UNIVERSITY OF PÉCS

Doctoral School of Chemistry

PhD Thesis

Pál Salamon

Supervisor: Dr. Beáta Albert D.Sc

Pécs, 2022

1. Introduction

From gene cloning to protein purification, the molecular tools used in all steps of the biosynthesis of recombinant proteins are available in many alternatives. Despite this abundance, failure of functional recombinant protein biosynthesis due to protein toxicity to the host or aggregation in inclusion bodies is not uncommon.

Various microbial systems are now widely used to produce recombinant proteins. Even though Escherichia coli (*E. coli*) is the most used host for the biosynthesis of recombinant proteins, as a prokaryotic system it has many limitations, of which the aggregation of foreign proteins into insoluble inclusion bodies is the main limiting factor. On the other hand, the advantages of *E. coli*'s low maintenance costs and simple genetic pool are essential for the design of new systems aimed at the overproduction of recombinant protein in the host cell. Gene and protein technology processes enable the design of new, personalized strategies according to the user or process requirements. Fusion technologies are widely used in *E. coli* expression systems to improve soluble recombinant protein production and/or purification. The use of the ubiquitin fusion partner enables the specific cleavage of the target protein and the biosynthesis of the affinity tag-free native protein in addition to protection against proteases and stability during the production of low molecular weight peptides.

This PhD thesis discusses strategies for the development of recombinant protein production in the *E. coli* expression system (Figure 1): 1. application of ubiquitin fusion for the biosynthesis of SMAC and BUFII proteins in their native form; 2. Denaturation of GST-XIAP inclusion bodies (aggregates), followed by its refolding to produce the active form of the protein; 3. application of autoinduction against the aggregation of AURKA into insoluble inclusion bodies. To adjust the conditions of the auto-induction fermentation, we chose EGFP, a comprehensively studied recombinant protein, which results in efficient active protein production in dissolved form even during traditional, IPTG-based induction.

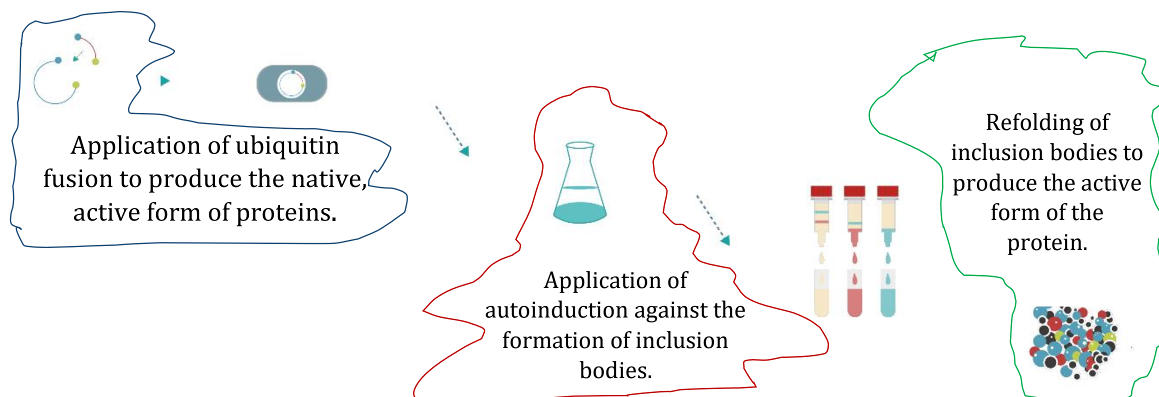

Figure 1: Strategies for the development of recombinant protein production in the E. coli expression system discussed in the thesis

2. Objectives

Pure, soluble, and functional proteins are in high demand in modern biotechnology. Natural protein sources rarely meet the requirements for quantity, ease of isolation, or price, so the recombinant technology is often the method of choice. *E. coli* is a commonly used host because of its relative simplicity, cheap and rapid high-density cultivation, well-known genetics, and large number of compatible molecular tools that contribute to efficient heterologous expression of the recombinant protein. Despite all these properties, expression of recombinant proteins with the *E. coli* host often results in insoluble and/or non-functional proteins. This PhD thesis discusses the following strategies for the development of recombinant protein production in the *E. coli* expression system: 1. application of ubiquitin fusion for the biosynthesis of SMAC and BUFII proteins in their native form; 2. denaturation of GST-XIAP inclusion bodies (aggregates), followed by its refolding to produce the active form of the protein; 3. application of autoinduction against the aggregation of AURKA into insoluble inclusion bodies.

During the synthesis of peptides containing 20-50 amino acids, it may happen that they do not adopt a suitable conformation and are less resistant against the activity of proteases. Our goal is to use a ubiquitin fusion tag in the case of a low molecular weight peptide (BUFORIN II) to increase the stability of the recombinant protein.

