



UNIONE EUROPEA
Fondo Sociale Europeo



REACT EU



UNIVERSITY OF "G. D'ANNUNZIO"

Chieti-Pescara

School of Advanced Studies "G. D'Annunzio"

PhD course in Biomolecular and Pharmaceutical Sciences

Curriculum: Molecular Medicine

XXXVII cycle

**“Development of SIL1 Protein Replacement
Therapy for Marinesco-Sjögren Syndrome:
production, characterization and cellular
delivery of recombinant protein”**

Department of Medical, Oral and Biotechnological Sciences

Scientific Disciplinary Field: BIOS-07/A

PhD Student

Laura Amodei

Coordinator

Prof. Assunta Pandolfi, PhD

Tutor

Prof. Luca Federici, PhD

Academic Years 2021/2024

ABSTRACT

Marinesco-Sjögren syndrome (MSS) is a rare autosomal recessive disorder characterized by cerebellar ataxia, myopathy, and cataracts. It is mainly caused by mutations in the SIL1 gene, which encodes a nucleotide exchange factor for BiP, a key endoplasmic reticulum (ER) chaperone. SIL1 deficiency causes the accumulation of misfolded proteins, leading to cellular stress and ultimately cell death. The present study aimed to develop a protein replacement therapy for MSS using cell-penetrating peptides (CPPs) fused to the SIL1 protein. To this end, we evaluated the expression of SIL1 lacking the initial 31 N-terminal amino acids fused to various CPPs. These constructs were cloned into the pGEX6P1 vector, resulting in the production of recombinant proteins with an N-terminal GST tag. We optimized expression in BL21(DE3) bacterial cells and developed an affinity chromatography purification protocol for the selected recombinant SIL1 fusion protein. We provided evidence of correct folding, thermal stability, and tetrameric structure of the protein by circular dichroism and dynamic light scattering.

Safety and efficacy assessments were conducted to evaluate the suitability of SIL1-TAT for both *in vitro* and *in vivo* applications. Preliminary *in vivo* studies using the *woozy* mouse model demonstrated promising protein distribution in peripheral tissues, particularly spleen and muscle, suggesting potential systemic distribution. However, limited penetration of SIL1-TAT into the central nervous system (CNS) was observed, highlighting a critical challenge for effective CNS-targeted therapy. Complementary *in vitro* immunofluorescence microscopy revealed predominantly extracellular localization of SIL1-TAT in HF-MSS patient cells, suggesting potential limitations in cellular uptake. Furthermore, this research delved into the molecular pathogenesis of MSS by performing transcriptomic analysis on patient fibroblasts. Transcriptomic analysis revealed significant disruptions in extracellular matrix (ECM) organization, membrane trafficking, and cell adhesion. We also aimed to characterize the *woozy* mouse model, establishing its utility both for elucidating disease mechanisms and evaluating therapeutic interventions.

This project demonstrates the potential of recombinant SIL1-TAT as a basis for protein replacement therapy in MSS.

Contents

1. INTRODUCTION	8
1.1 Marinesco-Sjögren Syndrome	8
1.1.1 Understanding the etiology and phenotype of the disease.....	8
1.1.2 Mouse models of Marinesco-Sjögren Syndrome: key tools for understanding pathophysiology.....	9
1.1.3 The role of SIL1 in the pathogenesis of Marinesco-Sjögren Syndrome	10
1.1.4 Management strategies and current treatment for Marinesco-Sjögren Syndrome.....	13
1.2 Protein Replacement Therapy: an emerging approach to genetic diseases.....	15
1.2.1 The power of Protein Replacement Therapy: key mechanisms.....	15
1.2.2 Potential drawbacks of Protein Replacement Therapy	16
1.2.3 The promise of Protein Replacement Therapy: clinical applications.....	19
1.3 Recombinant protein technology: a key to Protein Replacement Therapies.....	21
1.3.1 Selecting the right host for recombinant protein production	21
1.3.1.1 Vector selection for recombinant protein expression.....	23
1.3.2 Tailoring recombinant protein expression in Escherichia Coli.....	24
1.3.2.1 Optimizing recombinant protein purification: key considerations.....	25
1.3.3 Methods for evaluating recombinant protein purity, aggregation, and structure	27
2. AIM OF THE PROJECT	30
3. MATERIAL AND METHODS	31
3.1 Generation of recombinant SIL1 constructs for protein expression	31
3.2 BL21(DE3) and BL21(DE3)pLysS Strains: Expression System Details	32
3.3 Expression and purification of recombinant SIL1-KDEL using the pET21a+ vector in E. coli BL21(DE3)pLysS	32
3.3.1 Preparation of BL21(DE3)pLysS Competent cells.....	32
3.3.2 Heat-shock transformation of BL21(DE3)pLysS bacterial cells	32
3.3.3 Optimal induction parameters for SIL1-KDEL in BL21(DE3)pLysS	33
3.3.4 Cell lysis and differential recovery of soluble and insoluble SIL1-KDEL recombinant protein.....	34
3.3.5 Purification of SIL1-KDEL by AKTA Prime Plus Chromatography system	34
3.4 Expression and purification of recombinant SIL1-CPPs using the pGEX6P-1 Vector in E. coli BL21(DE3).....	35
3.4.1 Heat-shock transformation of BL21(DE3) bacterial cells.....	35
3.4.2 Expression of SIL1-CPPs using pGEX6P-1 in BL21(DE3)	35
3.4.3 Soluble and insoluble recombinant SIL1-CPPs recovery	36
3.4.4 Optimization of purification of GST-tagged SIL1-TAT protein by affinity chromatography	36

