IPS Instituto Politécnico de Setúbal
Escola Superior de Tecnologia do Barreiro

BÁRBARA FILIPA AGOSTINHO ALBUQUERQUE

PRODUCTION AND CHARACTERIZATION OF INTERLEUKIN-9 AND ITS RECEPTOR

Relatório de Estágio do Mestrado em Engenharia Biológica e Química

SUPERVISOR
Doutor Pavel Mikulecky

CO-SUPERVISOR
Professor Doutor Bohdan Schneider

COORIENTADORA
Professora Doutora Marta Justino

Dezembro de 2019

BÁRBARA FILIPA
AGOSTINHO
ALBUQUERQUE

PRODUCTION AND CHARACTERIZATION OF INTERLEUKIN-9 AND ITS RECEPTOR

JÚRI

Presidente: Professora Doutora Raquel Alexandra Galamba Duarte (ESTBarreiro/IPS)
Coorientadora: Professora Doutora Marta Sofia Guedes de Campos Justino (ESTBarreiro/IPS)
Vogal: Doutora Célia Valente Romão (Instituto de Tecnologia Química e Biológica/UNL)

Dezembro de 2019

Acknowledgments

This thesis became a reality with the kind support and help of many individuals. I would like to express my sincere thanks to all of them.

First, I must thank the Institute of Biotechnology CAS, v. v. i. (part of BIOCEV research center) for receiving me over 5 months for my master's internship. I have to thank Professor Bohdan Schneider, the head of the Laboratory of Biomolecular Recognition, to gave me this great opportunity to work with the team.

Next, I have to thank Pavel Mikulecky, my master's advisor during these 5 months, for guidance and patience over this time. For helping me conclude this master thesis and give me the guidance and support that I needed it.

I also have to thank my college advisor Professor Marta Justino for all the support in developing this thesis.

And finally, I must thank my parents because without them I couldn't have this amazing experience. I also thank them and my friends for the support and strength given during this time. And my friend Catarina Domingos who embarked on this adventure with me.

Resumo

O principal objetivo deste projeto foi proceder á descoberta da estrutura da Interleucina-9 e do seu respetivo recetor. Esta Interleucina mostrou uma ligação positiva com inflamações, asma e alergias. Por este motivo a descoberta das suas estruturas é essencial para o desenvolvimento de um inibidor eficaz.

Para atingir este objetivo, este projeto foi dividido em dois objetivos principais. O primeiro baseou-se na produção e purificação da Interleucina-9 e do seu recetor. No segundo objetivo, procedeu-se à caracterização biofísica e cristalização das proteínas produzidas.

Para alcançar o primeiro objetivo, procedeu-se à utilização de vários vetores de modo a obter construções finais com o fim de melhorar a produção e a solubilidades destas duas proteínas. Ambas as proteínas foram expressas em todas as estirpes testadas. No entanto, o principal problema que enfrentamos ao usar células de *E. coli* foi o facto de todas as proteínas expressas encontravam-se em corpos de inclusão. Resultado em proteínas insolúveis. Embora não tivéssemos conseguido o refold da proteína para uma proteína solúvel, os resultados obtidos foram bastante promissores. No futuro, seria necessário repetir mais técnicas de refold de proteínas ou construções de vetores diferentes.

Além da produção em *E. coli,* também executamos a produção destas proteínas em células Schneirder 2 (S2), sendo estas células de inseto. Obtivemos uma boa expressão de ambas as proteínas e estas encontravam-se na forma solúvel.

Com estas proteínas produzidas procedemos para a caracterização biofísica e cristalização destas. Primeiro foi medida a espectro de dicroísmo circular que nos mostrou que ambas as proteínas estavam devidamente bem enroladas. Após a confirmação do complexo da Interleucina-9 com o recetor, lemos a temperatura de fusão da Interleucina-9 sozinha e do complexo. Estas medições mostraram que o complexo é mais estável que o recetor da Interleucina-9 sozinho.

A prova final de interação entre estas duas proteínas foi medida através da afinidade entre estas por uma Termoforese em Microescala (MST). As curvas medidas evidenciaram uma ligação adequada entre estas proteínas.

Finalmente, procedeu-se à cristalização da Interleucina-9 sozinha e do complexo. Infelizmente não obtivemos nenhuma forma cristalina do complexo após 51 dias de controle continuo. Ao contrário da Interleucina-9, com esta proteína for possível detetar a presença de duas formas cristalinas após dois dias. Os nossos ensaios iniciais deram resultados, porém os nossos esforços futuros serão aplicados em maior profundidade na otimização das condições de cristalização de ambas as proteínas e do complexo.

Abstract

The main purpose of this project was to discover the structure of the Interleukin-9 and its receptor. This Interleukin has shown a positive linkage with inflammation, asthma and allergies. For this reason, the discovery of these structures is very important for the development of an effective inhibitor.

In order to achieve such a goal, this work attempts to accomplish two major objectives. First, the production and purification of Interleukin-9 and its receptor. Second, the biophysical characterization and crystallization of the proteins.

For the first part, we used several vectors to conduct several constructs in order to help with the production and the solubility of these two proteins. Both proteins were expressed in all tested strains. However, the main problem that we faced using *E. coli* was the fact that all the expressed proteins were in inclusion bodies, resulting in an insoluble form protein form. Although we could not refold the proteins to a soluble form, our refolding trials showed promising results. In the future, we could try to refold these proteins with other techniques or different constructs.

Additionally, we tried the production of Interleukin-9 and its receptor in insect cells, Schneider 2 (S2) cells. We obtained a good expression of both proteins with these cells and they were in a soluble form.

With these proteins, we turned our efforts to the second objective, the biophysical characterization and crystallization of the proteins. First, we measured the Circular Dichroism Spectrum that showed us that both proteins were folded. Then, after the confirmation of the complex with the Interleukin-9 and the receptor, we analyzed the melting temperature of the Interleukin-9 alone and the complex. These measurements showed that the complex is more stable than the Interleukin-9 receptor alone. The final proof of interaction was direct measurement of the affinity between those two proteins by a MicroScale Thermophoresis (MST). The measured curves evinced a proper binding between these proteins.

Finally, we proceeded to the crystallization of the Interleukin-9 alone and the complex. Unfortunately, we did not obtain any crystal form of complex after 51 days of continuous controlling. For the Interleukin-9 alone as opposed to the previously observed results with the complex, we could detect the presence of two crystal forms of IL-9 alone after two days. Our crystallization trials gave initial hits and our further efforts will be applied in greater depth to the optimization of crystallization conditions for complex and both proteins alone.

INDEX

FIGURE INDEX

TABLE INDEX

ABBREVIATIONS

IL-9 – Interleukin-9

IL-9R- Receptor of Interleukin-9

CD- Circular Dichroism

MST-MicroScale Thermophoresis

TEV- Tobacco Etch Virus

TUES- Buffer on section 2.4.

MWCO- Molecular weight cut-off

GST- Glutathione S-transferase

1. Theoretical Introduction

1.1. Cytokines

Cytokines are pleiotropic proteins or small glycoproteins with a molecular weight less than 30 kDa. These proteins are produced by several cell types such as leukocytes (which regulate immunity and inflammation) and haematopoiesis [1].

They are categorized according to their production, either from Th1 cells or Th2 cells. Recently, a third subset of Th cells (Th17) and T regulatory cells (Treg) were categorized which show a different cytokine profile from Th1 and Th2 cells [2].

Cytokine is a general name, other names include lymphokine (cytokines made by lymphocytes), monokine (cytokines made by monocytes), chemokine (cytokines with chemotactic activities), and interleukin (cytokines made by one leukocyte and acting on other leukocytes) [3].

1.2. Interleukins

Interleukins (ILs) are secreted proteins that bind to their specific receptors and play a role in intercellular communication among leukocytes. ILs are assigned to each family based on sequence homology and receptor chain similarities or functional properties [4].

There are several interleukins and their functions mostly indicate that they are in control of immune reactions, hemopoiesis and lymphopoiesis [5]. They can be pro- or anti-inflammatory, and some also have a function as chemokines or chemoattractants for other cells. Various ILs are involved at different levels in the inflammatory pathway that ultimately leads to tissue destruction. Due to these key roles that ILs play in inflammation, they make an important potential as therapeutic targets [6].

The 37 Interleukins are divided by families, the IL-1 (composed by IL-1/IL-1R, IL18, IL-33, and IL-37), the common γ-chain cytokine family (composed by IL-2, IL-4, IL-7 IL-9, IL-15, and IL-21), IL-10 family (IL-10, IL-19, IL-20, IL-22, IL-24, and IL-26), IL-12 family (composed by IL-12, IL-23, IL-27, and IL-35), IL-17 family and then we have other ILs that do not belong to these families [4].

The activation of these interleukins consists in the activation of multiple and distinct intracellular cascades, including the MAPK (mitogen-activated protein kinase, also known as ERK for extracellular-signal-regulated kinase) cascade, the JAK/STAT pathway (Janus kinase/signal transducer and activator of transcription), and the PI3-K (phosphoinositide 3-kinase) cascade. With the presence of the ligand, signalling-proficient complex is generated and brings together the receptor-bond JAKs. JAK autophosphorylation as well as the receptor tyrosine phosphorylation generates specific docking sites for adaptor proteins with domains for building an intracellular signalling scaffold. The STAT proteins bind to these adaptor proteins and become tyrosine

phosphorylated. The phosphorylated STAT proteins then dissociate from the receptor and translocate into the nucleus where they stimulate the target gene transcription [7].

1.2.1. The common γ chain cytokine family

The common γ chain (γc) family consists of IL-2, IL-4, IL-7, IL-9, IL-15, and IL-21. This family was named for binding of these factors to the γc receptor. They act mainly as growth and proliferation factors for progenitors and mature cells. They also have roles in lineage-specific cell differentiation [4].

1.2.2. Interleukin-9

Interleukin-9 (IL-9) was first discovered in mice, and was found to be a potent antigen-independent growth factor for T cells and mast cells. It was initially termed P40 (peptide that exists as part of the Th2) based on molecular weight or Mast cell growth-enhancing activity [8-9].

The cloning and complete amino acid sequencing of P40 revealed that it is structurally distinct from the other T cells growth factors. For this reason, it was renamed as IL-9 based on biological effects on both myeloid and lymphoid cells [10-11].

The IL-9 belongs to the family of 4-helix bundle cytokines. The human IL-9 consists of a 14 kDa glycoprotein, the mature form is composed of 144 amino acids along with a signal sequence of 18 amino acids. This protein contains a high proportion of cationic amino acid residues and 10 cysteines. It also contains 4 N-linked glycosylation sites. Its gene is composed of 5 exons and 4 introns, over approximately 4 kb. The 5' flanking region shows specific binding sequences for the activator protein 1(AP-1) and AP-2 transcription factors. The human IL-9 gene is located within the Th2 cytokine cluster region [12].

1.2.3. IL-9 Signalling Complex

The IL-9 signalling complex consists of two receptors, common gamma chain (γc, CD132) and a specific IL-9 receptor (IL-9R or IL-9Rα). The γc is shared by other cytokines including IL-2, IL-4, and IL-7. On the contrary, the IL-9R is specific only for IL-9. The human IL-9R gene contains 11 exons and encodes a 522-amino acid protein [12]. A possible binding sequence for AP-1, AP-2, AP-3, and nuclear factor NF-κB was demonstrated illustrating a potential transcriptional control of the IL-9R gene [13].

The activation of the IL-9R results in the phosphorylation of Jak1 (Janus kinase 1) and Jak3 (Janus kinase 3) which subsequently leads to the activation of transcription (STAT) complexes. Specifically, STAT1-STAT5 and STAT1-STAT3 heterodimers, leading to the upregulation of IL-9-induced gene transcription [14].

It has also been reported that MAPK and insulin receptor substrate – PI3K pathways are activated through IL-9R signalling and play key roles in cell growth, survival and differentiation [4,15]. The subsequent activation of STAT factors is important for the cell growth, differentiation and development of IL-9-targeting effector cells. These changes contribute to the upregulation of IL-9 gene transcription in the IL-9/IL-9R signalling pathway [15].

1.2.4. Sources of IL-9 and IL-9 Receptor

A variety of cell types including T cells, mast cells, eosinophils and neutrophils can produce IL-9. However, the major cell source of IL-9 is CD4+ T lymphocytes. IL-9 Receptor is expressed by T cells lineages. Several studies indicated that multiple cell types can also express this receptor including macrophages, mast cells, dendritic cells and Th9 cells [4,15].

1.2.5. lL-9 Functions

As suggested by the patterns of the receptor expression, IL-9 has biological effects on several distinct cell types (Figure 1). IL-9 can affect immune cells as well as resident tissue cells that can contribute to the development of the inflammation [16].

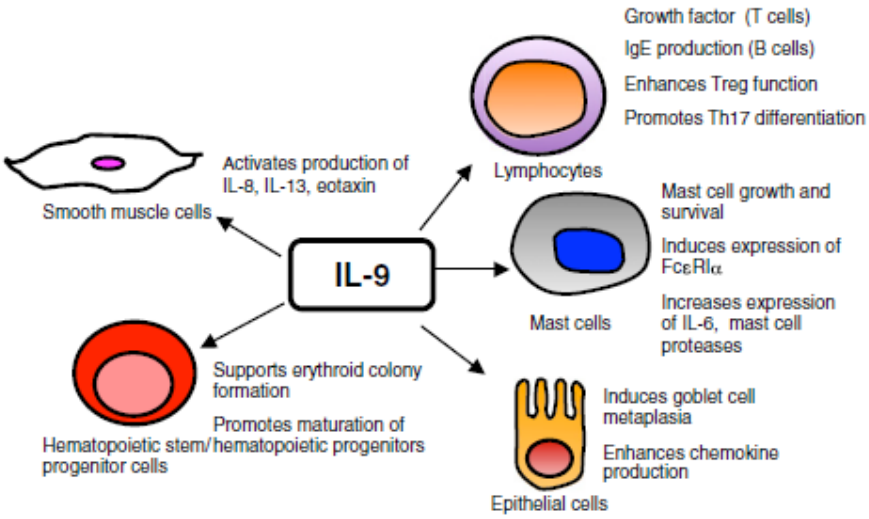


Figure 1 - Biological effects of IL-9 on several distinct cells [16].

The following points describe each biological effect of IL-9 in greater detail.

T-cells

Although IL-9 was originally cloned as a T cell growth factor, its function was more restricted than cytokines with similar functions like IL-2 and IL-4. IL-9 could promote T cell growth but it functioned preferentially on CD4+ T cells, and not on CD8+ T cells. It was also unable to support long-term antigen-independent growth of T helper lines [16].

B lymphocytes (B cells)

B lymphocytes are a population of cells that express clonally diverse cell surface immunoglobulin (Ig) receptors recognizing specific antigenic epitopes. IL-9 has effects on both B cell development and function. This cytokine increases IgE and IgG generation mediated by IL-4 from human B cells, without affecting IgM generation [17].

Mast cells

One of the main functions of IL-9 is to promote mast cell growth and function. IL-9 can promote the growth of mast cells from bone marrow and mast cell progenitors in concert with Stem Cell Factor. IL-9 alone or in combination with Stem Cell Factor or FcεRI, promotes the expression of mast cell proteases and pro-allergic cytokines in cultured mast cells [16].

Haematopoietic Stem cells

IL-9 has been widely described as a regulator factor for haematopoiesis. It has been demonstrated that human IL-9 promotes the activation of megakaryocyte progenitor cells [16].