Our aim is to use the ubiquitin fusion tag in the case of a low molecular weight peptide (BUFORIN II) or a low molecular weight protein (SMAC) to specifically cleave the target protein. Affinity tags are often used to facilitate the detection and purification of expressed proteins. The most used affinity tag is the hexa/deca histidine sequence, which binds the fusion protein to the Ni-nitrilotriacetate matrix. Bacterial proteins can be removed by washing, and then the pure protein can be eluted from the matrix with an imidazole solution. In expression vectors, a protease cleavage site is also incorporated between the affinity tag and the target protein. Commonly used proteases are thrombin or factor Xa. During the use of these enzymes, it is possible that the target protein also undergoes proteolysis. Since proteases usually cut in the middle of the cleavage sequence, the amino terminal sequence of the target protein is usually changed after the affinity tag is removed. The solution to the problems mentioned above is the use of a ubiquitin fusion partner. The cleavage of ubiquitin from ubiquitinated proteins at their amino terminus is performed by one of the ubiquitin hydrolases, such as the YUH1 enzyme (ubiquitin carboxyl-terminal hydrolase) used in this thesis. This enzyme is specific for the ubiquitin sequence and is not sensitive to the amino-terminal sequence of the target protein.

Our goal is to use the ubiquitin fusion tag in the case of a low molecular weight peptide (BUFORIN II) or a low molecular weight protein (SMAC) for the biosynthesis of an affinity tag-free target product.

Another objective of this thesis is to investigate the applicability of autoinduction fermentation for the active production of the recombinant AURKA protein formed in the inclusion body. To adjust the conditions of the auto-induction fermentation, we chose EGFP, a comprehensively studied recombinant protein, which results in efficient active protein production in dissolved form even during traditional, IPTG-based induction.

3. Materials and methods

Biosynthesis of SMAC and BUFII recombinant proteins

The pUbiq expression vector was a gift from the research group of prof. Dr. András Perczel (MTA-ELTE Protein Modelling Research Group and Structural Chemistry and Biology Laboratory). The designed oligonucleotides (primers) were manufactured by GeneriBiotech (Debrecen). Amplification of the SMAC gene from the pET20b_SMAC (Salamon et al., 2017) construct was performed by PCR (PRO FLEX PCRS System, Life Technologies). Snapgene software (version 1.1.3) was used for DNA sequence visualization and editing (https://www.snapgene.com/).

The PCR product was digested with restriction endonucleases SacII (Thermo Scientific) and BamHI (Thermo Scientific). The creation of the phosphodiester bond between the sticky ends of the SMAC inserts and the sticky ends of the pUbiq plasmid were carried out using T4 DNA ligase (Thermo Scientific), thus creating the pUbiq-SMAC vector DNA construct. Sequencing of the vector construct was performed by GeneriBiotech (Debrecen). Separation of PCR products and digestions with restriction endonucleases was performed by agarose gel electrophoresis, with 20,000-fold dilution stain (RedSafe Nucleic Acid Staining Solution; iNtRON Biotechnology, Korea) on a 1% agarose (Lonza SeaKem® LE) gel.

The bacterial culture was incubated in a Sartorius BIOSTAT® A plus 1-liter bioreactor, with the following operating parameters: the cells were cultured at 37°C in LB medium supplemented with 50 mg/L kanamycin for the selection and maintenance of *E. coli* transformants; mixing speed 300 rpm; air flow rate 0.2 vvm. The cell growth was monitored by optical density measurements at λ=600 nm with a UV-Vis spectrophotometer (GeneQuant pro Spectrophotometer, GE Healthcare). The expression of 10xHis-UBI_SMAC was induced by adding 0.5 mM IPTG at an optical density of OD600=0.8, the duration of the expression was 5 hours.

The bacterial culture was collected by centrifugation (SL-40R, Thermo Scientific) at 4000 g for 10 min. The cells were suspended in 20 mL lysis buffer, and after ultrasonic lysis, they were centrifuged at 60,000 g for 60 minutes (Sorvall Lynx 6000, Thermo Scientific). The supernatant was stored at 4°C during the purification process. The (dissolved) proteins in the supernatant were separated on SDS-PAGE (12% acrylamide). Protein Marker I (peqGOLD, VWR) was used to identify the recombinant protein according to its size.

The 10xHis-UBI_SMAC recombinant protein expressed in the bioreactor was purified by FPLC (ÄKTA purifier, GE Healthcare) on a HisTrap (GE) 5 mL affinity chromatography column. The buffer system contains Wash I and Wash II buffers to minimize non-specific interactions between the proteins and the column. The absorbance of the eluted proteins was monitored with a UV detector (UPC-900, GE Healthcare) at a wavelength of 280 nm.

After successful purification of 10xHis-UBI_SMAC, the protein complex was digested with 1U YUH1 for 4 hours. The YUH1 digestion buffer contained 20 mM Tris-HCl, 1 mM DTT, and 0.5 mM EDTA. The digestion reaction was verified on a 10% SDS-PAGE, and the resulting image was analyzed with GelAnalyzer 19.1 software (http://www.gelanalyzer.com/). To isolate the native SMAC, we used another affinity chromatography purification, thus removing the 10xHis-ubiquitin and the YUH1 enzyme from the mixture.

The absorbance of the eluted proteins was monitored with a UV detector (UPC-900, GE Healthcare) at a wavelength of 280 nm.

