Quigley, A. (2020). Changes in membrane protein structural biology. *Biology, Volume (11),* 1-18. DOI: 10.3390/biology9110401

Grinkova, Y. V., Denisov, I. G., and Siligar, S. G. (2010). Engineering extended membrane scaffold proteins for self-assembly of soluble nanoscale lipid bilayers. *Protein Engineering, Design & Selection, Volume (23),* 1 -6. DOI: 10.1093/protein/gzq060

Korz, D.J., Rinas, K., Hellmuth, K., Sanders E.A., Deckwer, W.D. (1995). Simple fed-batch technique for high cell density cultivation of *Escherichia coli*. *Journal of Biotechnology, Volume (39),* 59-65. DOI: 10.1016/0168-1656(94)00143-Z.

Kovermann, M., Schmid, F. X., Balbach, J. (2013) Molecular function of the prolyl cis/trans isomerase and metallchaperone SlyD. *Biological Chemistry. Volume (394),* 965-975. DOI: 10.1515/hsz-2013-0137.

Ritchie, T. K., Grinkova, Y. V., Bayburt , T. H., Denisov, I. G., Zolnerciks, J. K., Atkins , W. M., and Sligar S. G. (2009). Reconstitution of Membrane Proteins In Phospholipid Bilayer Nanodiscs. *Methods in Enzymology, Volume (464),* 211-228. DOI:10.1016/S0076-6879(09)64011-8.

Luckey, M. (2014). Membrane Structural Biology (second edition). Cambridge University press.

Shenkarev, Z. O., Karlova, M. G., Kulbatskii, D. S., Kirpichnikov, M. P., Lyukmanova, E.N., Sokolova, O. S. (2018). Recombinant Production, Reconstruction in Lipid-Protein Nanodiscs, and Electron Microscopy of FullLength αSubunit of Human Potassium Channel Kv7.1. *Biochemistry (Moscow). Volume (83),* 562-573. DOI: 10.1134/S0006297918050097.

Studier, F. W., Rosenberg, A. H., Dunn, J. J., Dubendorff, J. W. (1990). Use of T7 RNA polymerase to direct expression of cloned genes. *Methods in Enzymology. Volume (185),* 60-89. DOI: 10.1016/0076-6879(90)85008-C

Suo, Y.,Wang, Z., Zubcevic, L., Borgnia, M., Ji, R., and Lee, S. (2020). Structural Insights into Electrophile Irritant Sensing by the Human TRPA1 Channel. *Neuron. Volume (105*), 882-894. DOI: https://doi.org/10.1016/j.neuron.2019.11.023

Terpe, K. (2003) Overview of tag protein fusions: from molecular and biochemical fundamentals to commercial systems. *Applied Microbiology and Biotechnology. Volume (60), 523-533.* DOI: 10.1007/s00253-002-1158-6.

Appendix 1.

Detailed Material and methods.

To enable reconstitution trials of HaTRPA1 in nanodiscs, MSP2N2 needs to be isolated, overexpressed and purified. The starting point for the isolation of MSP2N2 is an *Escherichia coli (E-coli)* glycerol stock (LU85) containing the plasmid pMSP2N2 (Addgene plasmid #29520 ; http://n2t.net/addgene:29520 ; RRID:Addgene_29520, UK), which carries the gene encoding the desired MSP2N2 protein. After an initial period of growth of bacterial culture to gain sufficient cell mass, expression of MSP2N2 will be induced, and subsequently purification will be done using IMAC. Quantity and purity will be evaluated with SDS PAGE, Western blot and Mass spectrometry.

Plasmid preparation for purification of plasmid DNA and transformation of pMSP2N2 from DH5-α to the expression system BL21 (DE3) / pLysS.

For the expression of MSP2N2 the bacterial expression system BL21(DE3)/pLysS will be used. High-efficiency expression in *E-coli* is allowed by the expression system BL21 (DE3)/pLysS of genes that have a ribosome binding site and that are controlled by a T7 promoter. BL21 (DE3) produce T7 RNA polymerase under the control of the IPTG inducible lac UV5 promoter and the plasmid pLysS expresses T7 lysozyme which degrades low levels of T7 RNA polymerase which may leak from the promoter of uninduced cells, but does not notably affect the high expression induced by IPTG (Studier et al., 1990). Since the plasmid is delivered in DH5-α cells, plasmid preparation and transformation is needed to transfer the plasmid to the expression system BL21 (DE3) / pLysS to enable protein expression.