3.4.5 Purification of GST-tagged SIL1-TAT and SIL1-WT by Fast Protein Liquid Chromatography	38
3.5 SDS-PAGE electrophoresis	39
3.6 Protein detection by Western Blotting	39
3.7 Mass spectrometry analysis of SIL1-TAT recombinant protein	40
3.8 Characterization of protein stability and folding using CD spectroscopy	40
3.9 Evaluation of recombinant SIL1 protein aggregation and molecular size distribution by dynamic light scattering	41
3.10 X-ray Crystallography of SIL1-TAT protein	42
3.11 Determination of Endotoxin Levels in Purified SIL1-TAT and SIL1-WT	42
3.12 Cells culture	43
3.13 Analysis of extracellular matrix remodelling in Sil1-deficient fibroblasts	43
3.14 Evaluation of <i>in vitro</i> cytotoxicity of SIL1-TAT	43
3.15 Immunofluorescence to evaluate SIL1-TAT cellular uptake	44
3.16 Mouse model of Marinesco-Sjogren syndrome	44
3.16.1 Rotarod Acceleration Test	45
3.16.2 Beam Walking Test	45
3.16.3 <i>In vivo</i> SIL1-TAT administration and tissue collection	45
4. RESULTS	46
4.1 Selection of TAT and KDEL sequences for SIL1 delivery	46
4.2 Impact of induction parameters on SIL1-KDEL production in BL21(DE3)pLysS	46
4.2.1 Purification of recombinant SIL1-KDEL from inclusion bodies	51
4.3 Development of a pGEX6P-1-based expression system for SIL1-CPP fusion proteins	52
4.3.1 Expression of GST-tagged SIL1-CPP constructs in E. coli BL21(DE3)	53
4.3.2 Optimization of GST-tagged SIL1-TAT purification: key parameters and scale-up	55
4.3.2.1 Impact of binding buffer composition on SIL1-TAT purification	55
4.3.2.2 Impact of resin-to-protein ratio on SIL1-TAT purification	56
4.3.2.3 Enhanced purification of SIL1-TAT via automated FPLC	58
4.3.3 Evaluation of SIL1-TAT purity by mass spectrometry	59
4.3.3.1 Addressing GroEL contamination in SIL1-TAT purification	64
4.3.4 Structural and Biophysical Characterization of SIL1-TAT	65
4.3.4.1 Investigation of SIL1-TAT folding stability: buffer, thermal and cryopreservation effects	67
4.3.4.2 Structure determination of SIL1-TAT via X-ray crystallography	71
4.4 Structural integrity and stability studies of purified SIL1-WT protein	72
4.5 Assessing the therapeutic potential of SIL1-TAT: <i>in vivo</i> and <i>in vitro</i> studies	75
4.5.1 Preliminary tissue distribution of SIL1-TAT recombinant protein in mice	75
4.5.2 <i>In vitro</i> studies on cytotoxicity and cellular uptake of SIL1-TAT	76

4.6 Molecular and functional characterization of Marinesco-Sjögren syndrome	78
4.6.1 Unravelling disease mechanisms through transcriptomic analysis	79
4.6.2 Baseline motor function in the Woozy Mouse	83
5. DISCUSSION.....	85
5.1 Development of a robust purification protocol for SIL1: from structural characterization to therapeutic interventions for MSS.....	85
5.2 Preliminary evaluation of the safety and efficacy of SIL1: insights from <i>in vitro</i> and <i>in vivo</i> studies.....	87
5.3 Towards a complete characterization of Marinesco-Sjogren syndrome	89
6. CONCLUSION	91
ABBREVIATIONS.....	92
FUNDING	94
ACKNOWLEDGEMENTS	94
REFERENCES	95

1. INTRODUCTION

1.1 Marinesco-Sjögren Syndrome

Marinesco-Sjögren Syndrome (MSS) is a rare autosomal recessive neurological condition that mostly affects children. The disease's onset can be detected in the first months of a newborn's life when the first motor milestones are unmet.¹ Despite being a severely disabling disorder, the symptoms tend to stabilize over time, and patients can live to old age.

The syndrome was named for the two doctors who discovered it in the early 1900s: Georg Marinesco and Harald Sjögren. Romanian neurologist Marinesco reported a group of patients who had symptoms such as muscle weakness, cataracts, and cerebellar ataxia². Harald Sjögren, a Swedish neurologist, independently reported his findings of comparable instances³.

1.1.1 Understanding the etiology and phenotype of the disease

The only known genetic cause of MSS is mutations of the SIL1 gene, found on chromosome 5q31.2^{4,5}. Both homozygous or compound heterozygous point mutations and larger genomic deletions that result in protein loss in MSS patients are the most commonly reported mutations in the SIL1 gene⁶.

The defining characteristics of MSS are congenital cataracts, cerebellar ataxia, and myopathy (Figure 1)^{4,5}. Muscle fibres degenerate with infiltration of fibroadipose tissue and altered fibre size, as well as abnormal interfibrillar membranous structures⁷. Cerebellar atrophy, deriving from the progressive degeneration of Purkinje cells and granular cells in the cerebellum, is the primary driver of this syndrome⁶.

Other differentiating traits include mild to severe mental retardation and hypergonadotropic hypogonadism⁸. Short stature, skeletal anomalies⁹, nystagmus, strabismus, and dysarthria are common findings¹⁰.

Marinesco-Sjogren syndrome "clinical triad"

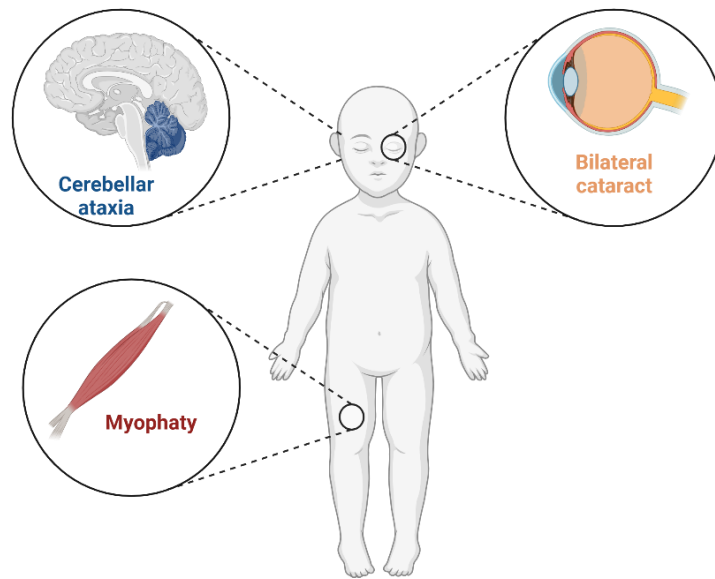


Figure 1. Characteristic clinical triad of Marinesco- Sjögren Syndrome. *From Cell-penetrating SIL1 protein replacement therapy for Marinesco-Sjogren syndrome. 21st Scientific Convention TELETHON Poster, Amodei L. et al.*

A clinical assessment of the primary symptoms is used to make the diagnosis, which includes muscle biopsies to detect myopathy, ophthalmologic exams to detect cataracts, and magnetic resonance imaging to detect cerebellar atrophy¹. Genetic tests that can detect biallelic mutations in the SIL1 gene are used to confirm clinical results ¹.