Airway epithelial cells

Transgenic expression of IL-9 in lungs results in gene expression changes in the airway epithelial cells, including goblet cell metaplasia. Many of the effects of IL-9 in lungs are thought to occur due to the indirect effects of IL-13. In this scenario, IL-9 promotes allergic inflammation through effects on hematopoietic cells which then produce IL-13 to have direct effects on airway epithelium. However, IL-9 can have direct effects on human primary airway epithelial cells and cell lines including the direct induction of mucus genes. Thus, IL-9 may have both direct and indirect effects in the airway [16].

Airway smooth muscle cells

IL-9 also acts on airway smooth muscle cells. Human airway smooth muscle cells express both IL-9 and IL-9R. IL-9 can also potentiate the ERK (extracellular-signal-regulated kinase) dependent release of CCL11 (eosinophil chemotactic protein) and IL-8 from these cells. IL-9 mediated CCL11 expression in primary smooth muscle cells is dependent upon STAT3. When the IL-9 is activated, STAT3 binds directly to the Ccl11 promoter [16].

Intestinal and Skin tissue

IL-9 function as a protective factor in immunity against intestinal parasites. It also has a role in fungal infection because Th9 cells can be induced in the skin in response to specifics fungal antigens [15].

1.3. Production of recombinant proteins

Recombinant proteins can be used as diagnostic tools, vaccines, therapeutic proteins or functional enzymes [18].

The production of these proteins is mainly motivated by the ambition to determinate the protein structure or investigate its activity mode of action. This production is accomplished by using an expression system. Due to the existence of several systems available, the choice of the right one is the major concern with the development of recombinant proteins [19].

The methodology for all expression systems is fundamentally similar. However, the simplest or most accessible system needs to meet minimum requirements. The simple systems are often chosen for the initial expression studies and if the expression is successful, these efforts are subsequently scaled up for downstream applications [19-20].

This methodology normally consists of having a coding sequence for the heterologous protein, a vector (regularly a plasmid in which the coding sequence is cloned behind a promoter that is active in the host) and an apt host that will express the desired protein. Regardless of the final host and the encoding sequences, the cloning of the coding sequence into the plasmid and the succeeding application is performed in *E. coli* [20].

1.3.1. Protein Production in Bacteria

Bacteria are usually the first type of system considered to produce longer peptides or proteins since high yields can be achieved in a short time and the cells don't need a lot of effort to live. The production of proteins can be challenging, although this can be achieved using several specialized bacterial expression systems [19].

In the protein production, if the structure and/or function of the protein depends on disulphide bonds, proteolytic cleavage or any other post-translational process, bacteria are less likely to be the ideal system. The targeting of recombinant proteins into the periplasmic space can encourage the formation of disulphide bonds. However, the yield tends to be much lower than that achieved by cytoplasmic expression. Under these circumstances, a eukaryotic expression system might be a better option [19,21].

An example of a bacterial organism is *E. coli*. This organism system has many advantages compared with other expression systems, such as easy growth conditions, rapid biomass accumulation, and a simple scale-up process. This prokaryotic organism is often used for industrial production of therapeutic or commercial based proteins [22-23].

1.3.2. Protein Production in Yeast

Two examples of yeast systems are *Saccharomyces cerevisiae* and *Pichia pastoris.* These systems combine the simple and inexpensive cultures conditions of the bacteria with the processing abilities of the eukaryotic cells. Thereby, increasing the probability of proper folding and post-translational modifications [19]. However, the glycosylation on yeast differs significantly from human cells. Resulting in a hyper-glycosylation which can mask the active sites of the enzymes and reduce their activity. In this case, the produced proteins cannot be used for therapeutic purpose. To generate a protein for vaccination purposes, modifications need to be made to the encoding DNA. For example, the amino acids that are the target for glycosylation must be substituted [21,24-25].

1.3.3. Protein Production in Mammalian Cells

Mammalian cells are suitable for demanding proteins [19]. The fastest way to produce recombinant proteins in mammalian cells is with transient transfection, although this is suitable for analytical experiments, since most of efficient transfection reagents are expensive for large-scale applications. These limitations can be addressed by stable transfection, which also allows the selection of individual high-yielding cells to increase overall productivity [19]. Generally, these cells are used to express a wide variety of proteins such as antibodies, antigens, hormones, and enzymes [26].

1.3.4. Protein Production in Insect Cells

The insect expression systems represent an adequate compromise between bacterial and mammalian systems. In the case of proteins secreted by insect cells, signal peptides are cleaved as in mammalians cells. Disulphide bonds are formed in the endoplasmic reticulum and proprotein-converting enzymes are available for proteolytic processing [27].

Glycosylation in insect cells is similar but not identical to mammalian cells. The expression system should resemble the glycan patterns of the source of the recombinant protein as far as possible. If the glycosylation is a critical property of the protein, that will affect its behaviour. Recombinant production of properly glycosylated peptides for functional studies required an appropriate expression system such as *Drosophila Schneider* cells [28].

Insect cell lines can produce recombinant proteins with a higher density than mammalian cells. For this reason, it is possible to use cultures with smaller volumes. With insect cells it is possible to use orbital shakers or spinners under standard laboratory conditions, since there is no need for a CO_2 atmosphere. However, the sterility is still equally important as with mammalian cells [27-28].

1.4. Drosophila melanogaster S2 cells for expression of heterologous genes

Heterologous gene expression by animal cells is of particular interest mostly when the synthesis of proteins requires a complex post-translational modification. Recombinant protein expression systems based on the use of *D. melanogaster* cells are increasingly used lately. To date, about 100 cell lines derived from *D. melanogaster* have been obtained. However, the only cell lines used for heterologous gene expression are Schneider's 2 and 3 (S2 and S3, respectively) [29].

Schneider's 2 and 3 are derived from late embryonic stages of *D. melanogaster* consisting of semi adherent cultures. Stably transfected S2 cells can grow in suspension reaching high concentrations. This not only allows the use of continuous culture approaches but also eases the scale-up process. In addition, these cells can be transfected with vectors caring inducible promoters which allows the design of high-performance bioprocess protocols for protein production [30].

The S2 cell line also has advantages in terms of maximum specific cell growth rates, reaching up to 0.084 h−1 and maximum cell concentrations (Xmax) attaining around 5.0×10^{7} cells mL^{-1} [26].

1.5. Biophysical characterization

The advantages in large-scale expression and purification of recombinant proteins have paved the way for structural genomics efforts. However, little is known about newly expressed proteins. For this reason, it is necessary to have a protein characterization to a better understanding of their biochemical roles and to enable structure and function relationship studies [31].

1.5.1. Circular Dichroism (CD) Spectrum

Circular Dichroism (CD) spectroscopy is an important method for examining protein secondary structure. A large number of algorithms and reference databases have been produced for the empirical analysis of percentages of alpha-helix, beta-sheet, beta-turns and other types of secondary structures from CD spectral data [32].

1.5.2. Melting Temperature

The melting temperature is the temperature at which the properties such as viscosity or the absorption of ultraviolet light will change abruptly. This temperature varies for different proteins, but the temperatures above 41ºC will break the interactions in many proteins and denature them [33].

1.5.3. MicroScale Thermophoresis (MST)

MicroScale Thermophoresis (MST) is a powerful technique to quantify biomolecular interactions. It is based on thermophoresis, the directed movement of molecules in a temperature gradient which strongly depends on a variety of molecular properties such as size, charge, hydration shell or conformation. In an all-optical approach, an infrared laser is used for local heating and the molecule mobility in the temperature gradient is analysed via fluorescence. In standard MST, one binding partner is fluorescently labelled. However, MST can also be performed label-free by exploiting intrinsic protein UV-fluorescence [34-35].

1.6. Crystallization

Many advances in our understanding of biological systems at the molecular level have been made possible through knowledge of the detailed structures of proteins and nucleic acids. However, a crucial requirement for this work is the production of macromolecular crystals of suitable sizes and perfection for an X-ray or neutron diffraction studies. Growing crystals is often the major bottleneck for structure determination. Unfortunately, our understanding of macromolecular crystallization still consists more of an armamentarium of recipes than a set of general principles. These principles mainly consist of changes in pH, ionic strength, temperature or concentrations of an auxiliary ion or molecule. These principles are of course unique to each protein [36].

2. Experimental methods

2.1. Equipment

- DeNovix DS-11 FX
- GSTrap HP (1 mL) column from GE Healthcare
- HisTrap excel (5mL) column from GE Healthcare
- Magnetic Stirrer RCT basic from IKA
- Mini Shaking Incubator NB-205 from N-Biotek
- NGC Chromatography System from BIORAD
- Ni-NTA Agarose from Qiagen
- ProFlex PCR system from Life Technologies
- Qsonica sonicators
- SnakeSkin Dialysis Tubing (3,500 MWCO) from Thermo Scientific
- Superdex 200 10/300 GL column from GE Healthcare
- Superdex 75 16/600 column from GE Healthcare
- Thermo micropipettes
- VivaSpin 2 Centrifugical Concentrators (10 kDa MWCO PE) from Sartorius

2.2. Kits and Reagents

- Arabinose ≥99% from Merck
- Arabinose from Sigma-Aldrich
- Cystamine dihydrochloride from Sigma-Aldrich
- Cysteamine hydrochloride from Sigma-Aldrich
- dNTPs from Sigma-Aldrich
- FavorPrep Plasmid DNA extraction Mini Kit from Favorgen
- Imidazole from Sigma-Aldrich
- IPTG from Enzo LifeSciences
- IPTG from Thermo Scientific
- L-Glutathione reduced from Sigma-Aldrich
- NucleoSpin Gel and PCR Clean-up from Macherey-Nagel
- Urea from Sigma-Aldrich

2.3. Enzymes

- *Bam*HI Restriction Enzyme from New England BioLabs
- FirePol DNA Polymerase from Solis BioDyne
- *Hind*III Restriction Enzyme from New England BioLabs
- *Nco*I Restriction Enzyme from New England BioLabs
- Q5 High-Fidelity DNA Polymerase from New England BioLabs
- T4 DNA Ligase from New England BioLabs
- *Xho*I Restriction Enzyme from New England BioLabs

2.4. Solutions and Buffers

- CutSmart Buffer from New England BioLabs
- Q5 Reaction Buffer (5x concentrated) from New England BioLabs
- T4 DNA Ligase Reaction Buffer (10x concentrated) from New England BioLabs
- TUES

Tris-HCl, pH 8.0……. 50 mM
Urea……..8 M
EDTA……5 mM
SDS…….. 2%

2.5. Vectors

- pMTH-IL-9
- pMTH-IL-9R
- pBADM10
- pBADM30
- pETM11
- pETM30
- pETSumo
- pQE80

2.6. Cell strains and lines

- Drosophila Schneider 2 (S2) Cells from Invitrogen
- *Escherichia coli* BL21(DE3)
- *Escherichia coli* C41(DE3)
- *Escherichia coli* C43(DE3)
- *Escherichia coli* JM109
- *Escherichia coli* SHuffle T7
- *Escherichia coli* SHuffle T7 Express
- *Escherichia coli* Top 10

2.7. Software

- SnapGene software software from GSL Biotech LLC
- MO.Affinity Analysis from NanoTemper
- ChromLab from Bio-Rad

2.8. Sequence of the expressed proteins

- Human IL-9; UniProt: P15248 (residues 19-144)

QGCPTLAGILDINFLINKMQEDPASKCHCSANVTSCLCLGIPSDNCTRPCFSERLSQMTN
TTMQTRYPLIFSRVKKSVEVLKNNKCPYFSCEQPCNQTTAGNALTFLKSLLEIFQKEKMR
GMRGKI

- Human IL-9R (extracellular domain); UniProt: Q01113 (residues 41-270)

SVTGEGQGPRSRTFTCLTNNILRIDCHWSAPELGQGSSPWLLFTSNQAPGGTHKCILRGS
ECTVVLPPEAVLVPSDNFTITFHHCMSGREQVSLVDPEYLPRRHVKLDPPSDLQSNISSG
HCILTWSISPALEPMTTLLSYELAFKKQEEAWEQAQHRDHIVGVTWLILEAFELDPGFIH
EARLRVQMATLEDDVVEEERYTGQWSEWSQPVCFQAPQRQGPLIPPWGWP

2.9. Cloning of new constructs

In order to explore the expression of both IL-9 and IL-9R, we prepared six new constructs of each protein. These constructs included the gene of interest for the expression of these proteins and additional tags to enhance the expression and solubility.

2.9.1. Design of primers

The SnapGene software was used for the construction of the primers (Table 1) and for a simulation of the PCRs. These primers included the sites for the restriction enzymes, TEV recognition sequence, and stop codons (Figure 2A and Figure 3A). The primers used for pETsumo did not include the TEV recognition sequence (Figure 2B and Figure 3B). Finally, the inserts were amplified from pMTH-IL-9 and pMTH-IL-9R using these primers.

Table 1 -Designed primers to conduct the PCR of the inserts.

Name	Sequence
FW1-IL9	5'- CAC GGA TCC GAG AAT CTT TAT TTT CAG GGC GCC ATG GGT CAG GGG TGT CCA ACC -3'
RV1-IL9	5'- CAC AAG CTT CTA TTA TAT CTT GCC TCT CAT CCC TC -3'
FW2-IL9	5'-CAC AAG CTT ATG CAG GGG TGT CCA ACC-3'
RV2-IL9	5'- CAC CTC GAG CTA TTA TAT CTT GCC TCT CAT CCC TC-3'
FW1-IL9R	5'-CAC GGA TCC GAG AAT CTT TAT TTT CAG GGC GCC ATG GGT TCT GTC ACA GGG G-3'

RV1-IL9R	5'- C ACA AGC TTC TAT TAT GGC CAC CCC-3'
FW2-IL9R	5'- CAC AAG CTT ATG TCT GTC ACA GGG G-3'
RV2-IL9R	5'- C ACC TCG AGC TAT TAT GGC CAC CCC-3'