Two Luria Agar (LA) plates were made to be able to create colonies for the production of starter culture prior to plasmid preparation and transformation. This was done by preparing 50 mL LA containing 50 µl of 50 mg/ml Kanamycin, LU85 was spread on one of the two plates and placed in a heating cabinet set to 37°C for incubation overnight. Two starter cultures were prepared as follows: 20 ml Luria Broth (LB) was divided into two 50 ml falcon tubes and 6 µl of 50 mg/ml Kanamycin were added to each tube. The tubes were then inoculated from the same colony and shaken at 37°C overnight.

The two starting cultures were centrifuged at 3600 rpm for 15 min and the supernatants were poured off. Each pellet was then used to extract plasmid DNA using the “Plasmid miniprep kit” protocol associated with GeneJET Plasmid Miniprep Kit from Thermo Fisher Scientific (Sweden). Once the DNA was extracted, DNA concentration was checked using nanodrop (see table 10). To verify that the plasmid was intact after purification from DH5-α, three FastDigest restriction enzyme digests (HindIII+XbaI, XhoI and EcoRI) were prepared in 3 tubes in accordance with the manufacturer's recommendations (ThermoFisher Scientific, Sweden), see table 1, 2 and 3. The tubes were placed for incubation in a water bath at 37°C for 15 min. 21 µl from each tube together with 4 µl loading dye (gelstar included, 1 µl gelstar

to 1 ml 6x loading dye) was loaded to a gel. 8 µl uncleaved plasmid with 1.6 µl loading dye were loaded to the gel as a control (see figure 1).

Table 1. Setup over double restriction enzyme digest with FastDigestion HindIII and XbaI

FastDigest buffer 10x	2 µl
Plasmid DNA (from tube 1)	15 µl
FastDigest HindIII	1 µl
FastDigest XbaI	1 µl
Water	1 µl

Table 2. Setup over restriction enzyme digestion with FastDigestion XhoI.

FastDigest buffer 10x	2 µl
Plasmid DNA (from tube 1)	15 µl
FastDigest XhoI	2 µl
Water	2 µl

Table 3. Setup over restriction enzyme digestion with FastDigestion EcoRI.

FastDigest buffer 10x	2 µl
Plasmid DNA (from tube 1)	15 µl
FastDigest EcoRI	2 µl
water	2 µl

6 LA plates were prepared for transformation of pMSP2N2 into the expression system BL21 (DE3) / pLysS. For two plates, 50 ml LA containing 50 µl of 50 mg/ml Kanamycin was used. The transformation was done in accordance with standard protocols from the research group. Competent cells (Bl21 (DE3) / pLysS) were allowed to thaw for 15 min on ice, 3 eppendorf tubes each containing 100 µl of cells was marked as follows: 0 (negative control), 1 and 2. Purified plasmid DNA from tube 1 versus 2 as well as water was added according to table 4. The tubes were incubated for 1h on ice before heat shock was applied in a 42°C water bath for 42 seconds. The cells were then cooled for 5 minutes on ice. 900 µl of NZY medium and

$MgSO_4$/$MgCl_2$ glucose were added to each tube (0,1 and 2) before incubation for 1h at 37°C in a shaker.

Table 4. Amount of purified plasmid DNA added to 100 µl of cells for transformation of pMSP2N2 into the expression system BL21(DE3) / pLysS.

	Purified plasmid DNA tube 1	Purified plasmid DNA tube 2	Water
0	-	-	2 µl
1	2 µl	-	-
2	-	2 µl	-

100 µl of each tube (0,1 and 2) were added to each of the three previously made plates marked as follows: 0 (100 µl), 1 (100 µl) and 2 (100 µl). Remaining cell mixture in tubes 0,1 and 2 were centrifuged at 4000 rpm for 5 minutes. Most of the supernatant was poured off, some was saved and used for resuspension of the pelett and the liquid in each tube (0, 1 and 2) was added to the remaining three previously made plates marked as follows: 0 (remaining), 1 (remaining) and 2 (remaining). All 6 plates were then incubated overnight at 37°C. Colonies on each plate were counted and inspected (see table 11). 1 colony from plate 1 (remaining) was transferred to a new plate labeled as “transformant 1” and 1 colony from plate 2 (remaining) was transferred to a new plate labeled as “transformant 2”.