The 3D structure of the 10xHis-UBI_SMAC protein complex was performed using a homology-based modelling algorithm. The Phyre2 online search engine was used to search for fusion recombinant protein sequence homologues and to construct the model (http://www.sbg.bio.ic.ac.uk/~phyre2/). During the creation of the model, we used the following structures as templates: 3L0W, 5Y3T, 1FEW, 3AI5, 1OGW, 1OQY, 6PX3, 1YX5. Using Phyre2, the spatial structure of the fusion protein can be determined (with high accuracy) based on the existing known crystal structures from the appropriate templates (DeLano, 2002; Kelley et al., 2015). To display the 3D structure, we used PyMOL 2.0 software (https://pymol.org/2/).

Recombinant GST-XIAP used for the in vitro study of the interaction between SMAC and GST-XIAP was obtained from commercial sources (Merck). The Glutathione Sepharose 4b resin was equilibrated with the following reaction buffer (50 mM Tris-HCl, 100 mM NaCl, 1 mM DTT, pH 7.4). GST pulldown experiments were performed at room temperature. GST-XIAP proteins (8 µg) and recombinant SMAC (8 µg) were incubated with glutathione beads for 1–2 h in the reaction buffer, followed by three washing steps. Bound proteins were eluted with 10 mM free reduced glutathione in Tris buffer. GST (8 µg) was used as a negative control. The eluted samples were verified on an 8% SDS-PAGE.

Due to the small size of BUFORIN II (21 amino acids, 63 bp), the insert (gene) required for the cloning is two single-stranded oligonucleotides (primers). A sequence recognized by the SacII enzyme was designed for the N-terminal-, and a sequence recognized by the BamHI enzyme on the C-terminal site of the BUFORIN II peptide. The designed oligonucleotides were produced by GeneriBiotech (Debrecen). Snapgene software (version 1.1.3) was used for DNA sequence visualization and editing (https://www.snapgene.com/). The insert was produced by hybridization of the two single-stranded oligonucleotides (DNA sequence encoding BUFORIN II peptide). Hybridization was performed in TE buffer (100 µL): 25 µL of "forward" (5') and 25 µL of "reverse" (3') sequences were added to the reaction mixture, which was supplemented with 4 µL of 1 M NaCl solution. The reaction mixture was placed in a PCR apparatus, which was slowly heated to 95°C, held at that temperature for 10 minutes, and then cooled to room temperature at low speed.

E. coli Rosetta™ (DE3)pLysS strain was used as a protease-deficient host cell for the expression of the 10xHis-UBI_BUFII protein. The recombinant protein was produced in large quantities in a Sartorius Biostat®A Plus bioreactor. The fermentation conditions were controlled using the BioPAT®MFCS/DA Supervisory Control software. The bioreactor was filled with 0.7 L of 2YT nutrient broth (16 g/L Triptone, 10 g/L yeast extract, 5 g/L NaCl), and then the entire system was sterilized in an autoclave (120 °C, 20 min). The fermentation parameters were: 400 RPM, 37 °C, pH=6.9. After the system was stabilized, the reactor was inoculated, and then the cell culture was grown at 37 °C, pH 6.9 and a dissolved oxygen level above 40% until an optical density of 26 (600 nm) was reached. The optical density was measured with a CamSpecM330 UV-VIS spectrophotometer. Protein expression was induced with isopropylthiogalactopyranoside (IPTG) at a final concentration of 1 mM. The protein expression took place at 37°C for 4 hours. The cells were collected by centrifugation (12000xg, 10 minutes, 4 °C, Sorvall LYNX 6000 Ultracentrifuge), and their analysis was performed using a Microfluidizer LM10 instrument. The composition of the lysis buffer used for detection is the following: 20 mM Tris-HCl (pH=8), 250 mM NaCl,

2 mM dithiothreitol (DTT), 1 mM phenylmethylsulfonyl fluoride (PMSF), 1X protease inhibitor (Thermo Scientific). The extract obtained during the exploration was centrifuged (60000xg, 60 minutes, 4 °C, Sorvall LYNX6000 Ultracentrifuge) for the removal of cell debris.

Affinity chromatographic purification was performed on a 5 mL HisTrap (GEHealthcare) column on an AKTA FPLC (Amersham Pharmacia Biotech) device. System control and data collection were performed using UNICORN 5.11 software. The composition of the washing buffer: 20 mM Tris-HCl (pH=8), 250 mM NaCl, 2 mM DTT, 1 mM PMSF, 20 mM imidazole; composition of the elution buffer: 20 mM Tris-HCl (pH=8), 250 mM NaCl, 2 mM DTT, 1 mM PMSF, 500 mM imidazole. The purification parameters used are: 1 mL/min flow rate, 1 mL/fraction. During the elution, step gradient and linear gradient separations were used. The protein solution obtained during the purification was subjected to dialysis, which lasted for 24 hours, at 4 ºC, with stirring, in the following buffer composition: 20 mM Tris-HCl (pH=8), 250 mM NaCl, 2 mM DTT, 1 mM PMSF. After successful purification of 10xHis-UBI_BUFII, the protein complex was digested with 1 U of YUH1 according to the conditions of the 10xHis-UBI_SMAC complex.