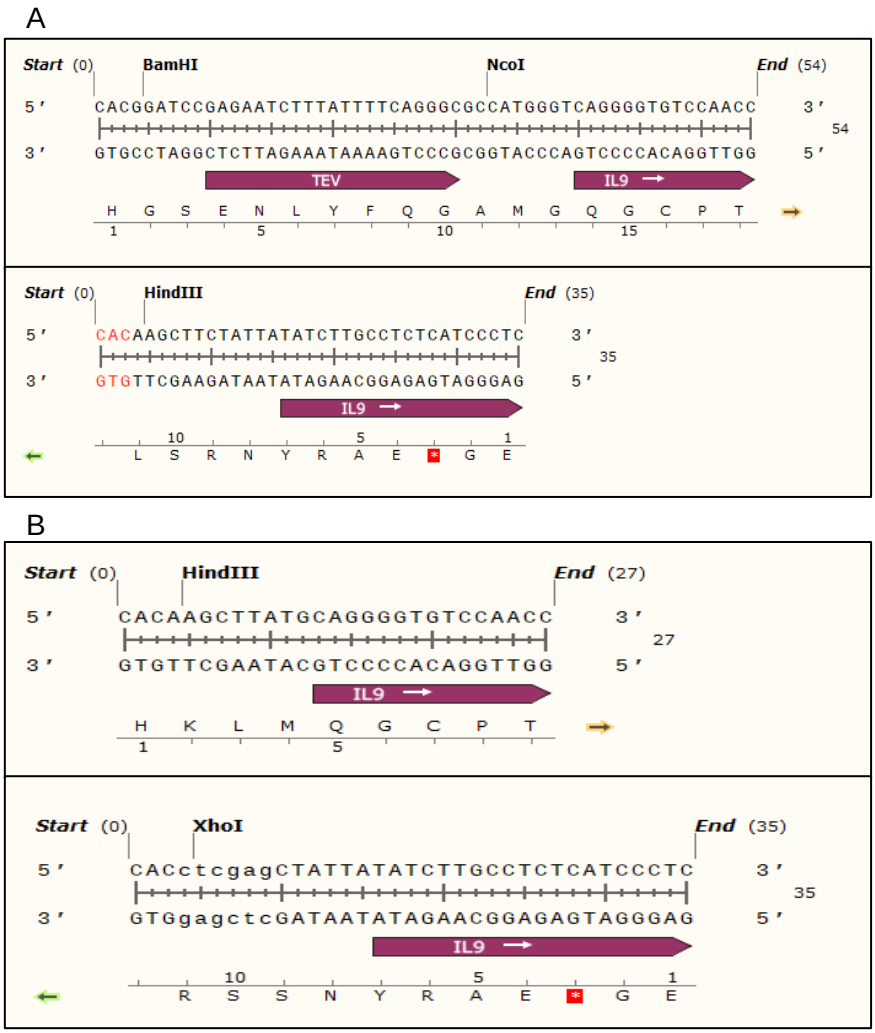
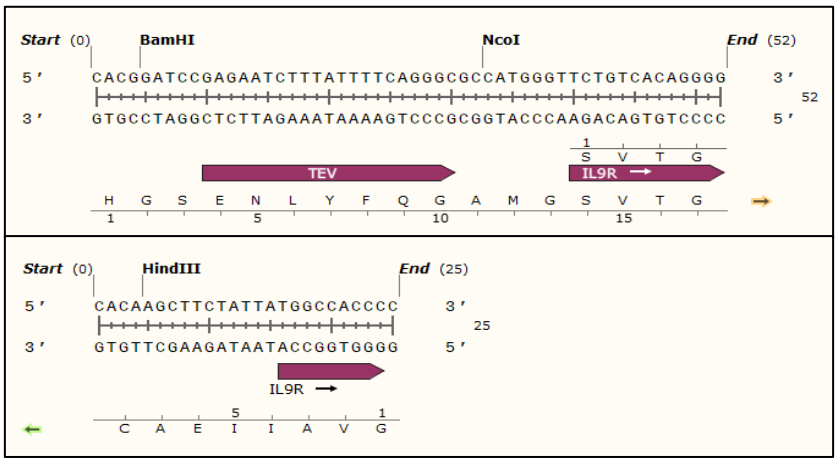


Figure 2 – A- Forward and Reverse Primers for the IL-9 gene (FW1-IL9 and RV1-IL9). Primers used in all vectors except pETsumo. B- Forward and Reverse Primers for the IL-9 gene (FW2-IL9 and RV2-IL9). Primers used in pETsumo.

A



B

Figure 3 – A- Forward and Reverse Primers for the IL-9R gene (FW1-IL9R and RV1-IL9R). Used in all vectors except pETsumo. B- Forward and Reverse Primers for the IL-9R gene (FW2-IL9R and RV2-IL9R). Used in pETsumo.

2.9.2. Cloning protocol

In a typical DNA cloning procedure, the gene or other DNA fragment of interest is first inserted into a plasmid.

i. Vectors

The first step consists in obtaining the necessary vectors. In this project the vectors used were pBADM10, pBADM30, pETM11, pETM30, pQE80 and pETsumo.

To produce the necessary amount of these vectors, *E. coli* JM109 cells were transformed with each vector grown and the DNA was isolated and purified using a commercial kit FavorPrep Plasmid DNA extraction Mini Kit.

ii. Digestion of the vectors

After having the vectors purified, the DNA was digested with restriction enzymes represented in Table 2. These enzymes needed to be the same as the ones in the preparation of the insert. Six reactions were made in order to obtain the DNA for the cloning, one for each plasmid. These reactions consisted of DNA (should be 1-2 µg), 2.5 µL of 10x CutSmart Buffer, 0.5 µL of each enzyme and distilled water to a final volume of 25 µL. The mixture was incubated overnight at a temperature of 37 ºC. The digested DNA was recovered from a 1% Agarose Gel using the commercial NucleoSpin Gel and PCR clean-up kit by company Macherey-Nagel.

Table 2 - Vectors and the respective restriction enzymes

Vectors	**Restriction Enzymes**
pBADM10	NcoI + HindIII
pBADM30	
pETM11	
pETM30	
pQE80	BamHI + HindIII
PETsumo	HindIII + XhoI

iii. PCR amplification of the Inserts

Before the digestion of the insert with the restriction enzymes, a PCR had to be conducted in order to amplify the correct region of the gene. This was done by using the primers that were previously designed (Table 3).

The vectors used in the PCR as template DNA were pMTH IL-9 and pMTH IL-9R, that respectively contained the IL-9 and IL-9R gene.

We used 50 µL of reaction mixture made in a 0.2 mL microcentrifuge tube as the total volume for PCR. This reaction was composed of 10 µL of 5x Q5 Reaction Buffer, 5 µL of 2 mM dNTPs (final concentration of 200 µM), 2.5 µL of 10 µM Forward and Reverse primers (final concentration of 0.5 µM of primers used represented in Table 3), 200 ng/ µL of DNA, 0.5 µL of Q5 Hight-Fidelity DNA Polymerase (final concentration of 1 U) and sterile water to 50 µL.

The temperatures used in the reaction were, in the initiation of denaturation 30 s at 98 °C followed by 30-35 cycles of a) a denaturation step for 10 s at 98 °C. b) The annealing for 30 s at 55 °C, and c) extension step for 1 min at 72 °C. Finally, the last step consisted in 5 min at 72 °C and then the reaction was cooled down to 4 °C. After the PCR the products were run in a 1% Agarose gel and the bands containing the genes of interest were purified using a commercial NucleoSpin Gel and PCR clean-up kit by company Macherey-Nagel.

Table 3 – Primers used in insert PCR amplification

	PCR IL-9		PCR IL-9R	
	Other vectors	Sumo vector	Other vectors	Sumo vector
10 µM FW primer	FW1-IL9	FW2-IL9	FW1-IL9R	FW2-IL9R
10 µM RW primer	RV1-IL9	RV2-IL9	RV1-IL9R	RV2-IL9R

iv. Digestion of the Inserts

The purified PCR products were digested with the respective enzymes. These reactions were performed with 20 µL of insert, 2.5 µL of 10x Buffer CutSmart, 0.5 µL of each enzyme and sterile water to 25 µL.

This reaction was executed in order to prepare the insert for the ligation. At the end of this reaction, we have the genes of IL-9 and IL-9R with the restriction sites needed for the transformation of the vectors. After the digestion, the samples were purified again using a commercial NucleoSpin Gel and PCR clean-up kit by company Macherey-Nagel.

v. Ligation

The ligation is the construction of the final construct. This new construct will be composed by the vector and the previously digested insert. The ligation reaction was set up in a tube with 2 µL of 10x T4 DNA Ligase Reaction Buffer, 50 ng of vector DNA, 3-fold molar excess of Insert DNA to vector DNA, 0.5 µL of T4 DNA ligase and sterile water to a final volume of 20 µL. All the reactions were incubated at room temperature for at least a minimum of 3 hours.

vi. Transformation

For the transformation of Top 10 cells, 1 µL of the construct DNA was gently mixed with the cells which were put on ice for 30 min. Afterward, a heat shock was performed by putting the cells at 42ºC for 45 seconds. In the end, they were put on ice again for 2-5 min. 950 µl of LB medium was added to the cells which were subsequently put in a shaker for 1 hour at 37ºC. Finally, the cells were centrifuged and the supernatant was discarded leaving just 100 µL of medium to resuspend the cells. The resuspended cells were spread in an LB agar plate with the appropriate antibiotic.

vii. Colony PCR

The last step of the cloning consists of confirming positive clones, by first choosing several colonies obtained from the transformation plates. Then, they are processed to a PCR with the respective primers that hybridize in the vector, upstream and downstream the region of the insert to confirm if the DNA from the colonies contains as the final construct.

The reaction mixture for the PCR colony was composed of 1.5 µL of 10x BD Buffer, 1.5 µL of 25 mM $MgCl_2$, 1.5 µL of 2 mM dNTPs (final concentration of 200 µM), 0.75 of 10 µM Forward and Reverse primers (final concentration of 1 µM, Table 4), 0.075 µL of FirePol DNA Polymerase and sterile water to reach the final volume of 15 µL. As DNA template a single colony was picked by a pipette tip and added to the PCR reaction mixture. The temperatures used in the reaction were, in the initial denaturation 5 min at 95 °C, followed by 30-35 cycles of amplification. The next step was the denaturation for 30 s at 95 °C, the annealing for 30 s at 55 °C, and the extension was 2 min at 72 °C. Finally, the last step consisted of 10 min at 72 °C and then the reaction was cooled down to 10 °C.

The final constructs were confirmed by sequencing.

Table 4 -Primers used for sequencing and PCR colonies.

T7	TAATACGACTCACTATAGGG
T7term	GCTAGTTATTGCTCAGCGG
pBADfor	ATGCCATAGCATTTTTATCC
pBADrev	GATTTAATCTGTATCAGG
pGEXfor	ATAGCATGGCCTTTGCAGG
pQE80-FW	CGGATAACAATTTCACACAG
pQE-BW	GTTCTGAGGTCATTACTGG

2.10. Expression tests (small scale)

A day before the expression was conducted, a small culture of 5 ml of LB medium with the appropriate antibiotics was inoculated with some colonies of transformed *E. coli*. This culture was incubated in the shaker at 37ºC over-night. The following morning, a 10 mL culture was prepared with fresh LB medium with antibiotics and 100 µL inocule of the small culture.

This new culture grew for approximately 3 hours and then the inductors 1 mM IPTG (for pET and pQE80 vectors) or 0.1% arabinose (for pBAD vectors) were added. The protein expression continued for additional 3 hours at a temperature of 37ºC or over-night at 18ºC. In the end, 1 mL of the culture was collected for a tube and was centrifuged. The supernatant was discarded and 200 µL of TUES buffer (section 2.4.) was added. Finally, the samples were analysed by SDS-PAGE.

2.11. Solubility test

If the protein was expressed, we would perform a solubility test. The cultures of the expression tests were centrifuged, and the supernatants were discarded. The pellet was resuspended in 1ml of buffer (50 mM Tris, pH 8.0, 300 mM NaCl, 0.3% Brij-35) and the lysis of the cells was executed using sonication. After the lysis, the cells were centrifuged at 20,000x g for 20 min at 15 ºC. The supernatants were collected into new tubes (these supernatants had the soluble proteins). 1 mL of buffer (50 mM Tris, pH 8.0, 8 M Urea, 300 mM NaCl) was added into the pellet, the tubes were rotated for 10-15 min and were centrifuged again in the same conditions as before. The supernatants were collected into new tubes (these supernatants had the insoluble proteins). The samples collected were analysed by SDS-PAGE.

2.12. Protein refolding for the pETM11-IL9R construct

In order to remove the Urea present in the protein, we perform a dialysis. The samples were introduced inside a snakeskin membrane of 3,500 molecular weight cut off (MWCO) and submerged in Buffer (50 mM, Tris, pH 8.0, 300 mM NaCl). The samples stayed in the buffer overnight and were analysed in an SDS gel, in order to conclude if the refolding of the protein was executed or not.

2.12.1. Second dialysis

For the second dialysis, we only changed the buffer following the same procedure as before. The new buffer was composed by 0.5 mM Cystamine dihydrochloride, 2.5 mM Cysteamine hydrochloride, 50 mM Tris, 300 mM NaCl.

2.12.2. Protein purification using Ni-NTA column

After the dialysis, the samples containing the proteins from pETM11-IL9R construct were purified using a Ni-NTA column. We used both water and equilibration buffer (50 mM Tris, 300 mM NaCl, pH 8.0) to wash the column before applying the sample. After the sample went through the column, we washed it again with the first buffer (binding buffer) and then with a washing buffer (50 mM Tris, 300 mM NaCl, pH 8.0, 20 mM Imidazole). The elution of the protein was done with an elution buffer (50 mM Tris, 300 mM NaCl, pH 8.0, 300 mM Imidazole) and the samples were collected in several tubes. The samples were analysed by SDS-PAGE.

2.12.3. Size-exclusion chromatography

For the size exclusion chromatography, we used a Superdex 200 10/300 column, equilibrated with buffer (50 mM Tris, 300 mM NaCl, pH 8.0) at a flow rate 0. mL/min. The samples collected when necessary were concentrated using VivaSpin 2 (10,000 MWCO PES) concentrator.

2.13. Protein refolding and purification for the pBAD30-IL9 construct

In order to remove the Urea present in the protein we performed a dialysis. The samples were introduced inside a snakeskin of 3,500 MWCO and were submerged in Buffer (0.5 mM Cystamine dihydrochloride, 2.5 mM Cysteamine hydrochloride, 50 mM Tris, pH 8.0, 300 mM NaCl). The samples stayed in the buffer overnight and were analysed by SDS-PAGE, in order to conclude if the refolding of the protein was executed or not.

2.13.1. Glutathione S-Transferase affinity chromatography

The samples from the dialysis were purified using a GSTrap HP 1 mL column. The binding buffer consisted in a buffer of 50 mM Tris, pH 8.0, 300 mM NaCl. The elution buffer (50 mM Tris, pH 8.0, 300 mM NaCl, 10 mM L-Glutathione reduced) was used at a flow rate 0.5 mL/min.

2.13.2. Size-exclusion chromatography

For the size exclusion we used the Superdex 200 10/300 column, in buffer (50 mM Tris, 300 mM NaCl, pH 8.0) at a flow rate 0.7 mL/min. When necessary, the samples collected were concentrated using VivaSpin 2 (10,000 MWCO PES) concentrator.

2.14. Transfection of S2 cells

In the morning of the transfection day, 1.0 x 10^6 cells/mL were seeded per 5-well (IL9, GFP, IL9R, complex, cells (neg control)) in 0.6 mL in SFX + 10% FBS. The cells were left to attach for approximately 3 hours.

0.5 µg of expression DNA (DNA concentration: 0.2 µg/µL) and 0.03 µg selection plasmid (pCoBLAST) with the DNA-condensation buffer (Buffer EC) were mixed to a total volume of 50 µL. And 4 µL Enhancer was added and the solution was mixed for 5 s. This solution was incubated at room temperature (15–25°C) for 5 min. 12.5 µL Effectene Transfection Reagent was added to the DNA-Enhancer mixture. This mixture was mixed by pipetting up and down 5 times. The samples were incubated for 5–10 min at room temperature (15–25°C) to allow transfection-complex formation.

We gently aspirated 0.3 mL of the growth medium from each well. 0.3 mL of SFX + 10% FBS were added to the tube containing the transfection complexes. Again, the mixture was mixed by pipetting up and down twice, and the transfection complexes were immediately added dropwise onto the cells on the plate. The plate was gently swirled to ensure uniform distribution of the transfection complexes

The cells were incubated with the transfection complexes under their normal growth conditions for 48 hours. We gently aspirated 0.4 mL the growth medium and add 0.5 mL of fresh SFX + 10% FBS containing an appropriate antibiotic (e.g. 40 µg/mL blasticidin).