Due to its rarity, MSS is often misdiagnosed or misunderstood, which makes accurate diagnosis and treatment difficult.

1.1.2 Mouse models of Marinesco-Sjögren Syndrome: key tools for understanding pathophysiology

The same histopathological findings characteristic of MSS patients are present in the spontaneous mutant *woozy* mouse, which lacks a functional Sil1 gene¹¹. These mice develop cerebellar ataxia between 12 and 16 weeks of age, resulting from degeneration of Purkinje cells in cerebellar lobules I to VIII ¹¹. *Woozy* mice also show progressive degenerative myopathy, functionally similar to the human MSS phenotype, with morphological

abnormalities in the sarcoplasmic reticulum, autophagic vacuoles, enlarged mitochondria, and altered myonuclei ⁷.

The SIL1Gt mouse, generated via a gene trapping approach on the C57BL/6J strain, is another model of MSS¹¹. The mutant SIL1 transcript is truncated after exon 7, confirming the role of SIL1 in skeletal muscle physiology and neurodegeneration, producing a phenotype very similar to that of the *woozy* mouse¹². Around 24 weeks of age, these mice begin to show a progressive loss of muscle strength, accompanied by myofibrillar disorder with fatty infiltration and structural abnormalities in the muscle tissue, including irregular fibers ¹².

1.1.3 The role of SIL1 in the pathogenesis of Marinesco-Sjögren Syndrome

The Sil1 protein acts as a nucleotide exchange factor (NEF) for an endoplasmic reticulum (ER) chaperone, immunoglobulin-binding protein (BiP) ¹³. BiP is a stress-inducible molecular chaperone that controls a variety of essential processes in the endoplasmic reticulum, including translocation of nascent proteins into the endoplasmic reticulum, accurate subsequent folding of newly synthesized proteins in the endoplasmic reticulum lumen, elimination of proteins that fail to mature properly, response to cell stress, and calcium homeostasis ¹⁴.

Sil1 catalyses the release of ADP from BiP by directly interacting with BiP's ATPase domain (Figure 2). ATP hydrolysis stabilizes BiP's binding and promotes protein folding, while the exchange of ADP for ATP allows the release of the folded substrate^{13,15}.

ATP-bound BiP cycle

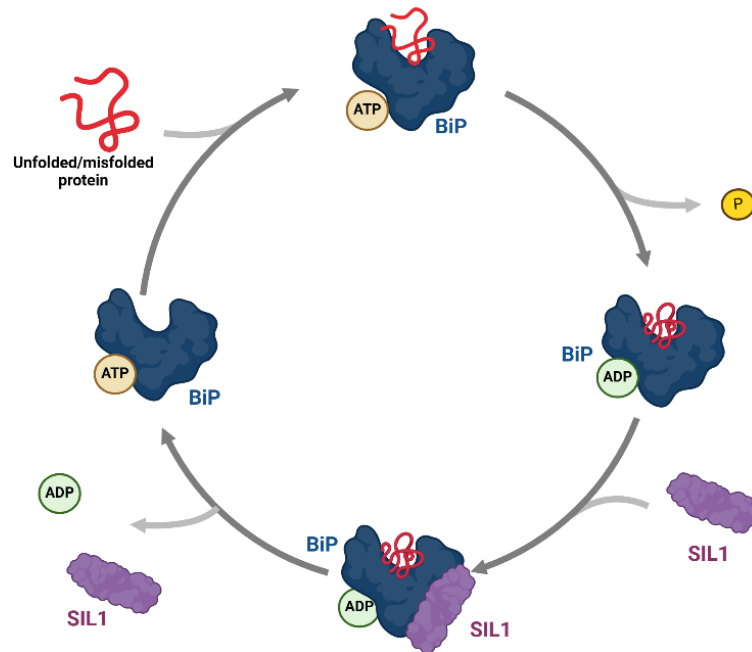


Figure 2. The ATP-bound BiP cycle. Sil1 facilitates ADP release from BiP's ATPase domain. ATP hydrolysis strengthens BiP binding and aids protein folding. ADP to ATP exchange triggers folded substrate release. Defective nucleotide exchange, as seen in MSS, results in persistent BiP-client protein association. *From Cell-penetrating SIL1 protein replacement therapy for Marinesco-Sjogren syndrome. 21st Scientific Convention TELETHON Poster, Amodei L. et al.*

If this nucleotide exchange is defective, as in MSS, BiP remains associated with its client protein, leading to the accumulation of unfolded proteins, ER stress, and activation of the unfolded protein response (UPR)^{13,16}.

The UPR is a signalling pathway aimed at restoring ER homeostasis (Figure 3). This pathway reduces protein synthesis, increases ER folding capacity, and enhances the disposal of unfolded proteins via ER-associated degradation (ERAD). Three ER transmembrane proteins initiate the UPR: endoplasmic reticulum kinase-like PRK (PERK), inositol-requiring enzyme 1 (IRE1), and activating transcription factor 6 (ATF6)¹⁷. By interacting with these proteins, BiP inactivates them. When unfolded proteins accumulate, BiP disengages from ER stress sensors, allowing activation of cytosolic effector domains. Activated PERK, ATF6, and IRE1 have specific cytosolic activities related to the ability to activate their respective transducers, also influencing protein translation, cellular metabolism, and cell survival or death programs^{17,18}.