Control cleavage of transformants is performed to verify that pMSP2N2 is present in BL21 (DE3) / pLysS. Two new starter cultures were prepared and the plasmid DNA was purified as described above. One starter culture was inoculated with a colony from plate 1 (remaining) and the other starter culture was inoculated with one colony from plate 2 (remaining). Once the DNA was extracted, DNA concentration was checked using nanodrop (see table 12). Both products from the two plasmid preparations were cleaved with both FastDigest EcoRI and FastDigest XhoI and the digestion was prepared in two tubes in accordance with the manufacturer's recommendations (ThermoFisher Scientific, Sweden) (see table 5 and 6). The tubes were placed for incubation in a water bath at 37°C for 15 min. 20 µl from each tube together with 4 µl loading dye (gelstar included, 1 µl gelstar to 1 ml 6x loading dye) was charged to a gel (see figure 2).

Table 5. Setup over restriction enzyme digestion with FastDigestion EcoRI. One setup was made with purified plasmid DNA from tube 1 and another setup was made from tube 2.

FastDigest buffer 10x	2 µl
Plasmid DNA (tube 1 vs tube 2)	14 µl
FastDigest EcoRI	2 µl
Water	2 µl

Table 6. Setup over restriction enzyme digestion with FastDigestion XhoI. One setup was made with purified plasmid DNA from tube 1 and another setup was made from tube 2.

FastDigest buffer 10x	2 µl
Plasmid DNA (tube 1 vs tube 2)	14 µl
FastDigest XhoI	2 µl
Water	2 µl

MSP2N2 expression

MSP2N2 expression was done according to Ritchie et al. (2009) page 216. A starter culture of 30 mL LB and 18 µl of 50 mg/ml Kanamycin was prepared and inoculated with a colony from plate “transformant 2” to then be incubated for 5 hours. Growth was stopped by placing the tube in the refrigerator (4°C) overnight. The fermenter was prepared and a sterile solution of 250 ml containing 0.17 M KH_2PO_4 and 0.72 M K_2HPO_4 was added together with 0.5 ml of 50 mg/ml Kanamycin and a few drops of antifoam before the fermenter was inoculated with the prepared starting culture. The pH of the fermenter was set between 7.2-7.4. Four hours from start, optical density (OD) is inspected, then the OD is checked every hour (see table 13). For each sample taken to check optical density, 1 ml is saved and stored at -20°C to be able to run a western blot. After 4h and 10 min, protein expression was induced by the addition of 3.4 ml of 0.8 M IPTG. After another two hours, the cells are harvested by centrifugation in a Jla-8.1000 rotor at 6000 rpm for 20 min. The generated cell pellet was weighted and stored at -80°C.

The tubes with samples from the fermenter from 4, 5 and 6 hours after start were centrifuged in an eppendorf centrifuge at 15 000 rpm for 5 min and the supernatants were poured off (pelett from the tube associated with 6h were lost). To 5 ml of 20 mM phosphate buffer, 50 µl of 100 mM PMSF, 50 µl of Triton-X-100, 0.125 mg of deoxyribonuclease and a pinch of lysozyme was added. To the tubes corresponding to 4h and 5h after start, 1 ml of the previous buffer was added to resuspend the pellets and the tubes were then incubated for