Transfected cells were selected, adding fresh SFX + 10% FBS containing an appropriate antibiotic (e.g. 40 µg/mL blasticidin) every third or fourth day.

After the stabilization of the transfected cells, several scales-up were performed until we reach a volume of 700 mL of medium. After the induction of these cells, the medium was centrifuged, in order to remove the S2 cells. The medium was stored ate -80 ºC.

2.15. IL9/IL9R purification from S2 cells

The 700 mL of culture medium from S2 cells were stored at -80ºC. The sample was thawed in warm water.

When the medium was completely defrosted, the following components were added:

- CaCl2 (5 mM final → 7.7 mL of 0.5 M stock)

- NiSO4 (1 mM final → 7.7 mL of 0.1 M stock)
- NaCl (200 mM final → 31 mL of 5 M stock)
- Tris pH 8.5 (50 mM final → 39 mL of 1 M stock)

The mixture was stirred for 1 hour at rotation 5,000x g for 30 min. The final mixture was then introduced in to an HisTrap excel 5 mL column after the equilibration buffer (50 mM Tris, pH 8.0, 300 mM NaCl. 20 mM Imidazole). The elution of the proteins was executed with an elution buffer (50 mM Tris, pH 8.0, 300 mM NaCl, 300 mM Imidazole) at a flow rate of 2 mL/min. The samples collected from the HisTrap were introduced in a Superdex 75 16/600 column with the elution buffer (50 mM Tris, pH 8.0, 300 mM NaCl) at a flow rate of 1.0 mL/min.

2.16. Confirmation of complex by SEC

Both IL-9 and IL-9R were introduced at the same time in a Superdex75 column, in order to purify the complex formed. For the size exclusion, we used the Superdex75 10/300 column, in buffer (50 mM Tris, 300 mM NaCl, pH 8.0) at a flow rate 0.7 mL/min. When necessary, the samples collected were concentrated using VivaSpin 2 (10,000 MWCO PES) concentrator.

2.17. Crystallization

The complex of IL-9 with its receptor and the IL-9 alone were used in crystallization trials. For the complex and the IL-9 alone we used plates of JCSG+, Crystal Screen and Jena XP. We also used for the IL-9 the PACT premiere. Each plate has a different compound composition, using more plates we increased the chances of obtaining crystals from our proteins.

2.18. Circular Dichroism (CD) spectroscopy

The measurement of CD spectra was done in co-operation with Centre of Molecular Structure at BIOCEV by Tania Charnavets, the platform specialist. The CD spectra were recorded using “Chirascan-plus” (Applied Photophysics) spectrometer in steps of 1 nm over the wavelength range of 195-260 nm. Samples at a concentration of 0.2 mg/mL were placed into quartz cell to the thermostated holder and individual spectra were recorded at the temperature of 25 °C. The CD signal was expressed as the difference between the molar absorption of the right- and left-handed circularly polarized light and the resulting spectra were buffer-subtracted.

2.19. Melting temperature

The melting temperatures of IL-9R and IL-9/IL-9R were determined using the nano differential scanning fluorometry (nanoDSF) implements in the Prometheus NT.48 instrument. The samples were loaded into Standard capillaries and heated from 20 to 95 °C at a rate of 1 °C/min. The melting temperatures (Tm) were estimated from the first derivative of the melting curves.

2.20. Affinity measurement

We utilize the MicroScale Thermophoresis technique to evaluate affinity between IL-9 and its IL-9R receptor. In our case, the IL-9R protein was labelled using the Monolith Protein Labelling Kit RED-NHS (Amine Reactive) according to manufacturer's protocol, including the buffer exchange into PBS, pH 7.4. IL-9R was purified in Tris buffer that is not compatible with this labelling kit. The labelled IL-9R (final concentration 100 nM) was further titrated with sixteen different concentrations of IL-9 diluted in Assay buffer (50 mM Tris, pH 8.0, 300 mM NaCl, 0.05% Tween-20). The resulting samples of IL-9/IL-9R complex were loaded into Standard type capillaries and inserted into the Monolith NT.115 instrument (NanoTemper). All measurements were done in triplicates. The final affinity was evaluated using the MO.Affinity Analysis software, v2.3 (NanoTemper).

3. Results

3.1. Production of new expression constructs

Several vectors were utilized to prepare new constructs of IL-9 and IL-9R (see Table 5). They differ in the promoter region, antibiotics resistances, and N-terminal tags for easier purification. Additionally, each vector contains a recognition element for tag removal. In total, we prepared twelve new DNA constructs.

Table 5 - Vectors used for preparation of IL-9 and IL-9R constructs. Vectors contain Ara (Arabinose), T5 or T7 promoter. All constructs possess HisTag, the M30 versions include also GST (Glutathione S-Transferase) tag. Most of the vectors use TEV (Tobacco Etch Virus) proteases mine uses SUMO (Small Ubiquitin-like MOdifiers) protease. Table also shows expected sizes of our proteins with tags and fusion proteins

Vector	Promoter	Antibiotics	N-terminal tags	Protease	Exp. Size for IL-9 (kDa)	Exp. Size for IL-9R (kDa)
pBADM10	Ara	Ampicillin	HisTag	TEV	17.5	29.3
pBADM30	Ara	Ampicillin	HisTag – GST	TEV	43.3	55.1
pQE80	T5	Kan	HisTag	TEV	16.6	28.4
pETM11	T7	Kan	HisTag	TEV	17.4	29.3
pETM30	T7	Kan	HisTag – GST	TEV	43.3	55.1
pETsumo	T7	Kan	HisTag	SUMO	27.9	39.7

These vectors were digested with the respective restriction enzymes (Table 2). The gene of interest (IL-9 and IL-9R) was amplified by a PCR method and the resulting DNA fragment was digested with the same enzymes as used for the vectors.

The ligation mix (containing the digested vector and the inserts) was transformed into *Escherichia coli* JM109 strain and several single colonies of each construct were tested by PCR colony to find positive clones. The DNA from selected clones was isolated and correct cloning procedure was successfully confirmed by DNA sequencing. All constructs were used for expression tests.

3.2. Expression and solubility of IL-9 and IL-9R constructs in *E. coli* strains

Distinct *E. coli* strains were used in order to express the proteins. The cells were cultivated at both 37 ºC and 18 ºC after the induction of protein expression. In general, the temperature of 37 °C is better for production of proteins while the 18 °C is better for the solubility of proteins.

The proteins cloned into the pETM11, pETM30, and pETsumo were expressed in *E. coli* BL21(DE3), C41(DE3), C43(DE3), Shuffle T7 Express, and Shuffle T7 strains. The proteins from pBADM10 and pBADM30 were expressed in TOP10 strain and finally the proteins from pQE80 were expressed in all mentioned strains.

The results showed that both IL-9 and IL-9R were successfully expressed in every tested *E. coli* strain at 37 °C. In the case of expression at 18 °C, we detected low or no production only in a few cases.

Each expressed combination of strain and construct was tested for the protein solubility by ultrasound disintegration in TN buffer supplemented with Brij-35 detergent. This chemical was used in order to help with the solubilization of the proteins.

Unfortunately, all the solubility trials showed that the expressed proteins were in an insoluble state under the tested conditions.

3.3. Refolding of the protein expressed as pETM11-IL9R construct

Because we were not able to get a soluble protein directly by expression in any *E. coli* strains, we started the refolding trials of constructs with the highest expression rate. We chose over-night dialysis as a first refolding strategy.

The IL-9R protein cloned into the pETM11 vector was produced in *E. coli* BL21(DE3) in the form of inclusion bodies that were separated by centrifugation. The protein was then isolated by denaturation using urea. The renaturation and possible protein refolding was performed by continuous removal of the denaturation agent during dialysis against TN buffer. The results showed that only a smaller fraction of IL-9R protein precipitated during dialysis (Figure 4, sample P). The remaining soluble protein was purified using a NiNTA column (Figure 4). The eluted fractions were concentrated and further purified on size exclusion chromatography (SEC) column. The chromatography profile (Figure 5) and SDS-PAGE analysis (Figure 6) showed that protein formed oligomers with one dominant peak at the elution volume of 14 mL (Figure 5). This main peak consisted of IL-9R.

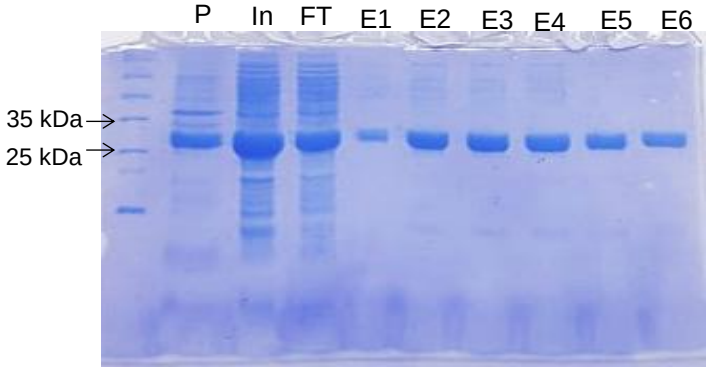


Figure 4 – SDS-PAGE from the refolding and following NiNTA purification of IL-9R protein derived from pETM11-IL9R construct. We performed the dialyzes of pETM11-IL9R construct and detected that only smaller proportion of protein precipitated (P-pellet). We further purified protein on NiNTA agarose using the soluble fraction (In- Input). Although, we can still observe protein in FT (Flow-through) fraction, we were able to elute (E1-E6) the protein. Expected size of pETM11-IL9R construct is 29.3 kDa.

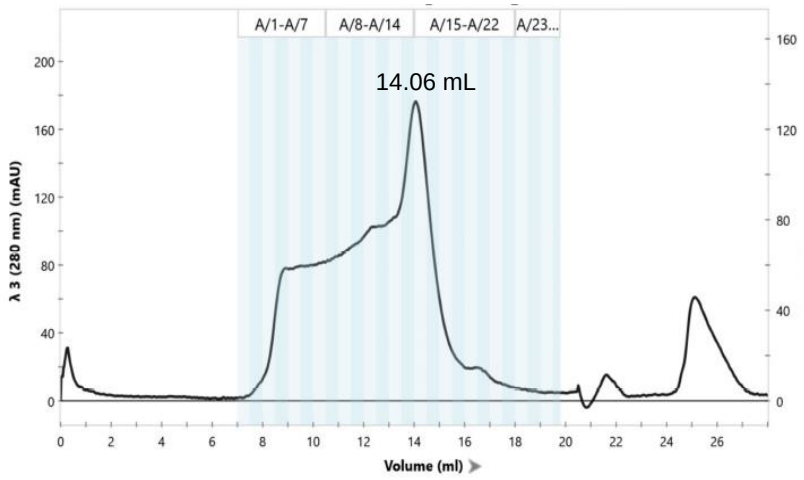


Figure 5 – Chromatogram of the size exclusion chromatography using a Superdex200 column and E2,E3, E4 as samples collected in the NiNTA column from the over-night dialyzes of the protein obtained from pETM11-IL9R construct.

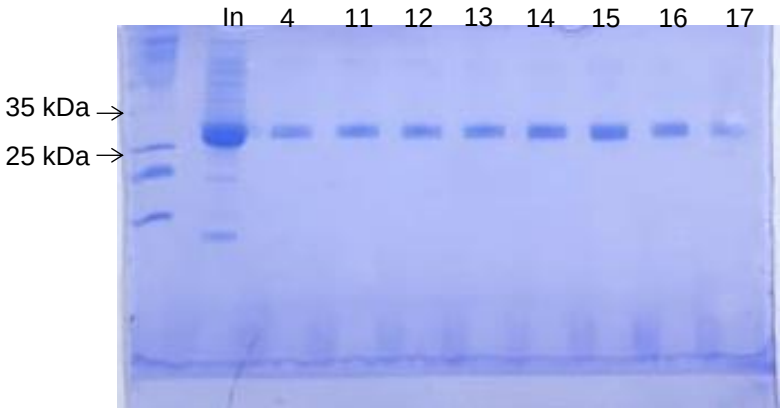


Figure 6 – SDS-PAGE from the size exclusion chromatography in a Superdex200 column using E2, E3, E4 as Input (In). The Input shows a big concentration of protein (size 29.3 kDa) with other contaminations. In the end of the size exclusion chromatography we obtained the protein purified in several fractions (4 – 17).

Because extracellular part of IL-9R contains seven cysteines, there was a possibility that oligomers were formed by wrongly made disulphide bonds. That is why we tried to refold IL-9R protein by dialysis but in the reducing-oxidation environment created by Cystamin and Cysteamine chemicals. Then, the dialyzed protein was purified on NiNTA column (Figure 7) and analysed by size exclusion chromatography (Figure 8). As in the previous case, we still detected the formation of oligomers but there was a shift towards lower molecular weights with the main peak being that of elution volume of 14 mL.

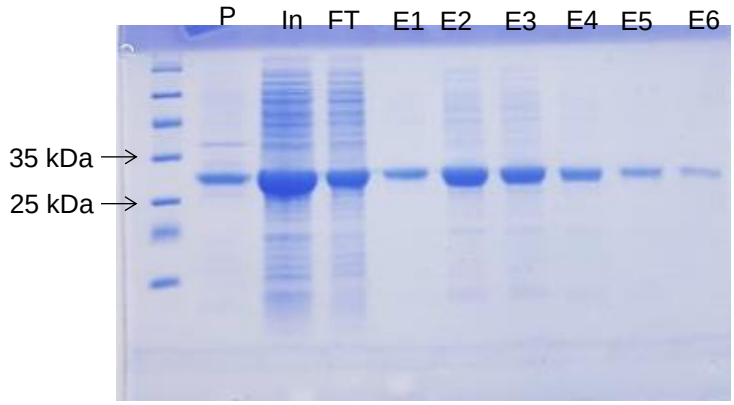


Figure 7 – SDS-PAGE from the second protein refolding and subsequent NiNTA purification of pETM11-IL9R construct. We executed a second dialysis with a different buffer in order to refold the protein. A small portion of precipitated protein was detected (P-pellet). The protein was purified on NiNTA agarose using the soluble fraction (In – Input). We were able to elute (E1 – E6) the protein, even though some of it could be observed in FT (Flow-through) fraction. Expected size of pETM11-IL9 construct is 29.3 kDa.

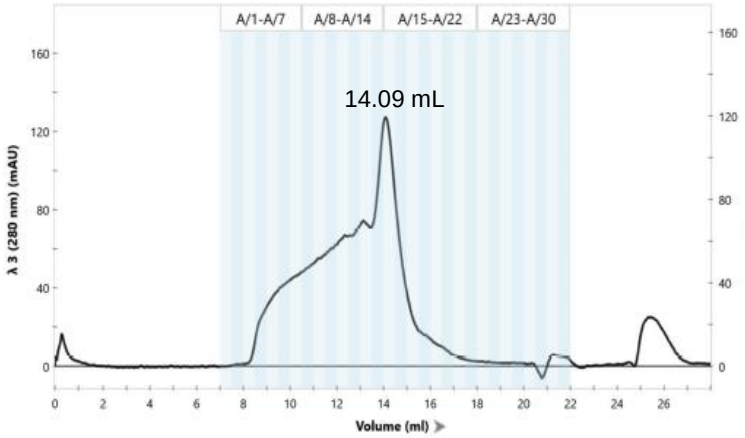


Figure 8 - Results obtained in the size exclusion chromatography using a Superdex200 column and E2,E3, E4 as samples collected in the NiNTA column from the second over-night dialysis of protein obtained from pETM11-IL9R construct.