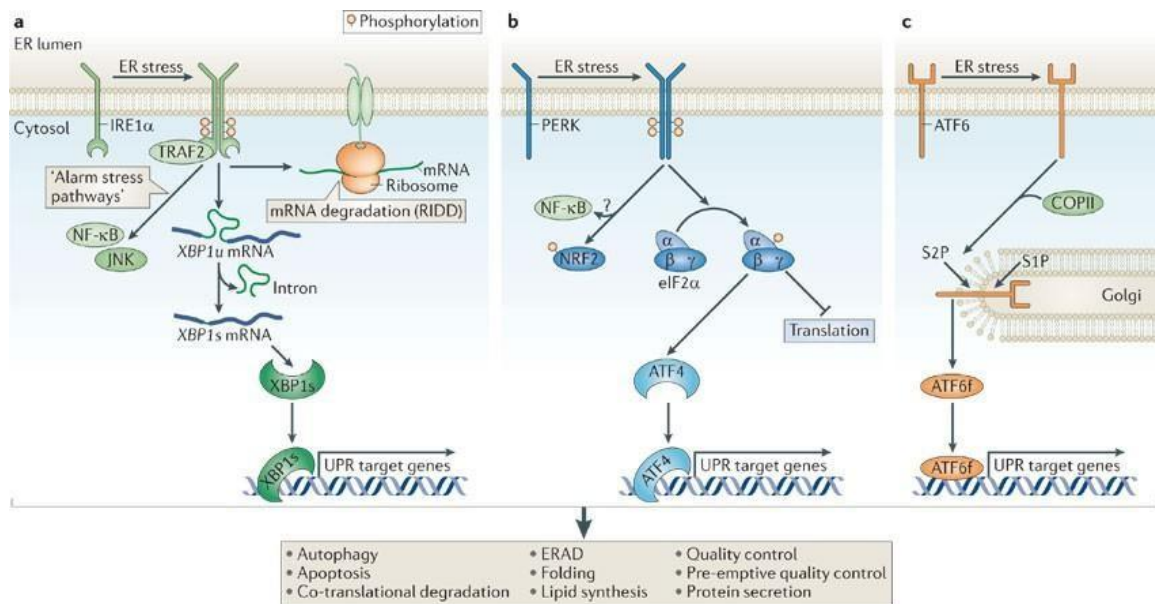


Figure 3¹⁹. Unfolded protein response. Three ER transmembrane proteins, PERK, IRE1, and ATF6, initiate the UPR. BiP normally inactivates these proteins, but upon accumulation of unfolded proteins, BiP disengages, activating the cytosolic effector domains of PERK, IRE1, and ATF6. These activated domains influence protein translation, cellular metabolism, and cell survival/death programs. *From “The unfolded protein response: controlling cell fate decisions under ER stress and beyond” Hetz C. DOI:10.1038/nrm3270*

Since BiP's ATPase activity is required for at least some of its functions, it is reasonable to assume that SIL1 regulates endoplasmic reticulum-associated processes by controlling the BiP ATPase cycle and that the pathogenic mechanism involved in Marinesco-Sjögren syndrome is impaired activation of BiP.

A recent proteomic study of skin fibroblasts from a young MSS patient (HF-MSS) with the SIL1 R111X mutation demonstrated mild UPR activation and altered cellular metabolism. The spliceosomal complex, ribosomal proteins, and RNA translation initiation were all clearly downregulated in HF-MSS cells that showed a senescence-like condition when SIL1 was absent. Lipid metabolism was also affected, exhibiting both increased beta-oxidation events and decreased beta-oxidation²⁰. These findings suggest that MSS may not solely be a cellular proteostasis disorder but also involves broader cellular dysfunction that impacts organelle dynamics, lipid metabolism, and energy production. The molecular mechanisms behind the pathology generally remain unknown, and it is still unclear why mutations in the widely expressed Sil1 protein only seem to affect certain tissues.

1.1.4 Management strategies and current treatment for Marinesco-Sjögren Syndrome

Currently, there is no cure for this disease. Except for cataracts, which can be surgically fixed, supportive therapies are the only way to prevent the progression of any other issues. Hypergonadotropic hypogonadism is treated with hormone therapy, and psychomotor difficulties are typically tracked and monitored by physiotherapists and neurologists¹. Given the high degree of patient variability and the rarity of the disease, it is undoubtedly difficult to comprehend the pathological mechanisms.

However, studies have shown that inhibiting the PERK signalling pathway, one of the key UPR pathways protects Purkinje cells, which would otherwise be vulnerable in the absence of SIL1^{21,22}. The PERK inhibitor (GSK2606414) prevents neurodegeneration and the onset of motor deficits, extending the asymptomatic phase while improving muscle pathology and motor performance during the symptomatic phase²². The cellular recovery process may encourage the synthesis of this alternative BiP cofactor, restoring BiP-mediated protein folding and synaptic protein transport, as evidenced by the elevated levels of the protein ORP150 in recovered Purkinje cells^{23,24}.

Unfortunately, it has been demonstrated that complete inhibition of PERK by GSK2606414 exhibits significant pancreatic toxicity. Other drugs, such as trazodone, dibenzoylmethane, and tauroursodeoxycholic acid, have been tested and found to be nontoxic but with only weak inhibition of PERK²⁵.

Misfolding diseases are caused by the accumulation of misfolded proteins, which damage cells and can lead to loss of function. To overcome this problem, small compounds called chemical chaperones can restore the altered protein to its correct conformation²⁶.