10 minutes before being stored at -20°C.
Western blot was performed on samples from the fermenter in accordance with standard protocols in the research group. The gel used was a Mini-Protean TGX Precast gel 10%, 10-well comb, 30 µl/well (Bio-Rad, Sweden). A master mix containing 33 µl of the sample from the fermenter 4h from start and 16.5 µl solubilization buffer was prepared, from this mastermix, 27 µl were taken and charged to one well, and 15 µl was charged to another well. In the same way a master mix was prepared with the sample corresponding to 5h from start, also here 27 µl and 15 µl were charged. After the samples had migrated on the gel for 35 min, the proteins were transferred to a membrane using the Trans-Blot Turbo Transfer system (Bio-Rad, Sweden). The membrane was then blocked for 50 minutes using 0.6g BSA in 20ml 1xTBS before washing the membrane with 1xTBS for 2x5 minutes. 2 µl of the primary antibody (6xHis monoclonal clontech 631212 from mouse) was added to 20 ml 1xTBS with 0.6 g BSA which was then added to the membrane and placed on a shaking table overnight. The membrane was washed with TBS-T for 3x10min. 2 µl of the secondary antibody (Anti-mouse lgG (whole molec.) Alkaline phosphatase Sigma-Aldrich A9316 from goat) was added to 20 ml TBS-T with 0.6 g BSA and added to the membrane which was placed on a shaking table for 2h. The membrane was washed with TBS-T for 3x10 minutes before development which was done by adding 10 µl NBT and 5 µl BCIP to 15 ml developing buffer and pipetting this over the membrane in a dark room. The membrane was documented (see figure 3).

MSP2N2 purification with Immobilized metal affinity chromatography (IMAC) (Attempt 1)

For purifications of MSP2N2, protocol from Ritchie et al. (2009) page 217 is used. 200 ml of 20 mM phosphate buffer was used to resuspend the cell pellet from the fermenter (84 g), to which PMSF, 1% Triton-X-100 and deoxyribonuclease 1 were added. The resuspension was sonicated and centrifuged in a JA-25-50 rotor at 30 000xg and the supernatant was poured off and stored at 8°C (total of 200 ml). The column was prepared with 2 bed volumes (corresponding to 4 ml) of 0.1 M nickel(II)-chloride-hexahydrate, followed by 2 bed volumes of water. 10 ml of 40 mM phosphate buffer was drawn through the column to equilibrate it before 10 ml of lysate (from the 200 ml supernatant) was added to the column. When the lysate passed through, 1 ml of flow through was collected and stored at -20°C. The column was then washed with 10 ml of each: 40 mM Tris/HCl, 0.3 M NaCl, 1% Triton-X-100; 40 mM Tris/HCl, 0.3 M NaCl, 50 mM Na-cholate, 20 mM Imidazole and 40 mM Tris/HCl, 0.3 M NaCl, 50 mM Imidazole. Of all washes, 1 ml flow through was collected and stored at -20°C. For the elution, 10 eppendorf tubes were marked with E0-E9. 1 ml of 40 mM Tris/HCl, 0.3 M NaCl, 0.4 M Imidazole was added to the column, this was repeated 10 times and each fraction was collected in the corresponding labeled tube. The protein concentration in each fraction was checked using nanodrop (see table 14).

Flow through of the lysate as well as flow through of the third wash (40 mM Tris/HCl, 0.3 M NaCl, 50 mM Imidazole) and flow through of eluate 0-5 from the purification was used to charge to a gel which was then stained with coomassie in accordance with standard protocols in the research group. The gel used was a Mini-Protean TGX Precast gel 10%, 10 well comb, 30 µl/well (Bio-Rad, Sweden). 20 µl of each sample together with 10 µl solubilization buffer were prepared and 27 µl of each was charged to the gel. After the samples had migrated on the gel for 30 minutes, it was stained with coomassie overnight before being documented (see figure 4).

Size exclusion chromatography (SEC) was performed on E1 and E2 to see if it is possible to make the sample a little less complex, this is done in accordance with standard protocols in the research group using the NGC 10 Medium Pressure Chromatography system (Bio-Rad, Sweden). A buffer of 0.5 L containing 20 mM Tris/HCl, 0.1 M NaCl, 0.5 mM EDTA was prepared. All solutions that would be used for the SEC were filtered with a 0.45 µm syringe filter and were then de-gased under vacuum. E1 and E2 were pooled to be concentrated and filtered before the SEC, which was done using an Amicon Ultra-4 Centrifugal filter Unit (30kDa cutoff) (Sigma-Aldrich, Sweden) and a 1 ml syringe with a 0.45 µm filter. The capped filter device was centrifuged at 3220xg for 4 min until 500 µl remained in the filter. The sample was then filtered with the 1 ml syringe and 0.45 µm filter before being collected in an eppendorf tube. Protein concentration on what went through the filter versus what did not go through the filter was checked using nanodrop (see table 15). The column used was a SuperDex 200 10/300GL (GE healthcare, Sweden) with a flow of 0.4 ml/min and 1 column volume corresponding to 23.46 ml. The column was equilibrated with 3 column volumes of 20 mM Tris/HCl, 0.1 M NaCl, 0.5 mM EDTA and then the injection loop was washed with 5 ml water before it was filled with 5 ml Tris/HCl, 0.1 M NaCl, 0.5 mM EDTA. The sample containing E1+E2 (100 µl) was then loaded and the sample collector was put in the correct position before the chromatography program was started. The chromatogram was evaluated.