3.4. Refolding of the protein of the pBADM30-IL9 construct

The pBADM30 vector is used to express proteins fused to Glutathione-S-Transferase (GST) protein that could help in protein solubilization. The idea was that GST could promote proper folding of IL-9 during dialysis in the reduction-oxidation environment. Considering the presence of a fusion GST protein, the purification was performed using a GST column (Figure 9) followed by size exclusion chromatography (Figure 11). The fractions collected from the GST column were analysed by an SDS-PAGE gel (Figure 10). Unfortunately, the SEC analysis displayed a non-standard profile and the experiment should be repeated for better results.

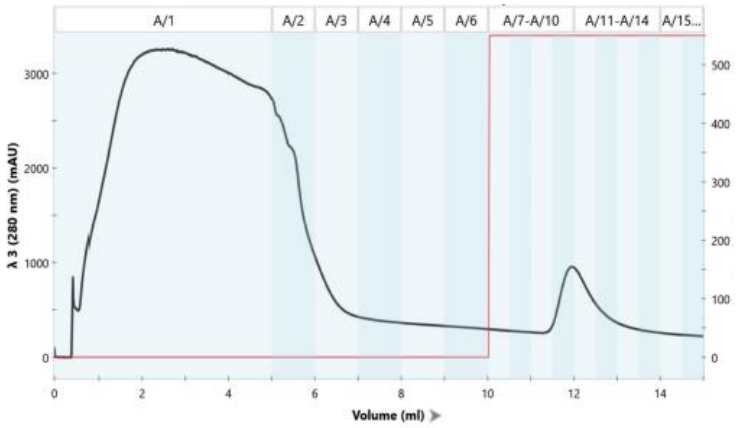


Figure 9 – Chromatogram of the purification (using a GST column) of the protein produced from pBADM30-IL9 dialysed over-night.

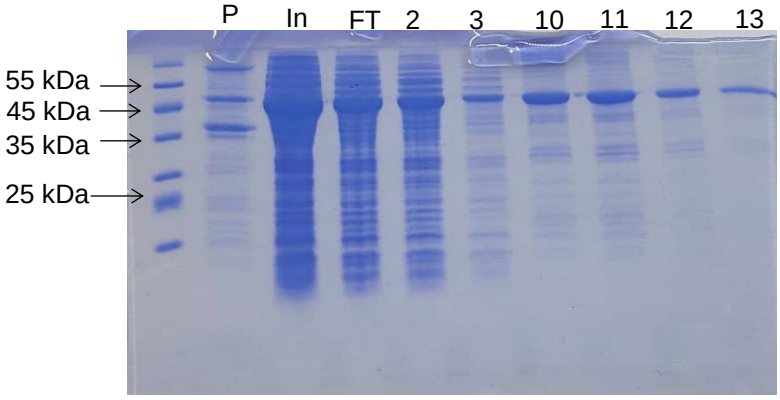


Figure 10 – SDS-PAGE from the refolding and subsquent GST purification of IL-9 protein production from pBADM30-IL9 construct. A dialysis of the protein was performed obtaining a small amount of protein precipitated (P-pellet). We then purified the soluble fraction of the protein (In-input) using a GST column. A more purified protein was collected in the elution fractions (2-13), the fractions 10, 11 and 12 were the ones with higher concentration of IL-9 protein. The FT (flow-through) also contained the protein. Expected size of pBADM30-IL9 derived protein construct is 43 kDa.

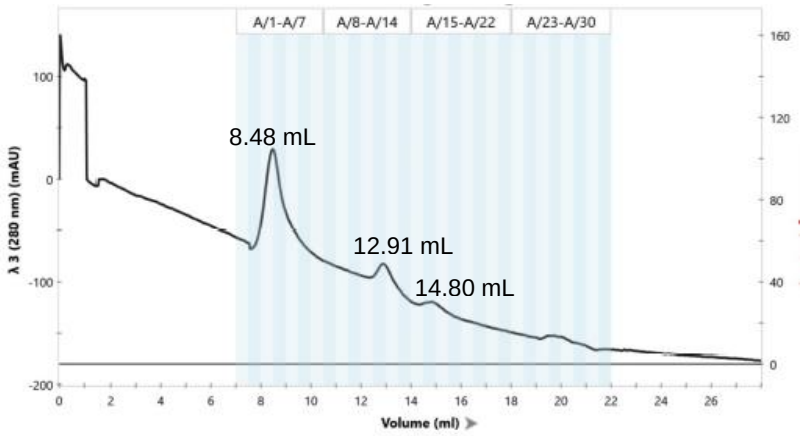


Figure 11 - Results obtained on the size exclusion chromatography using a Superfex200 column of samples 10,11, 12 from the GST column.

3.5. Expression of IL-9 and IL-9R in an insect Schneider 2 (S2) cells

The expression of IL-9 and IL-9R proteins were also tested in eukaryotic insect Schneider 2 (S2 cells). Proteins were cloned into the pMTH vector, the modified version of pMT/V5-His_A vector. Both proteins are extended to possess a C-terminal non-cleavable HisTag. The S2 cells were transfected and stable cell lines were established. The cultures were then scaled-up and addition of copper sulphate served as an induction of protein secretion into the medium. Figure 12 represents an SDS-PAGE analysis of protein expression into the 700 mL of medium. The medium with protein samples had to be seven-times concentrated to see the expression directly on SDS-PAGE.

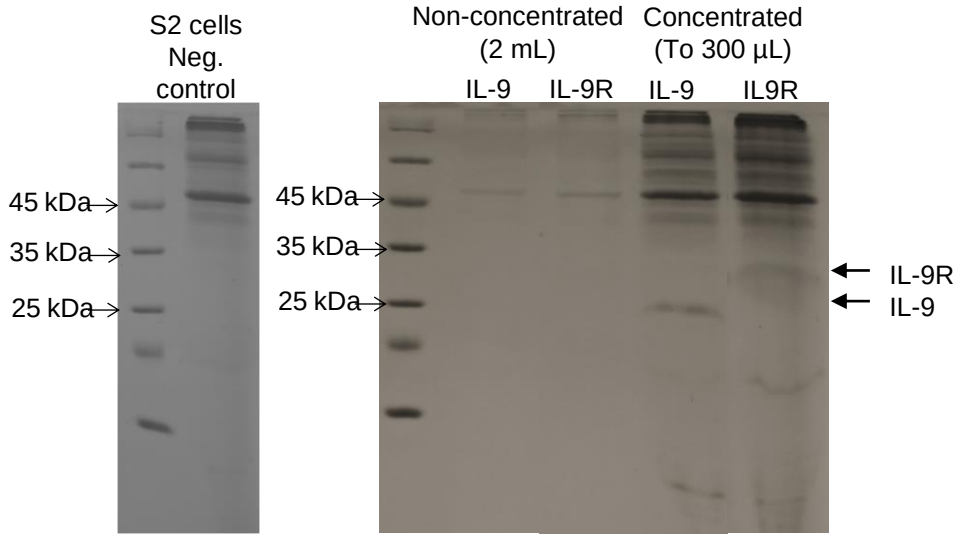


Figure 12 – SDS-PAGE from the expression of IL-9 and IL-9R in S2 cells in 700 mL of culture. Both of these proteins were expressed by the cells after the transfection.

The proteins obtained were used in different kind of experiments. The purification of these proteins consisted in the performance of two steps. First the proteins were purified using a HisTrap excel column and several fractions were obtained (Figure 13A). The fractions that contained the protein were used in the second step. This second step consisted in the purification through a Superdex 75 (16/600) obtaining again several fractions (Figure 13B and Figure 13C).

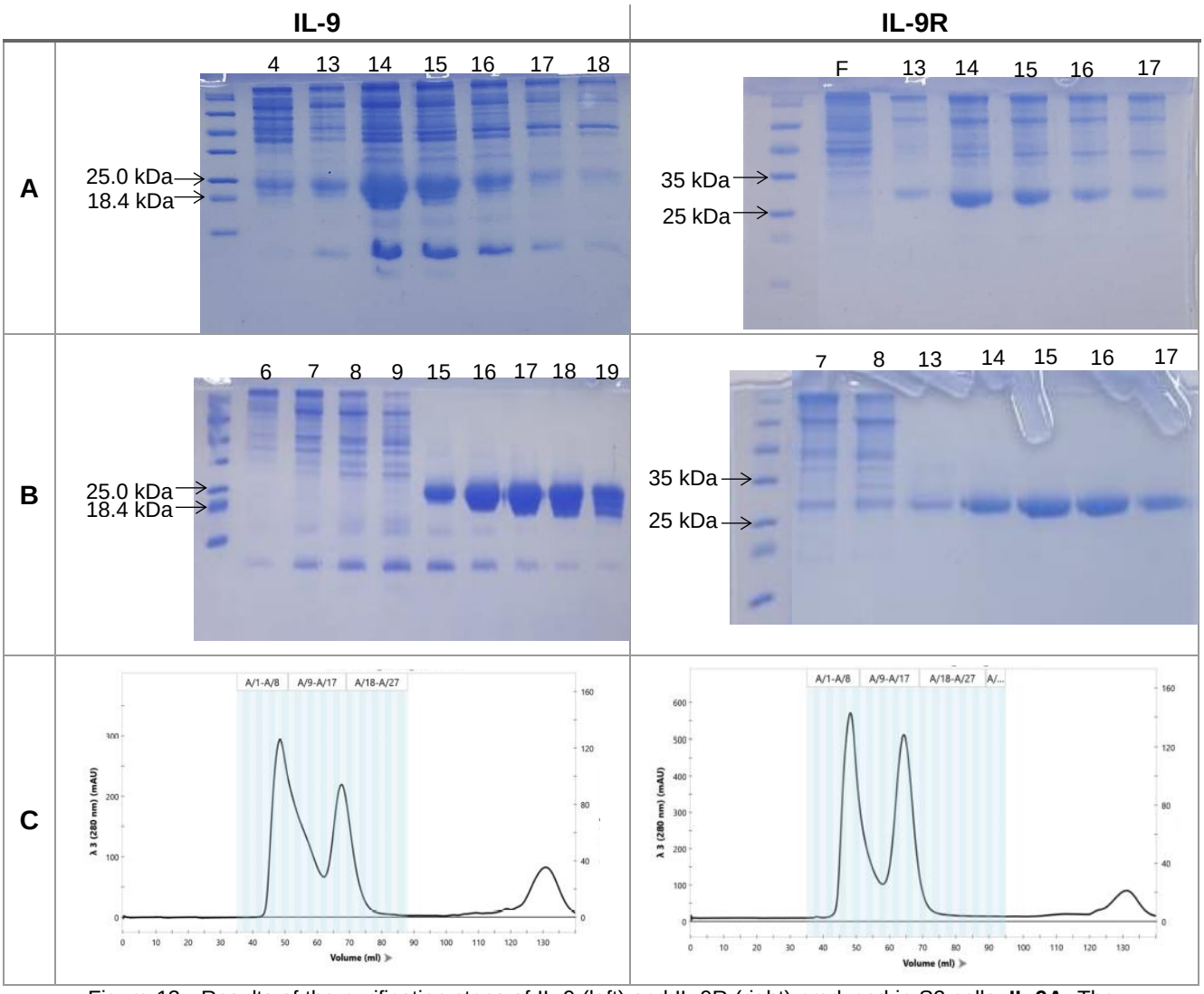


Figure 13 - Results of the purification steps of IL-9 (left) and IL-9R (right) produced in S2 cells. **IL-9A**- The purification of IL9 produced by S2 cells using a HisTrap excel column. We collected one fraction of the washing step (4) to analyse the amount of contaminations and IL9 protein. The remaining fractions (13-18) represent the eluted protein. **IL-9B**- The second purification in Superdex 75 column, using the fractions 14, 15 and 16 as the Input. In the washing steps we collected some samples to analyse the presence of contaminations (6-9). For the protein elution we collected the fractions were the protein was present (15-19). We achieved a good amount of IL9 protein in fractions 16, 17 and 18. **IL-9C**- Results obtained of the size exclusion chromatography using a Superfex75 column and 14, 15, 16 as samples from the HisTap excel column. The elution volumes were 48.48 mL and 67.53 mL. **IL-9RA**- The purification of IL9R produced by S2 cells using a HistTrap excel column. Nearly all the protein was collected in the elution step (13-17), due to an almost non-existence of it in the FT (Flow-through). **IL-9RB**- The second purification in Superdex 75 column, using the fractions 14, 15 and 16 as the Input. In the washing steps we collected some samples to analyse the presence of contaminations (7-8). For the protein elution we collected the fractions were the protein was present (13-17). We achieved a good amount of IL9 protein in fractions 14, 15 and 16. **IL-9RC**- Results obtained of the size exclusion chromatography using a Superfex75 column and 14, 15, 16 as samples from the HisTap excel column. The elution volumes were 48.20 mL and 64.31 mL.

3.5.1. Circular Dichroism (CD) Spectrum

Secondary structures of both IL-9 and IL-9R was confirmed by measurement of circular dichroism spectra. The results showed that IL-9 protein consists mainly of alpha-helices (Figure 14A) and IL-9R protein contains more types of secondary structures (Figure 14B). These results correlated with expected structural features of both proteins.

The measurements were done at the Centre of Molecular Structure at BIOCEV, directly by platform specialist Tania Charnavets.

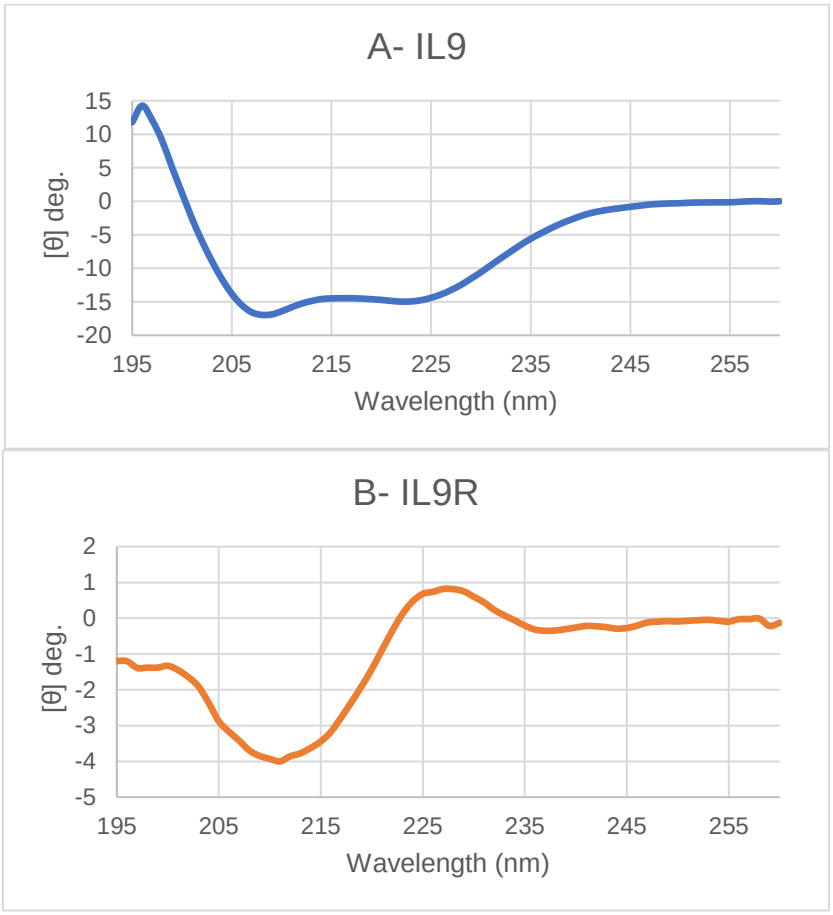


Figure 14 - CD spectrum obtained for IL9 and IL9R.