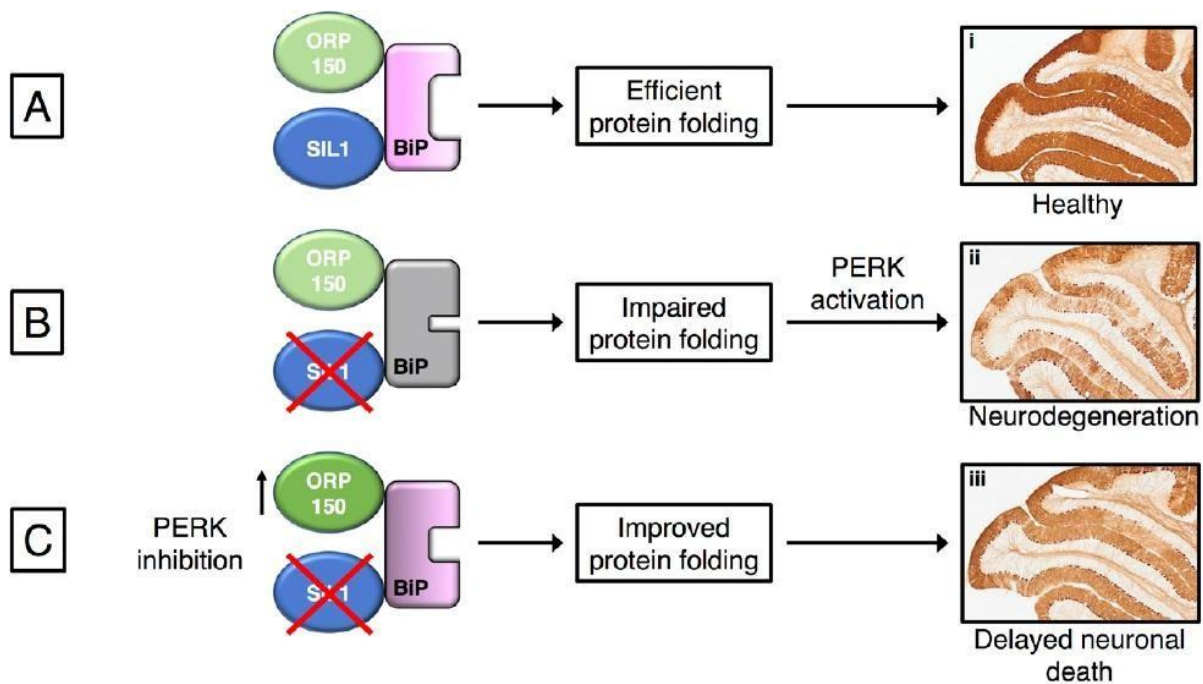


Figure 4²⁴. PERK inhibitor prevents neurodegeneration. A) Under normal SIL1 functioning conditions, Purkinje cells are in a healthy state; B) In the absence of SIL1, ORP150 function is insufficient, the PERK pathway is activated and neurodegeneration occurs; C) With PERK inhibition, ORP150 expression increases which, although still insufficient to sustain the BiP-dependent folding process, delays the cell death process. From “Neuroprotective modulation of the unfolded protein response in Marinesco-Sjögren syndrome: PERK signalling inhibition and beyond”. Restelli E. et al. DOI: 10.4103/1673-5374.243708.

Chemical chaperones were tested to restore the MSS phenotype in cultured cells. The study involved immortalized lymphoblastoid cell lines from patients with the p.Leu313fs SIL1 mutation. Chemical chaperones such as valproate, dexamethasone, 4-phenylbutyric acid, and tauroursodeoxycholic acid (TUDCA) were found to reduce tunicamycin-induced cell death. However, the treatment did not reduce ER stress markers²⁷.

New approaches that can more successfully participate in the treatment of these illnesses are needed to greatly enhance patients' quality of life.

1.2 Protein Replacement Therapy: an emerging approach to genetic diseases

Autosomal recessive genetic disorders often result from mutations or deletions in the affected genes, resulting in the loss or malfunction of the encoded proteins ²⁸. Protein replacement therapy (PRT) offers a promising solution to these protein deficiency disorders, providing a means to directly address the underlying genetic cause without the need for complex gene therapy techniques. This approach aims to restore necessary physiological functions and alleviate the debilitating symptoms associated with these genetic conditions.

1.2.1 The power of Protein Replacement Therapy: key mechanisms

PRT involves the administration of a therapeutic recombinant protein to patients with diseases caused by deficiencies of specific proteins ²⁹. This approach directly targets the biochemical defect, providing the body with an otherwise absent or malfunctioning protein. The administered proteins are typically produced *in vitro*, using techniques such as recombinant DNA technology, to create functional versions of the protein of interest.

By providing the missing or malfunctioning proteins, PRT can effectively compensate for the genetic deficiency and alleviate the associated health complications ³⁰. The mechanism of action of PRT can be explained through three key aspects:

- **Replacement of the missing protein:** In diseases where the body is unable to produce a functional protein due to genetic mutations (e.g., lysosomal storage), PRT directly provides the missing protein ³¹. This process compensates for the protein deficiency by providing the active form of the protein, which can then interact with cells and tissues.
- **Restoring cellular functions:** By replacing the missing or defective protein, PRT can restore normal cellular functions that are disrupted due to the absence of this protein. Once inside the cell, the protein can perform its catalytic or functional roles, helping with normal metabolic processes and cellular homeostasis^{32,33}.
- **Slowing or stopping disease progression:** Many genetic disorders are progressive, meaning that symptoms worsen over time due to ongoing cellular damage or metabolic disorders³⁴. PRT can help slow or even stop disease progression by replacing the missing protein and thereby mitigating the harmful effects of the underlying genetic defect³⁵.

In conditions where the body is unable to produce a functional protein due to genetic mutations, PRT provides a crucial lifeline. It aims not only to restore cellular function, but also to prevent disease progression.

1.2.2 Potential drawbacks of Protein Replacement Therapy

Although PRT offers significant advantages, it is not without limitations. The potential for immune responses to the therapeutic protein (especially if derived from non-human sources or if the protein is administered repeatedly over time) and limited tissue penetration (especially those involving hard-to-reach tissues, such as the central nervous system) are major challenges of this approach.

Immunogenicity reduction in drug therapeutics has been achieved through replacing non-human drugs with human sequences or improving manufacturing processes³⁶. However, these precautions are often insufficient, leading to the need for additional approaches. Poor protein solubility can be overcome by optimizing expression, purification, formulation, and solution conditions. Rational solubility engineering can identify mutations that minimize aggregation³⁷. Identifying antibody epitopes can also be challenging, but modifications can reduce binding by existing antibodies. Site-directed mutagenesis of antibody epitopes can reduce antigenicity, but repeated treatment may elicit new antibody responses³⁸.