MSP2N2 purification trial number 2

Purification is done according to Ritchie et al. (2009) page 217, but in comparison with trial 1, the purification was performed with 5 times as much lysate and the lysate is allowed to bind overnight instead. The column was equilibrated by passing a bed volume of 2 ml 0.1 M nickel(II)-chloride-hexahydrate through the column followed by 4 ml (two bed volumes) of water. 10 ml 40 mM phosphate buffer was added to the nickel agarose to resuspend it and the nickel agarose was then moved in equal parts to two 50 ml falcon tubes. To the falcon tubes, 25 ml lysate was added in each and the falcon tubes were then shaken in a cold room overnight. The two falcon tubes were then centrifuged at 1300xg for 10 seconds. The supernatants were poured off and stored at -20°C to be able to charge to a gel to ensure that protein has not been lost, however, some fluid was saved to resuspend the agarose and bring it back to the column. The lysate that had bound overnight was allowed to pass through the column and 1 ml of flow through was collected and stored at -20°C. The column was washed with 10 ml of 40 mM Tris/HCl, 0.3M NaCl, 1% Triton-X-100; 10 ml of 40 mM Tris/HCl, 0.3

M NaCl, 50 mM Na-cholate, 20 mM Imidazole and 10 ml of 40 mM Tris/HCl, 0.3 M NaCl, 50 mM Imidazole. 1 ml of flow through from each wash was collected and stored at -20°C. 10 eppendorf tubes were marked with E0-E9 for the elution. 1 ml of 40 mM Tris/HCl, 0.3 M NaCl, 0.4 M Imidazole was added to the column, this was repeated 10 times and each fraction was collected in the corresponding labeled tube. The protein concentration in each fraction was checked using nanodrop (see table 16).

A western blot and a gel stained with coomassie was done on flow through of the lysate, the third wash and E1-E3 in accordance with standard protocols in the research group. Gels used were Mini-Protean TGX Precast gels 10%, 10-well comb, 30µl/well (Bio-Rad, Sweden). The samples (flow through of the lysate, wash number 3 and E0-E3) were prepared (x2) by taking 20 µl sample together with 10 µl solubilization buffer. Of E1 and E2, two fractions of each were made where one fraction of each was boiled by placing the two samples on a heating block at 100°C for 5 minutes before being charged to the wells. 27 µl of each sample was loaded to the wells of the two gels and allowed to migrate for 35 min. One gel was taken for staining with coomassie overnight before being documented (see figure 5), the imprint on the other gel was transferred to a membrane which was then blocked for 50 min before the primary and secondary antibodies were added and development with BCIP/NBT was performed (see MSP2N2 purification with Immobilized metal affinity chromatography (IMAC) (Attempt 1)). The membrane was documented (see figure 6).

SEC was conducted to try to find out the size of the different aggregates that we can see on western blot/gel with coomassie staining. E1 and E2 from the second purification were pooled and concentrated from 2 ml to 0.7 ml to obtain 10 mg/ml protein (7.019 mg/0.7 ml), this was done using “Vivaspin 15R centrifugal concentrator” (Sigma-Aldrich) and “Spin-X 0.45µm Costar1" (Sigma-Aldrich, Sweden). Pooled E1 and E2 were added to the Vivaspin concentrator and centrifuged at 3000xg for 5 minutes. Protein concentration on what went through the filter versus what did not go through the filter was checked using nanodrop (see table 17). What did not pass through the filter (E1+E2) was spin filtered using “Spin-X 0.45 µm Costar1" until the sample visibly passed through. “Spin-X 0.45 µm Costar1” was used twice because the first time the filter was used it was clogged and the flow stopped, so what actually went through the filter both times was pooled (about 200 µl) and protein concentration was checked using nanodrop (see table 18), 20 µl was collected and stored at -20°C. What did not pass through the filter during the first attempt was lost, but what did not pass through the second time was collected and protein concentration was checked using nanodrop (see table 18) and 20 µl was saved and stored at -20°C. SEC was done in accordance with standard protocols in the research group (see above). The sample containing flow through from the Spin-X was loaded (100 µl). The chromatogram was evaluated.