3.6. Confirmation of complex formation by size exclusion chromatography

The size exclusion chromatography (Figure 15) was used to check the formation of IL-9/IL-9R complex and its purification for crystallization trials. The SDS-PAGE analysis confirmed that the main peak contained both proteins (Figure 16A). According to general column calibration, the elution volume of 14.82 mL (fraction16) agrees with the molecular size of complex (approximately 50 kDa). The excess of non-binding proteins was eluted at later elution volumes around 16 mL (fractions 19 and 20). The main peak was concentrated and analysed again on SDS-PAGE to verify the presence of both proteins (Figure 16B).

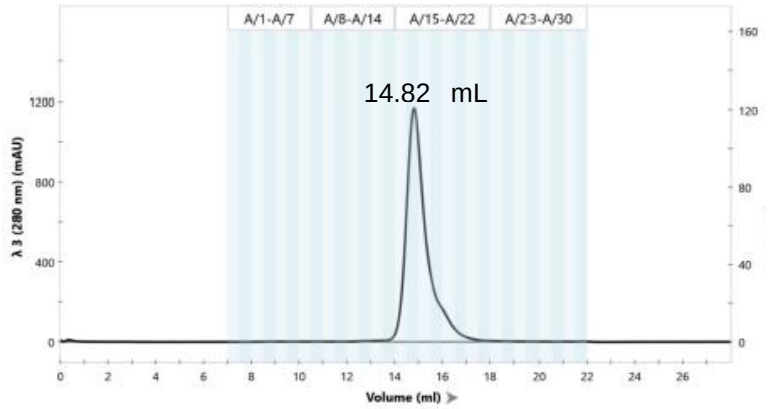


Figure 15 - Results from the size exclusion chromatography using Superdex200 to evaluate the formation of the complex of IL-9 and its receptor.

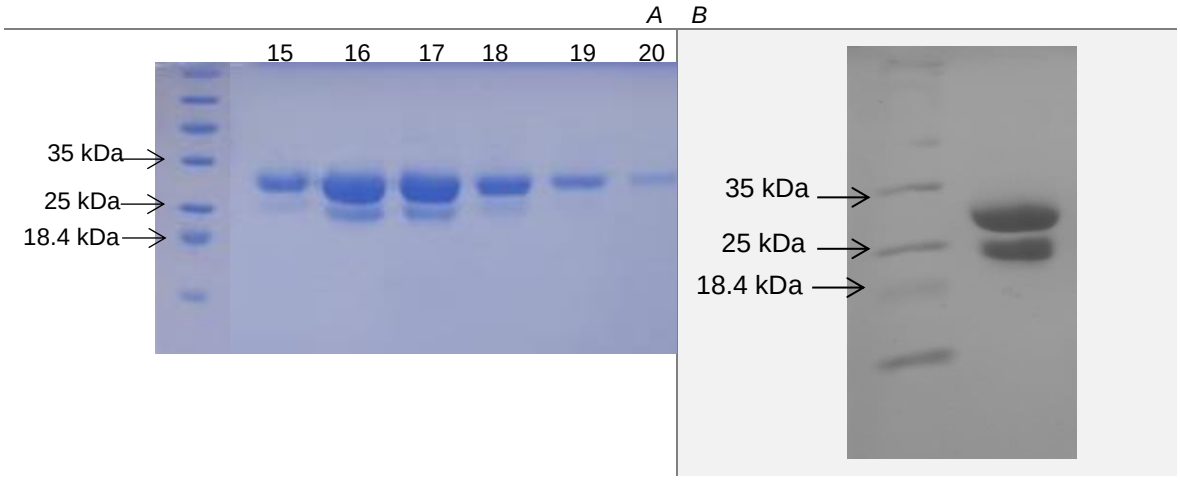


Figure 16 – SDS-PAGE analysis of the fractions collected from the size exclusion chromatography to obtain an isolate complex of IL-9 with its receptor. A- With both proteins produced in S2 cellsand purified (IL-9 and IL-9R), we introduced them in a Superdex 200 column. We were able to elute the proteins fractions (15-20), and the complex is presented in fractions number 16 and 17. B- The fractions 16 and 17 were concentrated and analysed again on SDS-PAGE.

3.6.1. Melting Temperature

The melting stability of both proteins was measured using the nano differential scanning fluorometry implemented in the Prometheus instrument. The method is based on detection of intrinsic tryptophan fluorescence in the temperature gradient. We can only check the signal of IL-9R because IL-9 doesn't contain any tryptophan. The results revealed that the melting temperature of IL-9R alone is 47 °C. However, the protein is stabilized by addition of IL-9 since the Tm of complex raised to the value of 59 °C (Figure 17).

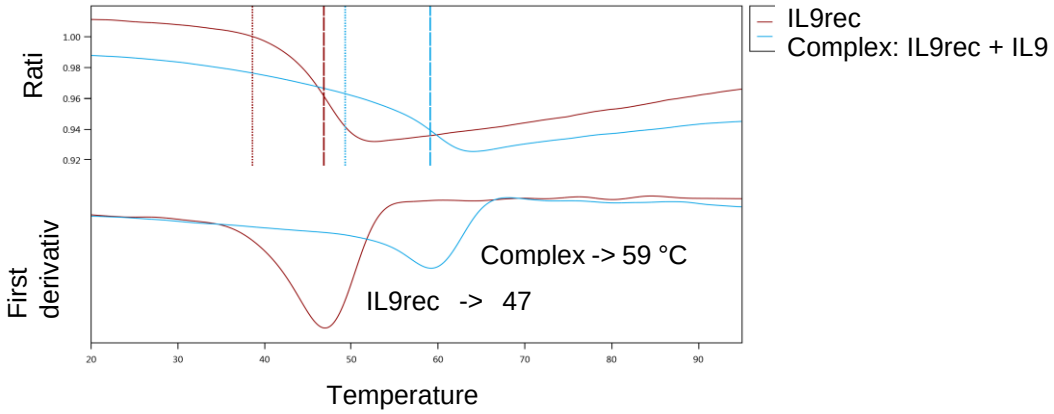


Figure 17 -Melting Temperature of IL-9R and the complex IL-9R+IL-9, measureD in Prometheus.

3.6.2. MicroScale Thermophoresis (MST)

The affinity between IL-9 and its receptor was measured by MicroScale Thermophoresis (MST) method. This technique is based on detection of movement of a fluorescently labelled molecule in the small temperature gradient. The analysis of interaction showed that the equilibrium dissociation constant Kd was 43 (±8) nM (Figure 18) under tested conditions.

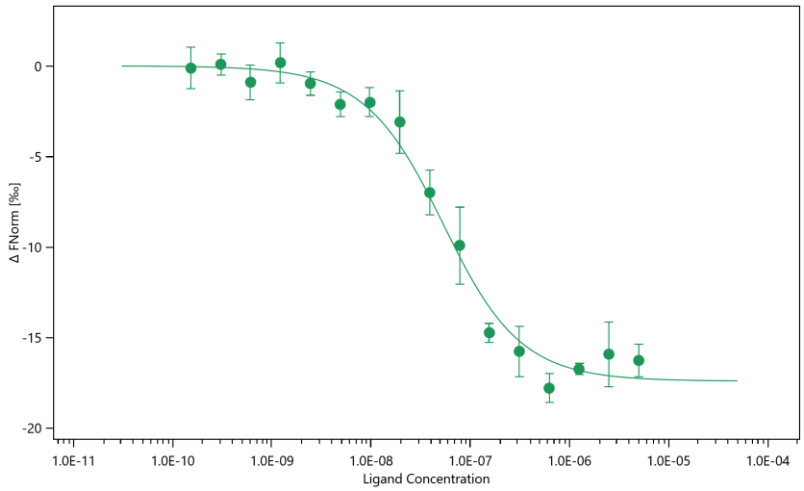


Figure 18 -Affinity results of the interaction between IL-9 and IL-9R. Labelling: IL9R with RED-NHS kit (on lysines) Buffer: 50 mM Tris, pH 8.0 + 300 mM NaCl + 0.05% Tween-20.

3.7. Crystallization of IL-9/IL-9R complex and IL-9 alone

The crystallization process of IL-9/IL-9R complex and IL-9R was initiated by sitting drop vapor diffusion method. The drops were pipetted by a crystallization robot in 96-well plates in three different ratios of protein: precipitant concentration. We setup three different crystallization screens of the complex but we didn't observe any crystal formation in the time scale of 51 days.

In case of IL-9 alone, four crystallization screens were prepared. We detected small "crystal needles" after three days in the Jena XP screen, H2 well with the ratio 1:1 (Figure 19A).The composition of the condition was 0.2 M Potassium Sodium Tartrate, 0.1 M tri-Sodium citrate, pH 5.6, 2.0 M Ammonium Sulfate, and 1 mM Anderson−Evans polyoxotungstate.

More importantly, we observed crystals' growth after two days in the Crystal Screen HT, A9 well with the ration 1:2 (Figure 19B). The composition of the condition was 0.2 M Ammonium Acetate, 0.1 M tri-Sodium citrate, pH 5.6, and 30% (w/v) PEG 4000.

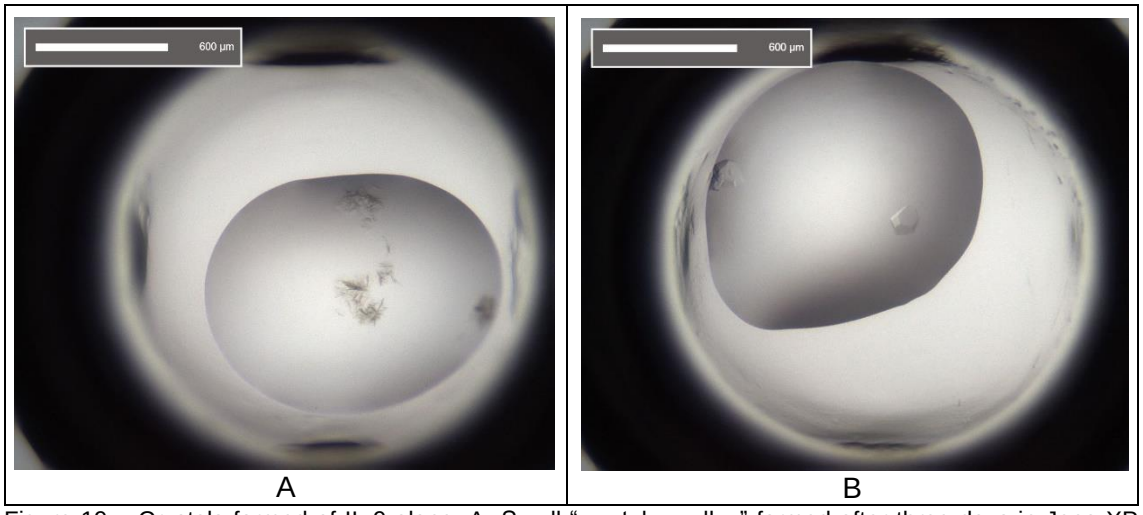


Figure 19 – Crystals formed of IL-9 alone. A- Small "crystal needles" formed after three days in Jena XP screen, H2 well with ratio 1:1. The composition of the condition was 0.2 M Potassium Sodium Tartrate, 0.1 M tri-Sodium citrate, pH 5.6, 2.0 M Ammonium Sulfate, and 1 mM Anderson−Evans polyoxotungstate. B- Two crystals growth after two days in the Crystal Screen HT, A9 well with the ration 1:2. The composition of condition was 0.2 M Ammonium Acetate, 0.1 M tri-Sodium citrate, pH 5.6, and 30% (w/v) PEG 4000.

4. Discussion

Interleukins (ILs) are secreted proteins that play a role in the intercellular communication among leukocytes. They can be pro- or anti-inflammatory, and some also have function as chemokines or chemoattractants for other cells. Various ILs are involved at different levels in the inflammatory pathway that ultimately leads to tissue destruction. Due to these key roles that ILs play in the inflammation process, they serve as important potential medicinal targets [6].

The ILs can trigger several signalling pathwaysl but the major one is the Jak/STAT signalling pathway. The starting point of activation is the binding of interleukin to its receptors to form a ternary complex. This event is the most critical and specific from a biotechnology point of view, thus making it a great therapeutic target for the inhibition of interaction between signalling partners [38].

We choose the IL-9 and its specific IL-9R receptor to be the main subject of our attention because of their great importance in pathogenesis, inflammation, asthma and allergies. A recent work has shown that IL-9 produced by T cells plays a role in both inflammation and immunosuppression. The data from this study showed that IL-9 neutralization and IL-9R deficiency attenuate the experimental autoimmune encephalomyelitis (EAE) disease, and this correlates with decreases in Th17 cells and IL-6 producing macrophages in the central nervous system, as well mast cell numbers in the regional lymph nodes. Accordingly, with this data IL-9 as a Th17-devived cytokine can contribute to inflammatory disease [37].

In addition, a splice variant in the coding region of IL-9R gene has been reported to alter the binding of IL-9 which demonstrates the importance of a functional receptor in IL-9 signalling. This finding together with the reported positive linkage between asthma and the Xq/Yq pseudoautosomal region shows the importance of the IL-9R gene region in genetic control for the susceptibility to asthma [38].

However, the lack of structural background hampers the development of effective inhibitors, which explains why it is so important to uncover these structural properties in order to enable further bioengineering and inhibitory studies [37].

We divided this project in two objectives. First, the production and purification of IL-9 and its receptor. Second, the biophysical characterization and crystallization of the proteins.

The first step consisted of choosing an expression system to produce our selected proteins. However, there was no study concerning the recombinant production of IL-9 or its receptor in high yields for further biophysical characterization. We needed to choose the starting expression system. There are several of them available to produce recombinant proteins. Nevertheless, the simplest or most accessible systems that meet minimum requirements are often chosen for the initial expression studies.

We started with the bacteria *Escherichia coli* as expression host because this system is usually the first choice to try. This expression system is the fastest and generally

the cheapest one. In addition, we also tried the insect Schneider 2 (S2) cells since it is an eukaryotic system, and we worked with human IL-9 and its receptor. Another reason was that we have a good experience with other interleukins and their receptors in this expression system.

Since the optimal conditions for certain proteins are still not evident, the most favourable conditions for protein production may vary with the gene product and cultivation conditions. In case of *E. coli*, the choice of promoter and bacterial strain will have a large effect on the production of the target protein [39].

Considering all these possible conditions, the first part of this project consisted in the preparation of several constructs of the IL-9 or IL-9R gene. Six distinct vectors were chosen in order to explore which could be the most beneficial to produce proteins in *E. coli*. The chosen cloning vectors contained various promoter regions, purification tags, and cleavage site for proteases to remove those tags.

Both the pBADM10 and pBADM30 vectors comprise the Arabinose promoter. The implication is that we need to add arabinose to start the transcription of genes coding our proteins [40]. Those vectors are suitable for expression in *E. coli* strains bearing the araBADC- mutation (deletion of both araBA and araC, and the gene for araD is inactivated), such as TOP10 strain.