Refolding of GST-XIAP inclusion bodies

The pGEX-GST-XIAP recombinant plasmid was a gift from Dr. Attila Reményi (MTA TTK, Institute of Organic Chemistry, Protein Interaction Research Group, Budapest, Hungary). During our experiments, we expressed the entire XIAP protein. Recombinant GST-XIAP used as a control for Western blot analysis was obtained from commercial sources (Merck). The electrophoresis apparatus was from Bio-Rad (Bio-Rad Mini ProteanTetraCell). A Sorvall LYNX 6000 centrifuge (Thermo Scientific) was used for centrifugation. Batch purification of proteins was performed using Glutathione Sepharose 4B affinity chromatography resin (GE Healthcare). All chromatographic experiments were performed with an AKTA FPLC system (Amersham Pharmacia Biotech). HiLoad 16/600 Superdex 75 pg column was purchased from GE Healthcare. A 2100 Bioanalyzer Instrument (Agilent) and a Protein 250 Kit (Agilent) were used for the chip electrophoresis tests.

The recombinant plasmid pGEX-GST-XIAP was transformed into competent *E. coli* BL21 (DE3) Rosetta plysS bacterial cells and plated on Luria Bertani (LB) medium containing 100 µg/mL ampicillin and 34 µg/mL chloramphenicol. A single colony was used to inoculate 5 ml of LB-ampicillin medium and grown for 8 h at 37°C at 250 rpm. Then, 1 mL of inoculum culture was transferred to 200 mL of M9 minimal medium containing 100 µg/mL ampicillin and 34 µg/mL chloramphenicol and cultured under the same conditions until the OD_{600} value reached 0.7. The culture temperature was reduced to 18°C and 0.2 mM IPTG was added to initiate recombinant protein expression. The bacterial culture was incubated for 8 hours at 18°C. Cells were harvested by centrifugation at 4500 rpm for 10 min at 4°C. The total protein content was analysed on a 10% (w/v) SDS-PAGE.

3 g of extracted cells were suspended in 30 mL of lysis buffer (50 mM Tris HCl, 200 mM NaCl and 2 mM DTT, pH 7.8). Lysozyme (1 mg/mL) was added to the cell suspension and incubated for 30 min at 4°C. Afterwards, DNase I (0.1 mg/mL final concentration) was added to the cell suspension. After 30 min of incubation, the cells were homogenized with a high-pressure homogenizer (LM10 Microfluidizer) at 18,000 psi for two cycles. The cell lysate was centrifuged for 20 min at 15,000 x g, 4°C, and the pellets were collected and washed twice with ultrapure water.

Solubilization of the target protein was performed by adding N-Lauroyl-sarcosine at different concentrations (2%, 3%, 4% and 10%) to the lysis buffer. These suspensions were incubated at room temperature with gentle agitation for 1 h and then centrifuged at 10,000 x g for 15 min. The supernatant containing the target protein was subjected to an immediate refolding procedure. The solubilized protein fractions were analysed on a 10% (w/v) SDS-PAGE.

The qualitative analysis of the total protein was performed by protein chip analysis. The experiment was conducted according to the experimental description of the Protein 230 Kit (Agilent) using a Bioanalyzer 2100. The conditions for recombinant XIAP-GST folding were optimized using TritonX-100, CHAPS and OTG in different proportions. The 6 different conditions were: 1: 1% TritonX-100 and 0.61% (10 mM) CHAPS; 2: 2% TritonX-100 and 1.22% (20 mM) CHAPS; 3: 3% TritonX-100 and 4.83% (30 mM) CHAPS; 4: 1% TritonX-100 and 1.5% OTG; 5: 1% TritonX-100 and 1% OTG; 6: 0.61% (10 mM) CHAPS and 1% OTG.

The folding conditions were tested in three replicates. The detergents were carefully added to the solubilized protein at room temperature with slow mixing. The refolded proteins were purified on Glutathione Sepharose 4B resin. The solubilized proteins were incubated with the glutathione beads for 2 hours at room temperature with gentle mixing. Purification of the refolded proteins was performed according to the manufacturer's (GE Healthcare) instructions, followed by 10% SDS-PAGE analysis. Protein was eluted with 50 mM Tris-HCl buffer (pH 8.0) containing 20 mM reduced glutathione. Eluted fractions were stored at 4°C and analysed by 10% SDS-PAGE. Purified proteins were dialyzed for 48 hours in 50 mM Tris-HCl buffer (pH 7.8) with three buffer exchanges. The target protein was further purified by size exclusion chromatography on a HiLoad 16/600 Superdex 75 pg column at a flow rate of 1 mL/min. The column was first filled with 2 column volumes of 50 mM Tris-HCl, pH 7.8 buffer, and then the sample was injected. Elution was monitored at 280 nm. Eluted fractions were stored at 4°C and analysed on a 10% SDS-PAGE. All samples were analysed by SDS-PAGE based on the Laemmli protocol using a 5% (w/v) stacking gel and a 10% (w/v) resolving gel at a constant voltage of 120 V.

Western blot analysis (Mini Trans-Blot®, Bio-Rad) was performed to confirm GST-XIAP refolding. The recombinant GST-XIAP protein was separated by electrophoresis, using native PAGE, and then transferred to a PVDF membrane via wet transfer. The transfer buffer contained 25 mM Tris-HCl, 192 mM glycine, 0.7% acetic acid, pH=7.4. The membrane was incubated overnight at 4 °C with monoclonal anti-XIAP antibody, clone 2F1 (Abcam), diluted 1:1000 in TBS buffer. After washing four times with TBS for 15 minutes, the membrane was incubated with Goat Anti-Mouse IgG H&L (HRP) secondary antibody (Abcam) diluted 1:1000 in TBS. The next step includedincludedincluded four four four washing sstepps with TBS for 15 min each each. Finally, the protein was detected with Pierce™ ECL substrate (Thermo Scientific) and the membrane was visualized using X-ray film (dark room technique).