A gel stained with coomassie was performed in accordance with standard protocols in the research group (see above) on flow through of the lysate, E0 and E3 from the elution, flow through from the concentration with Vivaspin, flow through from the Spin-X column, the sample that did not go through the Spin-X filter and fraction 29 and 30 from the SEC. The samples were prepared according to table 7. The gel used was a Mini-Protean TGX Precast

gel 10%, 12 well comb, 20µl/well (Bio-Rad, Sweden) and 15 µl of each sample was charged to the gel (see figure 7).

Table 7. Preparation and composition of samples for coomassie gel after SEC (second purification).

	Amount of sample (µl)	water (µl)	Solubilization buffer (µl)
Lysate Flow through	2	18	10
E0	20	-	10
E3	20	-	10
E1+E2 Flow through (Vivaspin)	20	-	10
E1+E2 flow through (Spinnex)	16	4	10
E1+E2 which did not pass through the filter (Spinnex)	9	11	10
A29	20	-	10
A30	20	-	10

To evaluate degradation, a western blot was also run, this was done in accordance to standard protocols in the research group (see above) on samples E0, E3, flow through of the lysate, E1+E2 (that did not go through the Spin-X column) and fraction A29+A30 from the SEC. The gel used was a Mini-Protean TGX Precast gel 10%, 12-well comb, 20µl/well (Bio-Rad, Sweden). The samples were prepared according to table 8 and 15 µl of each sample was charged (see figure 8).

Table 8. *Preparation and composition of samples for western blot after SEC (second purification).*

	Amount of sample (µl)	water (µl)	Solubilization buffer (µl)
Lysate Flow through	2	18	10
E0	20	-	10
E3	20	-	10
E1+E2	9	11	10
A29	20	-	10
A30	20	-	10

For the sake of comparison, a western blot was performed in accordance with standard protocols in the research group (see above) on samples before induction with IPTG, after induction with IPTG, cell lysate (after cultivation in the fermenter) and E1+E2 from the second purification (what did not go through the Spin-X filter). The gel used was a Mini-Protean TGX Precast gel 10%, 10-well comb, 30µl/well (Bio-Rad, Sweden). The samples were prepared according to table 9 and 27 µl of each sample was charged (see figure 9).

Table 9. *Preparation and composition of samples for western blot.*

	Amount of sample (µl)	water (µl)	Solubilization buffer (µl)
Sample before induction	33	-	16.5
Sample after induction	33	-	16.5
Cell lysate	2	18	10
E1+E2 (what did not go through the Spin-X filter)	9	11	10

To try to get a clarity in what the dominant bands could be that we have seen several times after purification of MSP2N2 around 30kDa as well as over 70kDa and 100kDa, these three bands (see figure 7) are sent away for mass spectrometry performed by Katja Bernfur (in house service) (Appendix 2).

Appendix 2.

Result of the analysis of the three gel bands (figure 7) by mass spectrometry.

Gel band 1:

Database Search based on MS data

MASCOT Search Results

Protein View: SLYD_ECO57

FKBP-type peptidyl-prolyl cis-trans isomerase SlyD OS=Escherichia coli O157:H7 GN=slyD PE=3 SV=1

Database:	**SwissProt**
Score:	**78**
Expect:	**0.0081**
Nominal mass (M_r):	**20840**
Calculated pI:	**4.86**

Taxonomy: Escherichia coli O157:H7
Sequence similarity is available as an NCBI BLAST search of SLYD_ECO57 against nr.