The pETM11, pETM30, and pETsumo have a T7 promoter whose T7 RNA polymerase is very selective and efficient, resulting in both high frequency of transcription initiation and effective elongation. This results in an RNA elongation five-fold faster than a native *E. coli* RNA polymerase. Additionally, this promoter is also much stronger than other *E. coli* promoters. We can induce the protein expression by addition of lactose or IPTG [41][40]. Considering that T7 RNA polymerase comes from T7 bacteriophage, we need to use *E. coli* strains that include this polymerase either on extra plasmid or inserted directly into the bacterial chromosome.

The pQE80 vector contains the bacteriophage T5 promoter that is also recognized by native *E. coli* RNA polymerase. It means that this vector could be used in any *E. coli* strain. In addition, this vector contains the LacIq gene coding the repressor of T5 promoter to lower the basal protein expression. The induction is done the same way as in the case of T7 promoter, by addition of IPTG.

Two vectors, pETM30 and pBADM30, also contain the GST (Glutathione-S Transferase) tag were used. This protein is frequently used for a greater expression and solubility of the chosen protein [42]. Each vector includes the HisTag that is widely used for affinity purification. The HisTag has a low molecular weight and does not usually affect the protein structure and function. This means that it is not necessary to separate the His-tag from the target protein [40]. Although in some cases it could influence further experiments, especially the crystallization trials.

This is the reason why most of our constructs comprise the ENLYFQ/G sequence that is recognize by TEV (Tobacco Etch Virus) protease. It cleaves the sequence between Q and G which leads to the removal of N-terminal fusion elements, such as HisTag or GST tag. Those tags could interfere with the protein activity and structure [40]. One vector, pETsumo, contains SUMO (Small Ubiquitin-like MOdifier) gene instead of TEV sequence. The resulted SUMO protein is deconjugated by SUMO-specific proteases. Furthermore, this system also increases the expression levels of many different proteins and provides distinct advantages over other gene fusion

systems. Additionally, it also increases the solubility and stability of recombinant proteins [43-44].

Once we move to the question of the *E. coli* strains, the first choice to try is BL21(DE3), this strain has been largely used for many years. Although thousands of proteins have been successfully expressed in this particular strain, it is often the case that significant over-production cannot be achieved because of the toxicity of the target protein, which may even cause bacterial cell death. From BL21(DE3), two mutant host strains were derived, called C41(DE3) and C43(DE3), to grew with high saturation density, and continued to produce proteins at an elevated level without toxic effects [45]. A newer derivate of BL21(DE3) is called SoluBL21, this strain was optimized for an expression of proteins in a soluble form [46]. Since, we worked with human versions of IL-9 and IL-9R, we also tried the genetically engineered SHuffle cells. They offer an expression solution to proteins that require disulphide bonds for their folding and activity. We tested two versions of this *E. coli* strain, SHuffle T7 that is a K-12 strain and SHuffle T7 Express that is a B strain [47].

Our main interest was to find such conditions where we could produce IL-9 and IL-9R, possibly in a soluble form. It is believed that soluble proteins are properly folded. For the expression phase, we used two different temperature conditions, 37 and 18 °C. Since *E. coli* originally lives in the human intestine, the temperature of 37 °C is the optimal grow condition so there is a higher probability of the protein expression itself. Unfortunately, some proteins tend to form inclusion bodies that are considered to be an insoluble form of expressed proteins.

To avoid the formation of these inclusion bodies, it is advised to try the protein expression at 18 °C after induction. The use of reduced temperatures is a strategy that has been proven effective in improving the solubility of several difficult proteins. In general, the aggregation reaction is favoured at higher temperatures due to the strong temperature dependence of hydrophobic interactions that determine the aggregation reaction. A direct consequence of temperature reduction is the partial elimination of heat shock proteases that are induced under overexpression conditions. Furthermore, the activity and expression of several *E. coli* chaperones are increased at temperatures around 30°C. The increased stability and potential for correct folding at low temperatures are partially explained by these factors [39].

In the case of IL-9 and IL-9R, we successfully expressed both proteins at 37 °C in every tested *E. coli* strain. We also tried the production at 18 °C to detect the expression in almost all the strains tested. It is obvious that the expression of both proteins in *E. coli* was not the bottleneck of our project. The bigger concern was the solubility. Unfortunately, we were not able to produce neither IL-9 nor IL-9R in a soluble form under the tested conditions.

There are numerous strategies to try to improve the solubility of proteins. Some of these strategies imply, changing the vector with the addition of a fusion tag gene, the host strain, the culture parameters or co-expression of both proteins.

Changing the protein fusion tags is very commonly used in order to enhance the solubility of recombinantly expressed proteins in *E. coli*. We tested the GST and Sumo fusion tags, though we could also have tried Thioredoxin (Trx). This tag has been used to enhance the solubility and facilitate the expression of many mammalian cytokines [48]. Another fusion protein could be a Maltose-Binding Protein (MBP). Comparing the MBP with the GST and the Trx, the MBP demonstrated the biggest solubilisation effect [49]. There is a new fusion tag that is very promising, the Fh8 protein. It acts simultaneously as an effective solubility enhancer partner and robust

purification handle. It also has a low molecular weight, this is a great advantage over other large fusion partners for recombinant protein production [49]. There are many more fusion proteins to try [50].

Although we used several *E. coli* strains for protein production, there are still many possible strains available with interesting features. For example, the ArcticExpress strain that is designed for the expression of proteins at very low temperatures such as 10 °C. In addition, the solubility of proteins could be enhanced by the presence of cold-adapted chaperonins.

In fact, one of the strategies to promote proper protein folding is the co-expression of the chosen protein with various chaperones [51]. One company [52] offers the Chaperone plasmid set consisting of five unique plasmids coding multiple chaperones that operate as a “chaperone team” [53]. Furthermore, there are other variables to change during protein expression. We could vary concentration of inductor, cultivation medium or its additives.

The expression of both IL-9 and IL-9R was relatively straightforward but resulted in insoluble proteins in the form of inclusion bodies. For this reason, we decided to explore the refolding method as another strategy to obtain soluble proteins. The refolding process is a competing reaction with misfolding and aggregation events, its success depends on several factors. Protein solubilization results in the loss of secondary structure leading to the random coil formation of the protein structure and exposure of the hydrophobic surface [54]. There are several techniques to refold a protein, some of them are dialysis, dilution, refolding by chromatography, and some others. All of them begin with the denaturation of the chosen protein by chemicals such as urea or guanidine hydrochloride. We used urea considering that it is more compatible with analysis on SDS-PAGE.

We started with a dialysis because it is a relatively easy and simple option. It consists in the removal of the denaturant by buffer exchange through a membrane of a defined molecular weight cut-off. In this project, the purpose of the dialysis was to remove the Urea that was used to extract proteins from inclusion bodies.

We used the IL-9R protein expressed from the cDNA cloned into pETM11 to execute the first refolding trials. In this construct, the IL-9R contained only the N-terminal HisTag followed by TEV protease recognition site. We started with this protein since, it was one of the first big scale expression performed in this work and where we obtained a good expression of the protein. After the over-night dialysis against a buffer without Urea, we observed on SDS-PAGE that a significant amount of the protein stayed soluble. We continued with the purification on an affinity NiNTA column followed by size exclusion chromatography. We observed that IL-9R formed series of oligomers but with one predominant size.

We worked with human IL-9R containing eleven cysteines that should create disulphide bonds. *E. coli* usually does not form this type of chemical bond. The idea was that we create redox conditions during dialysis and disulphide bonds would be formed to decrease the presence of oligomers. We performed the same purification procedure as in the previous experiment and we noticed only a small improvement. Again, there were several oligomeric forms of IL-9R protein but with one predominant size. Both refolding experiments resulted in one main peak consisted of IL-9R according to SDS-PAGE analysis. It means that there is a high probability that additional optimization of the refolding process could lead to a properly folded protein.

We tried the refolding of IL-9 protein expresses from the cDNA cloned in the pBADM30 vector. It means that our protein is fused to an N-terminal GST protein together with an HisTag and TEV recognition site. We selected this construct due to its high expression yield. The assumption was that GST could promote better folding of IL-9 during refolding. We used the purification on GST column since, only the properly refolded GST affinity proteins binds to it. We were able to elute the protein from the GST column, but the subsequent purification on size exclusion chromatography was inconclusive. We didn't observe a "normal" chromatography profile. We need to repeat the experiment to take a statement about meaning of these results.

Considering the results of the first refolding attempts, it is necessary to further optimize the refolding conditions. We can also try others constructs and different dialysis buffers containing detergents and additives. Additionally, we could switch to other refolding techniques such as frequently used dilution.

However, the finding of the correct expression conditions in *E. coli* or the proper refolding process could take years. That is why we also launched the expression of both IL-9 and IL-9R in an eukaryotic expression system. We started with insect Schneider 2 (S2) cells as a result of a good experience in the laboratory with other interleukins and their receptors [55][56]. The main disadvantage of using insect cells is the time of growth. It takes approximately one and half month from the transfection of cells to reach a large-scale culture, and to be ready for the purification of the expressed protein. The handling of cells is more demanding and sensitive to contaminations. Altogether, the overall cost for transfection chemicals and cultivation media are much higher than in the case of *E. coli*. However, some proteins cannot be produced in a soluble form in *E. coli* and you need to use eukaryotic expression system. This was also our case.

We successfully expressed both IL-9 and extracellular part of IL-9R in insect S2 cells. Since, both proteins originally present in extracellular space in human bodies, we produced them as secreted proteins in S2 cells. This also means that the proteins could be modified by glycosylation. The SDS-PAGE of purified IL-9R did not show any huge change of mobility compared to calculated size of 29 kDa. On the contrary, the expected size of IL-9 protein was 15 kDa but we could clearly identify several forms of purified protein on SDS-PAGE gels. The size ranges from 17 to 24 kDa, suggesting the presence of glycosylation which is quite normal for secreted proteins. We could confirm the glycosylation either by mass-spectrometry or by enzymatical deglycosylation where the size should decrease on SDS-PAGE (the experiments were not done at the time of this thesis).

For the second objective, different characteristics were studied in order to biophysically characterize the expressed proteins. The first study performed was a measurement of Circular Dichroism (CD) spectra of both proteins alone. This method is important for examining protein secondary structure and folding properties of proteins that have been obtained using recombinant techniques. We observed that the IL-9 showed a spectrum that is typical for helical structures. The spectrum of IL-9R was not that definite, displaying a mixture of helices and beta-sheets. Nevertheless, both spectra were in agreement with the predicted structures based on other interleukins and their receptors.

The IL-9R is a specific receptor of IL-9 and it is shown that this interaction is usually strong [57][58].We mixed both proteins in calculated equimolar concentration and tried to purify the formed complex by SEC. We could detect one main peak which

elution volume agreed with the size of complex according to general column calibration. The SDS-PAGE analysis also confirmed that the main peak included both proteins and that excess of proteins run in later fractions. The SEC analysis confirmed formation of stable IL-9/IL-9R complex.

Another examined protein feature was the melting temperature (Tm), this property is often linked to its stability and function. We measured the Tm of the IL-9R and the IL-9/IL-9R complex but not the IL-9 directly because it does not have any tryptophan in its sequence. Unfortunately, we used the nano differential scanning fluorometry technique implemented in Prometheus instrument, that is based on detection of tryptophan fluorescence. This method has advantage that it is label-free, easy to handle and analyse, fast, and you can run more samples in one run. We could measure the melting temperatures by other methods such as circular dichroism spectra of thermal shift assay. However, we decided that it was not necessary in our case. We could detect that the melting point of IL-9R alone was 47 °C but, more importantly, the Tm of IL-9R protein in the purified complex increased to 59 °C. It means that the temperature stability of the receptor is influenced by a presence of IL-9 confirming the formation of IL-9/IL-9R complex.

The final proof of interaction was the direct measurement of affinity between those two proteins. For such purpose, we chose the MicroScale Thermophoresis (MST) method since the measurements are fast, done in solution, easy to optimize with low sample consumption. This method is based on the measurement of the movement of fluorescently labelled in the small temperature gradient that is created by a laser. One type of this method is the using of the intrinsic fluorescence of tryptophan. We tested this type because IL-9 does not possess any tryptophan. Unfortunately, we experienced the autofluorescence of IL-9 that made it hard to analyse the results (data not shown). This was the reason why we continued with the protein labelling, that could be done in several ways. The popular approach is the non-covalent labelling of HisTag but we cannot applied it since both of our proteins had uncleavable HisTag. It means that the results could be influenced by transferring of labelling dye from one protein to another.

We needed to use the covalent labelling of primary amines that are present in lysine residues. Our previous experience with other interleukins and receptors guided us to start with labelling of the receptor instead of IL-9 as the interleukins tend to be destabilized by the labelling itself. This arrangement, labelled IL-9R titrated by increasing concentration of unlabelled IL-9, happened to be beneficial. The measured MST curves evinced proper binding event with the Kd value of 43 (±8) nM under the tested conditions. This value is very similar to the interaction of other interleukins with their primary specific receptors [57].

We started the crystallization process to acquire structural details of the binding between IL-9 and IL-9R. There are many commercial crystallization screens to test. We chose those screens that are commonly used by our and other groups. In the case of IL-9/IL-9R complex that was purified by SEC, we setup three plates – JCSG+, Crystal Screen HT, and JenaXP. Unfortunately, we did not obtain any crystal form of complex after 51 days of continuous controlling.

The expression yield of IL-9 was higher than of IL-9R so we had enough protein to begin with the crystallization of IL-9 alone. We setup four plates – JCSG+, Crystal Screen HT, JenaXP, and PACT premiere. On the contrary to the complex, we could detect the presence of two crystal forms of IL-9 alone after two days.

In first case, we observed the crystal “needles”. In second case, we detected the growth of shaped crystal that needs to be characterized to state if it is a protein crystal and what the diffraction quality is. Usually, the initial hits need to be further optimized to acquire enough crystals with good diffraction quality. The focus is to influence the crystal nucleation and growth conditions. There are many variables to change and no one can guess which parameter is the best one to try first. We could try to change the temperature of crystal growth or change the protein concentration. Another strategy is the addition of supplements such as other ions, second precipitant, and many others. Usually, the optimization strategy of first choice is to make a grid plate around the initial condition with changes of precipitant concentration and pH of buffer. We are going to optimize both of our initial conditions by this way.

5. Conclusion

The main purpose of this project was to discover the structure of the IL-9 and its receptor. This Interleukin has shown a positive linkage with inflammation, asthma and allergies. For this reason, the discovery of these structures is very important for the development of an effective inhibitor, in order to stop some of these processes, like asthma for example.

We achieved a good production of both these proteins in several *E. coli* strains. However, the proteins expressed were in an insoluble form. In order to refold these proteins, we proceeded to the execution of a dialysis followed by a size exclusion chromatography. Unfortunately, we couldn't refold the proteins into a soluble form, although, the obtained results were very promising. We should try other constructs or different dialysis buffers containing detergents and additives in future studies. Producing these proteins in a soluble form in *E. coli* would be the optimal outcome to achieve in future work, since the production of proteins with S2 cells is both more expensive and time-consuming.

We also used insect S2 cells to produce IL-9 and the extracellular part of its receptor. The expression of both these proteins was good as well. The first study performed with these produced proteins was a measurement of Circular Dichroism (CD) spectrum. This method revealed that both proteins were folded. After confirming that both the proteins were folded, we mixed them and purified the complex by size exclusion chromatography. With complex of IL-9 and IL-9R formed, we measure the melting temperature of the IL-9R alone and the complex. These measurements showed that the temperature stability of the receptor is influenced by the presence of IL-9, confirming the formation of IL-9/IL-9R complex. The final proof of interaction was the direct measurement of affinity between those two proteins. For such purpose, we choose the MicroScale Thermophoresis (MST). The value obtained was very similar to the interaction of other interleukins with their specific receptors.