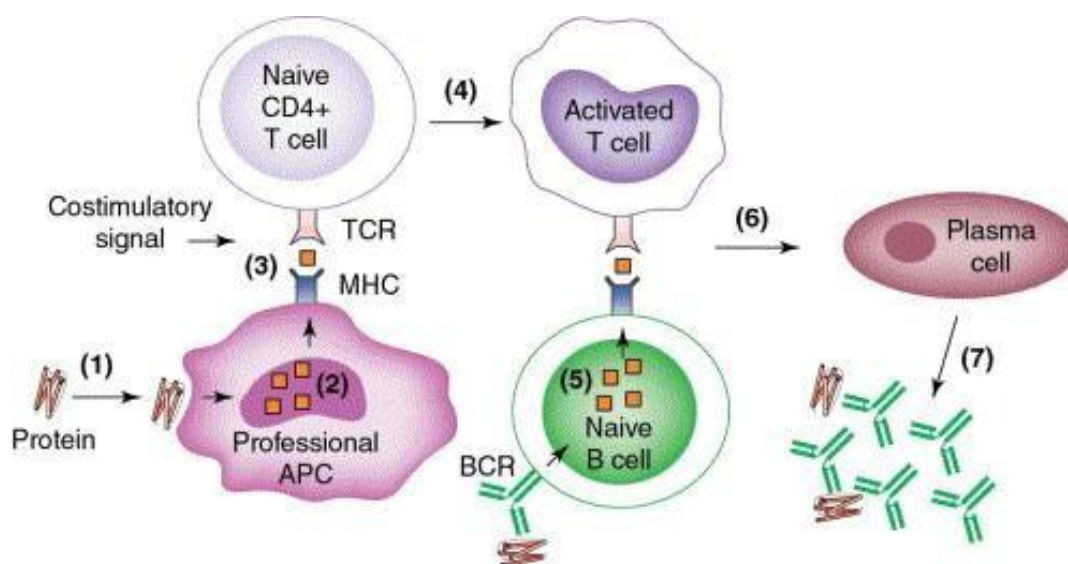


Figure 5³⁸. Targeting steps in the immune response to reduce protein drug immunogenicity. The process includes: (1) APC uptake, (2) antigen processing, (3) APC-T cell interaction (peptide-MHC-TCR complex and stimulation), (4) T cell maturation, (5) T cell-B cell interaction (protein binding to BCRs and peptide-MHC-TCR complex), (6) B cell maturation (isotype switching and affinity maturation), and (7) antibody binding. Blocking these steps is a potential strategy to improve protein therapeutics. From “*Rational design and engineering of therapeutic proteins*”, Marshall S.A. et al. DOI: [https://doi.org/10.1016/S1359-6446\(03\)02610-2](https://doi.org/10.1016/S1359-6446(03)02610-2).

Biomolecules are actively transported by the cell membrane. Cell penetrating peptides (CPPs) are a promising technology that facilitates protein transport across cell membranes, allowing for the development of a protein-based therapeutic modality for treating human disorders that have previously been avoided due to membrane permeability issues³⁹.

CPPs are short peptide sequences ranging from 5 to 30 aa that can bind to a variety of bioactive cargoes, including proteins and peptides, and have the capacity to pass through cell membrane barriers⁴⁰. CPPs can be categorized into linear or cyclic groups based on their characteristics. They can also be categorized based on their sequence, with three categories: cationic peptides, amphipathic peptides, and hydrophobic peptides^{41,42}. The mechanism of cell penetration is unclear and can vary based on the nature and type of CPPs, experimental setup, and fused cargo or drug. Two pathways have been proposed to explain cell penetration: direct penetration that involves charge interaction between negatively charged phospholipid and cationic CPPs⁴³ and endocytosis⁴⁴.

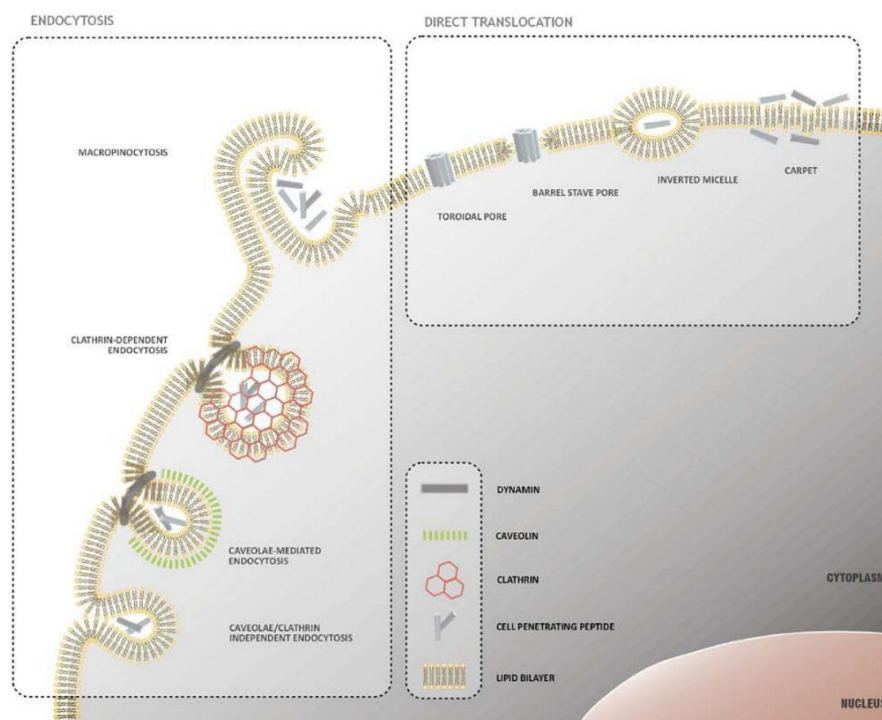


Figure 6⁴⁵. The schematic of cellular uptake mechanisms of CPP. From “Cell-Penetrating Peptides—Mechanisms of Cellular Uptake and Generation of Delivery Systems.” Trabulo, S. et al. DOI: 10.3390/ph3040961.