Application of autoinduction against aggregation of AURKA into insoluble inclusion bodies

During the preparation of the fermentations, we created pre-grown cultures. During our experiments, we expressed the entire AURKA protein. The first pre-cultivation was carried out in 10 mL of 100 mg/mL ampicillin-containing LB nutrient solution. The pre-cultivation of the cultures was carried out at 37℃, 250 rpm until the point when the cell density reached the value between OD_{600}= 0.1-0.2, this was usually achieved in 4 hours. Sterile glycerol was added to these pre-grown cultures at a final concentration of

15%, and then they were stored at -80℃. Additional pre-grown cultures were inoculated from these stored samples to work with cells from the same colony during the fermentations. Next, to create the pre-grown culture, 1 mL of the stored sample was inoculated into 5 mL of LB nutrient solution containing 100 mg/mL ampicillin and these were incubated under the same conditions as before.

During the autoinduction fermentations, LB and M9 nutrient solutions were used as controls, and we also examined the effectiveness of three autoinduction nutrient solutions: ZYM, ZYP, and autoinduction M9. During this experiment, we examined 12 different conditions, 2 cell lines and the expression of 2 recombinant proteins. The effect of different substrate/inducer ratios on cell growth and expression was investigated at 37 ℃ for 24 hours, 12 type pre-composed nutrient solution was measured, then 200-200 µL per sample was distributed on a 96-well microtiter plate (TPP TP29097) with 8 repetitions. A cycle time of 6 minutes is justified for continuous shaking.

A series of measurements took place for 24 hours, after fermentation the samples were collected, centrifuged at 18,000 RPM for 15 minutes, and then the supernatant was pipetted into a clean centrifuge tube. To extract the EGFP and AURKA proteins from cells, the cells were separated from the pellets applying the following protocol: the pellet (cell sediment) was suspended in 100 µL of lysis buffer, incubated at room temperature for 30 minutes, then placed on ice and sonicated for 2 x 20 pulses (Hielscher UP200S), centrifuged at 18,000 RPM for 10 minutes, then the supernatant was carefully pipetted into another centrifuge tube. The supernatant after centrifugation contains the soluble recombinant protein.

5 µL of 6X sample handling buffer was added to 25 µL of the supernatant obtained after dissection of the cells. The samples were boiled at 98 ºC for 5 minutes. During the run of the 10% SDS PAGE gel, 7 µL of molecular weight marker (peqGOLD Protein Marker I, VWR) and 10 µL of sample were applied to the gel. Running parameters of the SDS-PAGE gel: in the case of the upper gel, 80 V for 20 minutes, then after the arrival of the samples in the lower gel, at 120 V for approx. 1.5 hours. The staining of the gel was done while shaking for 30 minutes, then the destain solution was used also with shaking to detect the protein bands.

After the collection and analysis of the samples, the determination of the recombinant protein amount is measured via fluorimetric assay. We suspended 5 µL of the samples in 100 µL of lysis buffer, and then prepared a 20x dilution from this. The fluorescence of the prepared samples and dilutions was measured using a fluorimeter (BMG LABTECH FLUOstar OPTIMA) (405 nm excitation, 520 nm emission). During the calibration curve, we used a dilution series prepared from a protein solution with a known EGFP concentration. The basis of the measurement is the fluorescent property of EGFP, through which the concentration of the produced recombinant protein can be determined by measuring the fluorescence intensity.

4. Results

This PhD thesis discusses strategies for the development of recombinant protein production in the *E. coli* expression system in three segments:

- 1. application of ubiquitin fusion for the biosynthesis of SMAC and BUFII proteins in their native form.

- 2. denaturation of GST-XIAP inclusion bodies (aggregates), followed by its refolding to produce the active form of the protein.

- 3. application of autoinduction against the aggregation of AURKA into insoluble inclusion bodies. To adjust the conditions of the auto-induction fermentation, we chose EGFP, a comprehensively studied recombinant protein, which results in efficient active protein production in dissolved form even during traditional, IPTG-based induction.

The first step of molecular cloning was the design of the primers (oligonucleotides), which were used to amplify the target gene (SMAC) from the template DNA by polymerase chain reaction (PCR) simultaneously with the creation of appropriate restriction cleavage sites. Due to the small size of BUFORIN II (21 amino acids, 63 bp), the insert (gene) required for cloning is two single-stranded oligonucleotides. The insert used for the pUbiq-BUFII construct was produced by hybridization of two single-stranded oligonucleotides (DNA sequence encoding BUFORIN II peptide). The PCR/hybridization product was purified to remove the buffer solution and then double digested with the appropriate restriction enzymes to generate the needed sticky ends. After digestion of the pUbiq vector, the corresponding (large) fragment was isolated from the agarose gel.