Finally, we set some plates for crystallization with IL-9 alone and with the complex. After two days, we could observe the presence of two crystal forms in one plate and some crystal "needles" in another plate of IL-9. Our crystallization trials gave initial hits and our further efforts will be the optimization of crystallization conditions for complex and both proteins alone.

6. Conclusão

Este projeto teve como principal objetivo, descobrir a estrutura da Interleucina-9 e do seu recetor. Esta Interleucina mostrou uma ligação positiva com processos de inflamação, asma e alergia. Por este motivo, a descoberta de ambas estas proteínas são muito importantes para o desenvolvimento de um inibidor eficaz, de forma a que seja possível interromper certas doenças como a asma.

Conseguimos uma boa produção de ambas estas proteínas em várias estirpes de *E. coli*. No entanto, as proteínas expressas eram insolúveis. Com o fim de transformar estas proteínas insolúveis em solúveis, procedemos à execução de uma dialise seguida de uma Cromatografia de Exclusão por tamanhos. Infelizmente não foi possível fazer o refold destas para uma forma solúvel, embora os resultados obtidos tenham sido bastantes promissores. No futuro, de forma a solucionar este problema podemos tentar outras construções de vetores ou diferentes buffers de dialise que contenham detergentes ou aditivos. Produzir estas proteínas numa forma solúvel em *E. coli* seria o resultado ideal a ser alcançado em trabalhos futuros uma vez que a produção de proteínas em células de inseto S2 é muito cara e demorada.

Utilizamos células de inseto S2 para produzir a IL-9 e a zona extracelular do seu recetor. A expressão de ambas as proteínas também foram bastante positivas. O primeiro estudo realizado com estas foi uma Medida do Espectro de Dicroísmo Circular (CD), este método revelou que ambas as proteínas estavam enroladas corretamente. Após esta confirmação, misturamos estas proteínas e procedemos à purificação do complexo por uma Cromatografia de Exclusão por Tamanho. Com o complexo formado entre a IL-9 com o IL-9R foi possível medir a temperatura de fusão da IL-9 sozinha e do complexo, mostrando que a estabilidade do recetor é influenciada pela presença da IL-9. A prova final da interação do complexo foi medida diretamente através da Termoforese MicroScale (MST). O valor obtido nestas medições foi bastante semelhante aos valores de outras interleucinas com os seus recetores.

Finalmente, preparamos várias placas para a cristalização da IL-9 sozinha e do complexo. Após dois dias foi possível observarmos a presença de duas formas cristalinas numa placa e noutra placa diferente foi possível observar alguns cristais em forma de “agulhas”. Os testes seguintes serão focados na otimização das condições de cristalização.

7. Bibliography

[1] B. Deverman e P. Patterson, «Cytokines and CNS Development», *Neuron*, vol. 64, n. 1, pp. 61–78, 2009.

[2] Z. Zídek, P. Anzenbacher, e Kmonícková, «Current status and challenges of cytokine pharmacology», *Br J Pharmacol*, vol. 157, n. 3, pp. 342–361, 2009.

[3] J.-M. Zhang e J. An, «Cytokines, Inflammation and Pain», *Int Anesth. Clin.*, vol. 69, n. 2, pp. 482–489, 2009.

[4] M. Akdis *et al.*, «Interleukins, from 1 to 37, and interferon-γ: Receptors, functions, and roles in diseases», *J. Allergy Clin. Immunol.*, vol. 127, n. 3, pp. 701-721.e70, 2011.

[5] V. Bocci, «Interleukins: Clinical Pharmacokinetics and Practical Implications», *Clin. Pharmacokinet.*, vol. 21, n. 4, pp. 274–284, 1991.

[6] P. Maney e J. L. Owens, «Interleukin polymorphisms in aggressive periodontitis: A literature review», *J Indian Soc Periodontol*, vol. 19, n. 2, pp. 131–141, 2015.

[7] C. A. J. Vosshenrich e J. P. Di Santo, «Interleukin signaling.», *Curr. Biol.*, vol. 12, n. 22, pp. 760–763, 2002.

[8] C. Uyttenhove, C. PG, e V. Snick J., «No TitT cell growth and differentiation induced by interleukin-HP1/IL-6, the murine hybridoma/plasmacytoma growth factor», *J Exp Med.*, vol. 167, n. 4, pp. 1417–27, 1988.

[9] L. Hultner *et al.*, «No TitleMast cell growth-enhancing activity (MEA) is structurally related and functionally identical to the novel mouse T cell growth factor P40/TCGFIII (interleukin 9).», *Eur J Immunol*, vol. 20, n. 6, pp. 1413–6, 1990.

[10] V. Snick J. *et al.*, «Cloning and characterization of a cDNA for a new mouse T cell growth factor (P40)», *J Exp Med.*, vol. 169, n. 1, pp. 363–8, 1989.

[11] Y. Yang, S. Ricciardi, A. Ciarletta, J. Calvetti, K. Kelleher, e S. Clark, «Expression cloning of cDNA encoding a novel human hematopoietic growth factor: human homologue of murine T-cell growth factor P40», *Blood*, vol. 74, n. 6, pp. 1880–1884, 1989.

[12] H. P. Hauber e Q. Hamid, «The role of interleukin-9 in asthma», *Allergol. Int.*, vol. 54, n. 1, pp. 71–78, 2005.

[13] A. Kermouni *et al.*, «The IL-9 receptor gene (IL9R): genomic structure, chromosomal localization in the pseudoautosomal region of the long arm of the sex chromosomes, and identification of IL9R pseudogenes at 9qter, 10pter, 16pter, and 18pter.», *Genomics*, vol. 29, n. 2, pp. 371–82, 1995.

[14] J. Bauer, K. Liu, S. Lai, e M. Goldsmith, «Heteromerization of the gammac chain with the interleukin-9 receptor alpha subunit leads to STAT activation and prevention of apoptosis.», *J Biol Chem*, vol. 273, n. 15, pp. 9255–60, 1998.

[15] J. Boasberg *et al.*, «Features and roles of T helper 9 cells and interleukin 9 in immunological diseases», *Allergol Immunopathol*, vol. 47, n. 1, pp. 90–104, 2019.

[16] R. Goswami e M. H. Kaplan, «A Brief History of IL-9», *J. Immunol.*, vol. 186, n. 6, pp. 3283–3288, 2011.

[17] T. W. Lebien e T. F. Tedder, «B lymphocytes: How they develop and function», *Blood*, vol. 112, n. 5, pp. 1570–1580, 2008.

[18] G. Walsh, «Biopharmaceuticals: recent approvals and likely directions.», *Trends Biotechnol*, vol. 23, n. 11, pp. 553–8, 2005.

[19] C. Kollewe e A. Vilcinskas, «Production of recombinant proteins in insect cells», *Am. J. Biochem. Biotechnol.*, vol. 9, n. 3, pp. 255–271, 2013.

[20] R. Jayaraj e P. M. Smooker, «So you Need a Protein - A Guide to the Production of Recombinant Proteins», *Open Vet. Sci. J.*, vol. 3, n. 1, pp. 28–34, 2014.

[21] M. Wacker *et al.*, «N-linked glycosylation in Campylobacter jejuni and its functional transfer into E. coli», *Science (80-.).*, vol. 298, n. 5599, pp. 1790–3, 2002.

[22] P. Wingfield, «Overview of the purification of recombinant proteins produced in Escherichia coli», *Curr Protoc Protein Sci*, vol. 30, n. 1, 2003.

[23] F. Beneyx, «Keeping up with protein folding», vol. 3, n. 6, 2004.

[24] G. Cereghino e J. Cregg, «Applications of yeast in biotechnology: protein production and genetic analysis», *Curr Opin Biotechnol*, vol. 10, n. 5, pp. 422–7, 1999.

[25] P. Sudbery, «The expression of recombinant proteins in yeasts», *Curr. Opin. Biotechnol.*, vol. 7, n. 5, pp. 517–524, 1996.

[26] Invitrogen, «Schneider 2 (S2) Cells», vol. 2, 2002.

[27] C. Becker-Pauly e W. Stocker, «Insect cells for heterologous production of recombinant proteins», *Insect Biotechnol*, vol. 2, pp. 197–209, 2011.

[28] A. Clermont *et al.*, «Cloning and expression of an inhibitor of microbial metalloproteinases from insects contributing to innate immunity», *Biochem J*, vol. 382, n. Pt 1, pp. 315–322, 2004.

[29] L. Cherbas, R. Moss, e P. Cherbas, «Chapter 9 Transformation Techniques for Drosophila Cell Lines», *Mothods Cell Biol.*, vol. 44, pp. 161–179, 1994.

[30] I. Scheneider, «Cell lines derived from late embryonic stages of Drosophila melanogaster. Schneider I», *J Embryol Exp Morphol*, vol. 27, n. 2, pp. 353–65, 1972.

[31] M. Vedadi, C. Arrowsmith, A. Allali-Hassani, G. Senisterra, e G. Wasney, «Biophysical characterization of recombinant proteins: A key to higher structural genomics success», *J. Struct. Biol.*, vol. 172, n. 1–2, pp. 107–119, 2010.

[32] L. A., W. L., e W. B. A., «DICHROWEB: an interactive website for the analysis of protein secondary structure from circular dichroism spectra», *Bioinforma. Appl. Note*, vol. 18, n. 1 2002, pp. 211–212, 2001.

[33] C. O. Fagain, «Protein Stability and Stabilization of Protein Function», *Landes Biosci.*, 1997.

[34] S. A. I. Seidel *et al.*, «Microscale thermophoresis quantifies biomolecular interactions under previously challenging conditions», *Methods*, vol. 59, n. 3, pp. 301–315, 2013.

[35] M. Jerabek-Willemsen *et al.*, «MicroScale Thermophoresis: Interaction analysis and beyond», *J. Mol. Struct.*, vol. 1077, pp. 101–113, 2014.

[36] A. McPherson, «Protein crystallization», *Methods Mol. Biol.*, vol. 1607, pp. 17–50,

2017.

[37] E. C. Nowak *et al.*, «IL-9 as a mediator of Th17-driven inflammatory disease», *J. Exp. Med.*, vol. 206, n. 8, pp. 1653–1660, 2009.

[38] P. Kauppi, T. Laitinen, V. Ollikainen, H. Mannila, L. A. Laitinen, e J. Kere, «The IL9R region contribution in asthma is supported by genetic association in an isolated population», *Eur. J. Hum. Genet.*, vol. 8, n. 10, pp. 788–792, 2000.

[39] H. P. Sørensen e K. K. Mortensen, «Soluble expression of recombinant proteins in the cytoplasm of Escherichia coli», *Microb. Cell Fact.*, vol. 4, n. Figure 1, pp. 1–8, 2005.

[40] G. L. Rosano e E. A. Ceccarelli, «Recombinant protein expression in Escherichia coli: Advances and challenges», *Front. Microbiol.*, vol. 5, n. APR, pp. 1–17, 2014.

[41] H. Tegel, J. Ottosson, e S. Hober, «Enhancing the protein production levels in Escherichia coli with a strong promoter», *FEBS J.*, vol. 278, n. 5, pp. 729–739, 2011.

[42] «GST-tagged Proteins – Production and Purification». [Em linha]. Disponível em: https://www.thermofisher.com/pt/en/home/life-science/protein-biology/protein-biology-learning-center/protein-biology-resource-library/pierce-protein-methods/gst-tagged-proteins-production-purification.html. [Acedido: 07-Out-2019].

[43] R. T. Hay, «SUMO: A history of modification», *Mol. Cell*, vol. 18, n. 1, pp. 1–12, 2005.

[44] LifeSensors, «SUMOstar™ Gene Fusion Technology», vol. 8845, n. June, 2007.

[45] L. Dumon-Seignovert, G. Cariot, e L. Vuillard, «The toxicity of recombinant proteins in Escherichia coli: A comparison of overexpression in BL21(DE3), C41(DE3), and C43(DE3)», *Protein Expr. Purif.*, vol. 37, n. 1, pp. 203–206, 2004.

[46] M. R. Dyson, S. P. Shadbolt, K. J. Vincent, R. L. Perera, e J. McCafferty, «Production of soluble mammalian proteins in Escherichia coli: Identification of protein features that correlate with successful expression», *BMC Biotechnol.*, vol. 4, pp. 1–18, 2004.

[47] G. Ren, N. Ke, e M. Berkmen, «Use of the SHuffle strains in production of proteins», *Curr. Protoc. Protein Sci.*, vol. 2016, n. August 2016, pp. 5.26.1-5.26.21, 2016.

[48] E. LaVallie, E. DiBlasio, S. Kovaciv, K. Grant, P. Schendel, e J. McCoy, «A thioredoxin gene fusion expression system that circumvents inclusion body formation in the E. coli cytoplasm.», *Biotchnology (NY)*, vol. 11, n. 2, pp. 187–93, 1993.

[49] R. B. Kapust e D. S. Waugh, «Escherichia coli maltose-binding protein is uncommonly effective at promoting the solubility of polypeptides to which it is fused», *Protein Sci.*, pp. 1668–1674, 1999.

[50] M. Hammarström, N. Hellgren, S. van den Berg, H. Berglund, e T. Härd, «Rapid screening for improved solubility of small human proteins produced as fusion proteins in Escherichia coli», *Protein Sci.*, vol. 11, n. 2, pp. 313–321, 2009.

[51] G. J. Gopal e A. Kumar, «Strategies for the production of recombinant protein in escherichia coli», *Protein J.*, vol. 32, n. 6, pp. 419–425, 2013.

[52] «Takara», *Chaperone plasmid set*. [Em linha]. Disponível em:

https://www.takarabio.com/products/protein-research/expression-vectors-and-systems/protein-folding-kits/chaperone-plasmid-set.

[53] K. Nishihara, M. Kanemori, M. Kitagawa, H. Yanagi, e T. Yura, «Chaperone coexpression plasmids: Differential and synergistic roles of DnaK-DnaJ-GrpE and GroEL-GroES in assisting folding of an allergen of Japanese cedar pollen, Cryj2, in Escherichia coli», *Appl. Environ. Microbiol.*, vol. 64, n. 5, pp. 1694–1699, 1998.

[54] S. M. Singh e A. K. Panda, «Solubihzation and refolding of bacterial inclusion body proteins», *J. Biosci. Bioeng.*, vol. 99, n. 4, pp. 303–310, 2005.

[55] P. Mikulecký *et al.*, «Crystal structure of human interferon-γ receptor 2 reveals the structural basis for receptor specificity»:, *Acta Crystallogr. Sect. D Struct. Biol.*, vol. 72, n. 9, pp. 1017–1025, 2016.

[56] J. Zahradník *et al.*, «Flexible regions govern promiscuous binding of IL-24 to receptors IL-20R1 and IL-22R1», *FEBS J.*, pp. 1–16, 2019.

[57] P. Mikulecký *et al.*, «Increasing affinity of interferon- γ receptor 1 to interferon- γ by computer-aided design», *Biomed Res. Int.*, vol. 2013, 2013.

[58] J. Lee, R. Horuk, G. C. Rice, G. L. Bennett, T. Camerato, e W. I. Wood, «Characterization of two high affinity human interleukin-8 receptors», *J. Biol. Chem.*, vol. 267, n. 23, pp. 16283–16287, 1992.