They observed that purified or chemically synthesized HIV-1 TAT (trans-activator of transcription) protein could be easily transferred intracellularly and localized in the nucleus without the use of a transfection reagent⁴⁶. The short basic domain of the HIV-TAT protein, an 11aa peptide (YGRKKRRQRRR, residues 47–57) enriched in arginine (R) and lysine (K) residues, was later identified by additional research as the protein region necessary for cellular uptake^{47,48}. The “TAT” peptide has been extensively studied and utilized due to its exceptional cell internalization properties. Nevertheless, the susceptibility of the native TAT peptide to enzymatic degradation, particularly by furin proteases, has limited its efficacy in certain biological contexts. To address this limitation, the synthetic analogue TATκ (YARKAARQARA) was developed, incorporating specific amino acid substitutions to eliminate furin cleavage sites, thereby enhancing its stability and potential for intracellular delivery^{48,49}.

While CPPs offer a valuable tool for direct membrane translocation, alternative strategies that exploit endogenous cellular trafficking pathways hold great promise. For example, the potential to exploit the KDEL receptor-mediated retrograde transport system for targeted intracellular protein delivery.

The KDEL receptor plays a critical role in maintaining ER homeostasis by recognizing the canonical Lys-Asp-Glu-Leu (KDEL) retrieval signal present on ER-resident proteins that have inadvertently escaped the ER. Upon recognition, the KDEL receptor facilitates retrograde transport of these proteins into the ER lumen^{50,51}. This intrinsic cellular mechanism raises intriguing questions regarding the potential for exogenous KDEL-containing proteins to interact with and be internalized by these receptors.

1.2.3 The promise of Protein Replacement Therapy: clinical applications

PRT has demonstrated remarkable success in treating a variety of genetic and metabolic diseases. Enzyme replacement therapy (ERT) is a specific type of PRT used to treat diseases caused by enzyme deficiencies³¹.

ERT has been shown to be particularly effective in treating lysosomal storage diseases, such as Gaucher disease, Fabry disease, and mucopolysaccharidosis^{52–54}, which are caused by mutations in the genes that encode enzymes responsible for breaking down cellular waste products. Without functional enzymes, toxic substrates build up in lysosomes, causing organ damage. PRT has revolutionized the treatment of these disorders with enzyme replacement therapies that significantly improve outcomes for affected patients.

Pompe disease, a type of glycogen storage disease, is caused by a deficiency of the enzyme acid alpha-glucosidase. This causes glycogen to build up in the muscles, leading to progressive muscle weakness and respiratory failure.³⁵ ERT has significantly improved survival rates and quality of life for patients.

The success of ERT for these genetic disorders has paved the way for the development of protein replacement therapy for a broader range of autosomal recessive disorders, offering promising solutions to address the underlying causes of protein deficiency conditions.

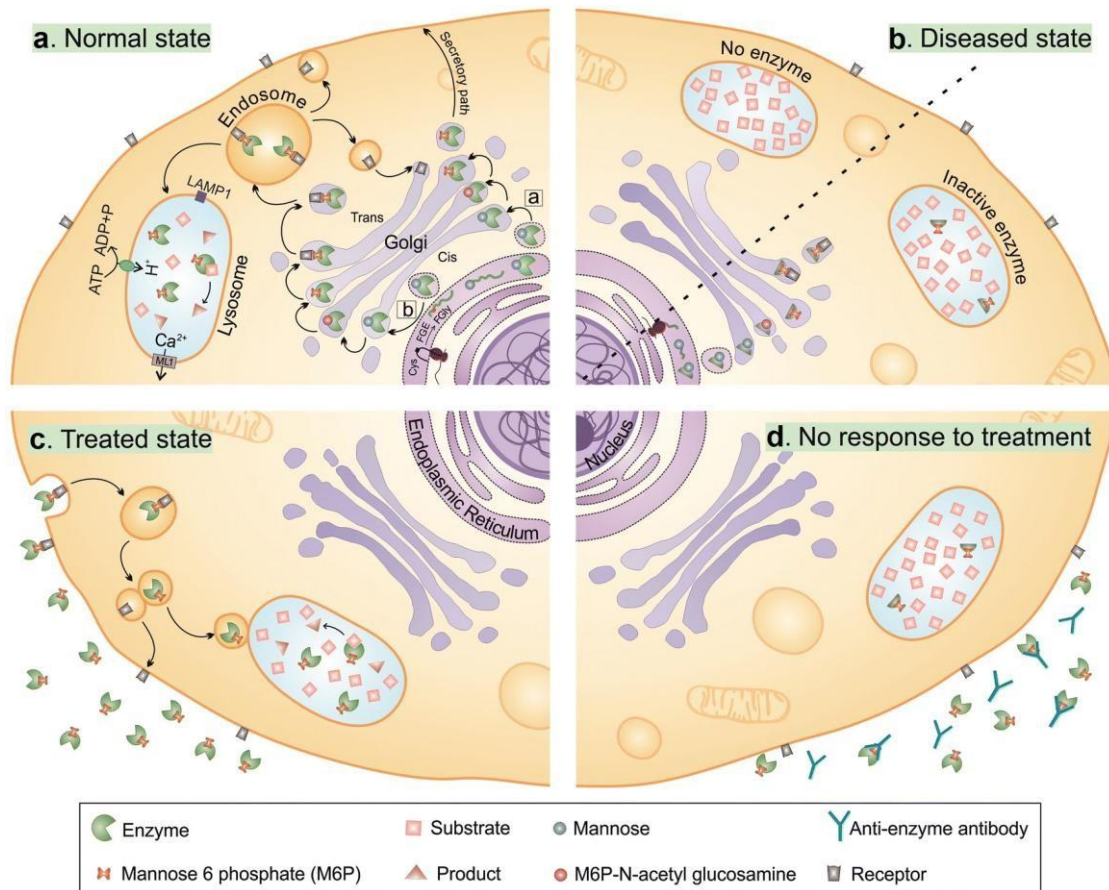


Figure 7⁵⁴. Schematic illustration of the mechanism of enzyme replacement therapy in mucopolysaccharidosis. (A) Healthy lysosomes process enzymes for transport via the mannose 6-phosphate receptors pathway. (B) In mucopolysaccharidosis, enzyme deficiencies lead to glycosaminoglycan buildup. (C) ERT utilizes the mannose 6-phosphate pathway to deliver recombinant enzyme to lysosomes. (D) ERT resistance can be caused by neutralizing antibodies targeting the therapeutic enzyme. From: “Targeted enzyme delivery systems in lysosomal disorders: an innovative form of therapy for mucopolysaccharidosis.” Safary, A., et al. DOI: <https://doi.org/10.1007/s00018-019-03135-z>.