A ligase enzyme was used to assemble the construct. The ligation product was transformed into competent *E. coli* TOP10 cells and then spread on LB plates containing kanamycin to check which colonies contained the plasmid. Plasmid was isolated from the colonies which appeared on the LB agar. After the control digestions, a sample suitable for DNA sequencing was prepared from the plasmid that showed appropriate insert incorporation. To express the recombinant construct encoded by the plasmid, we used *E. coli* BL21 (DE3) Rosetta plysS bacterial strain. The protein was purified on an FPLC system using affinity chromatography, followed by the enzymatic removal of SMAC/BUFII from the fusion protein. The native SMAC/BUFII was extracted from the digestion reaction mixture by affinity chromatography.

The use of the pUbiq expression system for the biosynthesis of native SMAC proved to be successful in a prokaryotic (*E. coli*) expression system (GST pull-down). By modelling the 3D structure of the recombinant protein and examining its interactions, the effectiveness of proteolysis on the complex (removal of the fusion partner) can be predicted. Molecular modelling of the 3D structure of the recombinant protein should be one of the most important steps in the design of biosynthetic processes (Kuhlman and Bradley 2019).

The implementation of an efficient expression system requires the use of several molecular biological tools and techniques (Demain and Vaishnav 2009; de Lorenzo 2009). In the present study, the main thesis point of the PhD dissertation is to investigate the efficient use of the pUbiq expression system through the human recombinant SMAC protein in dissolved form, high protein expression of a recombinant protein with a free N-terminal end in a prokaryotic (*E. coli*) system. The combination of *E. coli* used as the expression host and the pUbiq expression construct achieved adequate protein production despite the optimization of the expression conditions. YUH1 proteolytic digestion was used to isolate native SMAC and remove it from the fusion partner. The functionality of the native fusion partner-free recombinant SMAC protein was confirmed by the GST pull-down technique, using GST-XIAP as a bait, which is a known interaction partner (Song et al. 2003).

The protein expression system and purification strategy described in the study can be used to produce active recombinant proteins in a time- and cost-effective manner. The further applicability of the expression system examined in the study was also realized in the case of the production of the recombinant protein CCMV (Szövérfi et al., 2021).

The key result of the study is that the recombinant SMAC protein fused with ubiquitin was synthesized for the first time in a bacterial expression system, and the production of native SMAC protein was achieved using the present method. Instead of the entire protein structure, only the region that plays a role in the interaction of the SMAC protein with the N-terminal segment of AVPIA was cloned and expressed.

During our research, we realized the high-yield expression of the antimicrobial peptide BUFORIN II in a bioreactor, using the Escherichia coli BL21 DE3 Rosetta cell line as a host, provided with a ubiquitin fusion partner. The fusion complex is complemented with a decahistidine-tag, which enables purification of the antimicrobial peptide by affinity chromatography. Due to their structural and functional properties, the use of antimicrobial peptides is a promising therapeutic alternative to traditional antibacterial agents. During clinical trials, antimicrobial peptides are used in the treatment of diseases such as infections caused by Gram-positive bacteria (*Clostridium difficile, methicillin-resistant Staphylococcus aureus*), diabetic foot ulcers, rosacea, chronic bacterial otitis media or nail fungus (Fox 2013). BUFORIN II is a comprehensively studied antimicrobial peptide, which was first produced by modifying the structure of buforin I peptide isolated from the frog species *Bufo gargarizans* (Park et al. 1996). The peptide consisting of only 21 amino acids showed a significant antibacterial effect against Gram-positive (*Bacillus subtilis, Staphylococcus aureus, Streptococcus mutans, Streptococcus pneumoniae, Pseudomonas putida*) and Gram-negative (*Escherichia coli, Salmonella typhimurium, Serratia sp*.) bacteria, as well as fungi (*Candida albicans, Cryptococcus neoformans, Saccharomyces cerevisiae*). Unlike other antimicrobial peptides, BUFORIN II does not modify the membrane structure, but directly binds to bacterial DNA, or it exerts its antibacterial effect by binding to RNA. Due to this property, BUFORIN II and its analogues produced by structural modification also have an antitumor effect.

Mild solubilization of inclusion body aggregates is the key to improving the recovery of bioactive proteins, in our work we examined the effect of N-Lauroyl-sarkosyl concentration on the solubilization yield of the target protein. A refolding protocol based on the solubilization of proteins without chaotropic agents was optimized and applied to the GST-XIAP construct. Removal of N-Lauroyl-sarcosyl is required for efficient affinity purification and refolding of the target protein. There are several studies in the literature on the removal of N-lauroyl-sarcosine by dilution and dialysis (Yamaguchi and Miyazaki 2014). The CMC value of N-Lauroyl-sarcosine is >0.5% (14 mM), at lower concentrations (0.1-1%) it can be removed by dialysis. A low CMC indicates that the equilibrium between detergent monomers and detergent micelles is entirely on the micelle side, and the micelles are large, while only detergent monomers can easily diffuse into the surrounding buffer. Detergents containing more than 5 mM CMC cannot be removed by dialysis. Additionally, N-Lauroyl-sarcosine can be isolated by adding a non-ionic detergent such as OTG in an excess of at least five times the weight of N-Lauroyl-sarcosine.

In the dissertation, I also investigated the autoinduction fermentation conditions during which recombinant protein can be produced in an active form in a time- and cost-effective manner; enabling the

slow, regulated biosynthesis of proteins that are difficult to produce (formed in inclusion bodies) (qualitative optimization).