Search parameters

Enzyme:	Trypsin: cuts C-term side of KR unless next residue is P.
Variable modifications:	Oxidation (HW), Oxidation (M)
Mass values searched:	22
Mass values matched:	10

Protein sequence coverage: 37%

Matched peptides shown in ***bold red***.

```
  1 MKVAKDLVVS LAYQVRTEDG VLVDESPVSA PLDYLHGHGS LISGLETALE
 51 GHEVGDKFDV AVGANDAYGQ YDENLVQRVP KDVFMGVDEL QVGMRFLAET
101 DQGPVPVEIT AVEDDHVVVD GNHMLAGQNL KFNVEVVAIR EATEEELAHG
151 HVHGAHDHHH DHDHDGCCGG HGHDHGHEHG GEGCCGGKGN GGCGCH
```

Gel band 2:

Database Search based on MS data

MASCOT Search Results

Protein View: ARNA_ECOBW

Bifunctional polymyxin resistance protein ArnA OS=Escherichia coli (strain K12 / MC4100 / BW2952) GN=arnA PE=3 SV=1

Database:	**SwissProt**
Score:	**199**
Expect:	**6.9e-015**
Nominal mass (M_r):	**74242**
Calculated pI:	**6.39**

Taxonomy: Escherichia coli BW2952

Sequence similarity is available as an NCBI BLAST search of ARNA_ECOBW against nr.

Search parameters

Enzyme:	**Trypsin: cuts C-term side of KR unless next residue is P.**
Variable modifications:	**Oxidation (HW), Oxidation (M)**
Mass values searched:	**30**
Mass values matched:	**20**

Protein sequence coverage: 45%

Matched peptides shown in *bold red*.

```
  1 MKTVVFAYHD MGCLGIEALL AAGYEISAIF THTDNPGEKA FYGSVARLAA
 51 ERGIPVYAPD NVNHPLWVER IAQLSPDVIF SFYYRHLIYD EILQLAPAGA
101 FNLHGSLLPK YRGRAPLNWV LVNGETETGV TLHRMVKRAD AGAIVAQLRI
151 AIAPDDIAIT LHHKLCHAAR QLLEQTLPAI KHGNILEIAQ RENEATCFGR
201 RTPDDSFLEW HKPASVLHNM VRAVADPWPG AFSYVGNQKF TVWSSRVHPH
251 ASKAQPGSVI SVAPLLIACG DGALEIVTGQ AGDGITMQGS QLAQTLGLVQ
301 GSRLNSQPAC TARRRTRVLI LGVNGFIGNH LTERLLREDH YEVYGLDIGS
351 DAISRFLNHP HFHFVEGDIS IHSEWIEYHV KKCDVVLPLV AIATPIEYTR
401 NPLRVFELDF EENLRIIRYC VKYRKRIIFP STSEVYGMCS DKYFDEDHSN
451 LIVGPVNKPR WIYSVSKQLL DRVIWAYGEK EGLQFTLFRP FNWMGPRLDN
501 LNAARIGSSR AITQLILNLV EGSPIKLIDG GKQKRCFTDI RDGIEALYRI
551 IENAGNRCDG EIINIGNPEN EASIEELGEM LLASFEKHPL RHHFPPFAGF
601 RVVESSSYYG KGYQDVEHRK PSIRNAHRCL DWEPKIDMQE TIDETLDFFL
```

651 RTVDLTDKPS

Gel band 3

Database Search based on MS data

MASCOT Search Results

Protein View: HFQ_HAMD5

RNA-binding protein Hfq OS=Hamiltonella defensa subsp. Acyrthosiphon pisum (strain 5AT) GN=hfq PE=3 SV=1

Database:	**SwissProt**
Score:	**80**
Expect:	**0.0057**
Nominal mass (M_r):	**10204**
Calculated pI:	**9.25**
Taxonomy:	**Candidatus Hamiltonella defensa 5AT (Acyrthosiphon pisum)**

Sequence similarity is available as an NCBI BLAST search of HFQ_HAMD5 against nr.

Search parameters

Enzyme:	**Trypsin: cuts C-term side of KR unless next residue is P.**
Variable modifications:	**Oxidation (HW), Oxidation (M)**
Mass values searched:	**22**
Mass values matched:	**5**

Protein sequence coverage: 49%

Matched peptides shown in *bold red*.

1 MAKGQSLQDP FLNALRKERI PVSIYLINGI KLQGQIESFD QFVILLKSTS

51 PSLKNTICQV IYKHAVSTVV PSRAVSSSPH HAPPLQKPED DGS