During the growth curve of the cultures from the *E. coli* BL21 (DE3) ClearColi cell line, a slow exponential phase growth can be observed in the case of the aM9 nutrient solution, while a steeper exponential phase is observed in the ZYM, ZYP nutrient solutions. These results are explained by the fact that the ZYM, ZYP nutrient solutions have a richer nutrient composition than the synthetic aM9 nutrient solution. The nutrient-rich environment enables cells to adapt within a shorter period, while in the case of synthetic soup it takes longer. In the case of the *E. coli* BL21 (DE3) Star cell line, the amount of cell mass achieved is significantly higher in nutrient rich ZYM and ZYP nutrient solutions. By means of catabolite repression, the start of induction during autoinduction fermentation can be regulated by changing the glucose/lactose ratio. Based on the population dynamic phase analysis, the initial glucose concentration in the autoinduction M9 nutrient solution did not affect the cell growth in any of the tested cell lines. Increasing the glucose concentration in the ZYM and ZYP nutrient solutions had a positive effect on the cell density. The choice of glucose/lactose ratio is the key factor in an autoinduction fermentation experiment.

5. List of publications

Publications related to the PhD thesis:

P. Salamon, C. K. Orbán, K. Molnár-Nagy, Z. Kovács, K. Váncsa, E. Bálint, I. Miklóssy, B. Albert, G. Tar, S. Lányi, "Study of native SMAC protein production in the pUbiq expression system: molecular cloning, biosynthesis and molecular modelling", Electronic Journal of Biotechnology, vol. 56, pp. 39-46, 2022.,

IF.: 2,80

K. Nagy, Z. Kovács, I. Miklóssy, **P. Salamon**, C.-K. Orbán, B. Albert, and S. Lányi, "Detergent aided refolding and purification of recombinant XIAP from inclusion bodies," Studia Universitatis Babes-Bolyai Chemia, vol. 66, no. 4, pp. 355–368, 2021., **IF.: 0,447**

P. Salamon, I. Miklóssy, B. Albert, S. Lányi, and C. Orbán, "Comparative study on conventional and auto-induction fermentation" *Upb Scientific Bulletin, Series B: Chemistry and Materials Science*, vol. Vol. 83, Iss. 2, p. 77, 2021.

F. A. Boda, Z. I. Szabo, E. Szőcs, **P. Salamon**, C. Orbán, and E. Székely, "Heterologous expression and purification of the antimicrobial peptide BUFORIN II," *Bulletin of Medical Sciences*, vol. 92, 2020.

K. Nagy, Z. Kovacs, **P. Salamon**, C.-K. Orban, S. Lanyi, and B. Albert, "Enhanced heterologous expression in *E. coli*," *Studia Universitatis Babes-Bolyai Chemia*, vol. 64, no. 2, pp. 101–110, 2019., **IF.: 0,494**

P. Salamon, I. Miklóssy, B. Albert, M. Korodi, K. Nagy, I. Bakos, S. Lányi, and C. Orbán, "Heterologous expression and purification of recombinant proapoptotic human protein SMAC/Diablo with EGFP as fusion partner," *Studia Universitatis Babes-Bolyai Chemia*, vol. 62, no. 2, pp. 333–345, 2017., **IF.: 0,305**

Publications not related to the thesis:

Z. Bardocz-Veres, M. Székely, **P. Salamon**, E. Bala, E. Bereczki, and B. Kerekes-Máthé, "Quantitative and qualitative assessment of fluorescence in aesthetic direct restorations," *MATERIALS*, vol. 15, no. 13, p. 4619, 2022., **IF.: 3,748**

S. János, K. O. Csongor, A. Beáta, N. Katalin, **P. Salamon**, and L. Szabolcs, “In vitro study of the CCMV capsid protein: cloning, expression, and purification,” *Upb Scientific Bulletin, Series B: Chemistry and Materials Science*, vol. Vol. 83, Iss. 1, p. 135, 2021.

Z. Gáll, K. Kelemen, I. Mihály, **P. Salamon**, I. Miklóssy, B. Zsigmond, and M. Kolcsár, “Role of lacosamide in preventing pentylenetetrazole kindling-induced alterations in the expression of the gamma-2 subunit of the GABAA receptor in rats,” *Current Molecular Pharmacology*, vol. 13, p. 1, 2020., **IF.: 3,339**

S. Szentes, N. Zsibrita, M. Koncz, E. Zsigmond, **P. Salamon**, Z. Pletl, and A. Kiss, “I-Block: a simple Escherichia coli-based assay for studying sequence-specific DNA binding of proteins,” *Nucleic Acids Research*, vol. 48, no. 5, 2020., **IF.: 16,971**

B. Francisc-Andrei, **P. Salamon**, O. Csongor, B. Lavinia, C. Augustin, G. Serban-Andrei, and D. Maria, “Heterologous expression and purification of recombinant crotoxin b, the phospholipase a2 subunit of crotoxin,” *Studia Universitatis Babes-Bolyai Chemia*, Vol. 63, no. 1, pp. 7–20, 2018., **IF.: 0,275**

Presentations related to the thesis:

P. Salamon, I. Miklóssy, S. Lányi, Cs. K. Orban, and B. Albert, “Fehérje-fehérje interakció vizsgálata a sejthalál szabályozásában,” in *25th International Conference on Chemistry = XXV. Nemzetközi Vegyészkonferencia*, 2019, p. 105.

Á. Dósa, **P. Salamon**, C. Orbán, F. Boda, Z. Szabó, and E. Székely, “Antimicrobial Peptide Biosynthesis for The Development of a Nanofiber-Based Wound Dressings,” in *Fiatal Biotechnológusok Országos Konferenciája “FIBOK 2018”: Abstract Book*, 2018, p. 46.

P. Salamon, “Modern Biotechnológia Székelyföldön: Monoklonális Antitestek Bioszintézise,” in *Határhelyzetek XI.*, 2018, p. 423.

N. Katalin, O. Csongor-Kálmán, A. Beáta, M. Ildikó, **P. Salamon**, and L. Szabolcs, “Proteomimetic SMAC Peptides: Cloning, Heterologous Expression with Fusion Partner, Purification and Interaction with Inhibitor of Apoptosis Proteins,” in *20th Romanian International Conference on Chemistry and Chemical Engineering*, 2017, p. 16.

S. Réka, B. F. Márta, **P. Salamon**, A. Emőke, B. Zsolt, M. Ildikó, and L. Szabolcs, “Heterologous expression, purification, and primary characterization of an alcohol/aldehyde dehydrogenase suitable for 1,4 butanediol production,” in *20th Romanian International Conference on Chemistry and Chemical Engineering*, 2017.

Presentations not related to the thesis:

E. Antal, **P. Salamon**, and I. Miklóssy, “In silico examination of a malonyl-coa-reductase enzyme for metabolic engineering purposes,” in *ABSTRACTS BOOK, 19th International Symposium and Summer School on Bioanalysis, 2019*, 2019, p. 84.

M. Balázs, E. Antal, **P. Salamon**, I. Miklóssy, and Z. Bodor, “Analysis of Basfia Succiniciproducens Gene Expression Profile by RT-qPCR Method,” in *25th International Conference on Chemistry = XXV. Nemzetközi Vegyészkonferencia*, 2019, p. 34.

P. Salamon, F. M. Both, and I. Miklóssy, “Anyagcseremérnökséggel előállított, rekombináns *E. coli* törzsek génexpressziós profilelemzése és optimális tenyésztési körülményeinek meghatározása,” in *XXXIV. Országos Tudományos Diákköri Konferencia: Biológia Szekció*, 2019, p. 131.

P. Salamon, F. M. Both, R. Sinkler, E. Sipos, Z. Bodor, B. Albert, S. Lányi, and I. Miklóssy, “1,4-butándiol bioszintézisére alkalmas, anyagcsere mérnökséggel előállított, rekombináns *E. coli* törzsek génexpressziós profilelemzése és optimális tenyésztési körülményeinek vizsgálata,” in *XII. Szent-Györgyi Albert Konferencia kiadványa*, 2018, p. 32.

P. Salamon, “Az RT-qPCR alkalmazása iparilag hasznos vegyületek bioszintézisét megvalósító baktériumtörzsek vizsgálatára,” in *Határhelyzetek X.*, 2018, p. 379.

Posters related to the thesis:

Á. Dósa, **P. Salamon**, C. Orbán, F. Boda, Z. Szabó, and E. Székely, “Antimicrobial peptide biosynthesis for the development of a nanofiber-based wound dressings,” in Fiatal Biotechnológusok Országos Konferenciája “FIBOK 2018”: Abstract Book, 2018, p. 46.

P. Salamon, K. Nagy, Z. Kovács, B. Albert, I. Miklóssy, S. Lányi, and C. K. Orbán, “Nekroptózist moduláló humán fehérjék bioszintézise,” in *XXIV. Nemzetközi Vegyészkonferencia, 24th International Conference on Chemistry*, 2018, p. 42.

P. Salamon, F. M. Both, R. Sinkler, B. Albert, E. Sipos, and I. Miklóssy, “RT-qPCR módszer alkalmazása genetikai módosítások és tenyésztési körülmények hatásának vizsgálatára,” in *XXIII. Nemzetközi Vegyészkonferencia*, 2017, p. 119.

Posters not related to the thesis:

M. Cs. Csányi, **P. Salamon**, T. Feller, T. Bozó, J. Hársfalvi, and M. S.Z. Kellermayer, ”Force dependent multimer glycoprotein elongation” in *20th Int. Symposium and Summer School on Bioanalysis*, 2022

M. Cs. Csányi, **P. Salamon**, T. Feller, J. Hársfalvi, and M. S.Z. Kellermayer, “Von Willebrand multimer structure and extensibility analyzed by single-molecule AFM,” in PhD Scientific Days 2020, 2020

P. Salamon, F. M. Both, R. Sinkler, B. Albert, Z. Bodor, and I. Miklóssy, “RT-qPCR analysis of *E. coli* strains obtained by metabolic engineering” in *18th International Symposium and Summer School on Bioanalysis*, 2018, p. 31372038.

S. Sarolta, Z. Eszter, **P. Salamon**, and K. Antal, “; A simple genetic system to study sequence specific DNA-binding proteins in *E. coli*,” in *Hungarian Molecular Life Sciences 2017*, 2017, p. 45.