Hemophilia is a genetic disorder characterized by a deficiency of clotting factors, leading to excessive bleeding. PRT in the form of clotting factor concentrates can prevent bleeding episodes and allow patients to live relatively normal lives⁵⁵. Although therapies for hemophilia A and B have been available for decades, ongoing advances continue to improve the safety and effectiveness of these treatments.

Overall, PRT has revolutionized the treatment of various diseases by replenishing missing proteins, improving patient health outcomes, and offering hope for the development of even better therapies in the future.

1.3 Recombinant protein technology: a key to Protein Replacement Therapies

Through heterologous expression in both prokaryotic and eukaryotic hosts, recombinant DNA technology has opened up a world of possibilities for producing both new and existing proteins. A gene encoding a desired protein is inserted into an expression system, where it is synthesized, purified, and its biological activity is assessed⁵⁶.

Recombinant protein production is a fundamental aspect of modern biotechnology, with applications ranging from therapeutic drug development to the study of protein structure and function. The development of recombinant protein production systems has been instrumental in addressing the need for high-quality proteins for PRT.

1.3.1 Selecting the right host for recombinant protein production

Choosing the best method for expressing a recombinant protein is critical to achieving the desired yields and quality. Protein mass, disulfide bond number, desired post-translational modifications, and protein destination should all be considered. The purified protein's intended use, which can be divided into structural investigations, *in vitro* activity tests, antibody production, and *in vivo* investigations, is also crucial. Errors may cause proteins to be poorly expressed or misfolded⁵⁷. Commonly used host organisms include *Escherichia coli*, yeast species (such as *Pichia pastoris*), insect cells (using the baculovirus expression system), and mammalian cells.

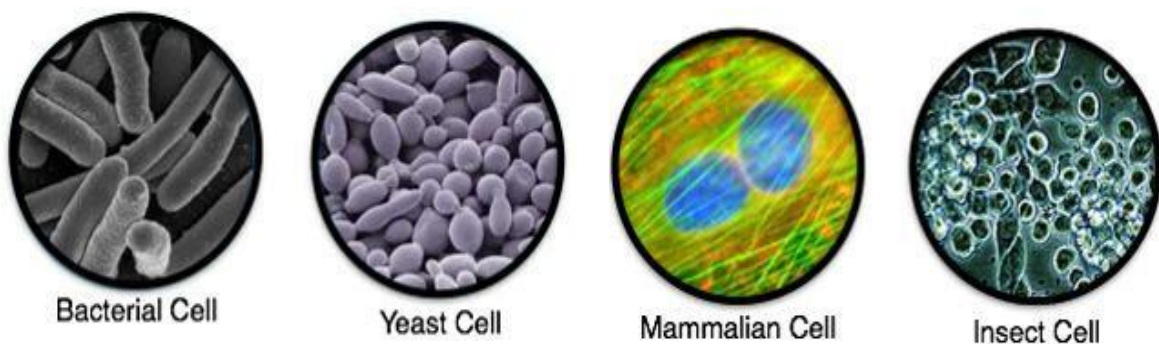


Figure 8. Commonly used host organisms used for Recombinant protein production.

Escherichia coli is by far the most widely utilized organism for recombinant protein production due to its rapid growth, ease of manipulation, and low-cost culture requirements. It is particularly suitable for expressing simpler proteins or those that do not require complex post-translational modifications (PTMs)⁵⁸. Bacterial systems' simplicity makes it possible to use simple, proven, and highly reproducible cloning and transformation methods⁵⁹. The main disadvantage of using *E. coli*, however, is that a large number of proteins expressed in this system create inclusion bodies, which are insoluble aggregates of the target protein that need to be solubilized and refolded using additional steps. Furthermore, bacterial systems are not appropriate for expressing eukaryotic proteins that need these modifications because they lack the machinery necessary for the production of PTMs, such as glycosylation, phosphorylation, or disulfide bond formation⁶⁰.

Mammalian cell lines, such as Chinese hamster ovary (CHO) cells, are often the expression system of choice for producing proteins that require complex PTMs⁶¹. These cell lines are ideal for the production of recombinant proteins that must be biologically active in human systems, including therapeutic proteins such as monoclonal antibodies⁶². Mammalian cells also provide a more authentic cellular environment for the folding and assembly of recombinant proteins, ensuring that they maintain their functional characteristics⁵⁸. However, the major drawbacks of mammalian cell systems are their higher costs, slower growth rates, and the complexities involved in scaling up production⁵⁹.

A flexible substitute for bacterial systems is provided by yeast species, especially *Pichia pastoris* and *Saccharomyces cerevisiae*⁶³. Since yeasts are eukaryotic, they have the cellular machinery necessary to perform PTMs, which may be essential for some proteins to function⁶⁴. The ability of *Pichia pastoris* to reach extremely high densities in low-cost media makes it an especially appealing option for large-scale production. Additionally, proteins can be secreted straight into the culture medium thanks to its effective secretory pathway, which makes downstream purification easier⁶⁵. But yeast systems can also be problematic, especially when it comes to glycosylation patterns, which can be different from those in mammalian systems and change the protein's biological activity. Furthermore, the growth rate is slower than that of *E. coli*, and the yield in yeast systems might not always be as high as in bacterial systems⁵⁹.

The baculovirus expression system is frequently used to produce recombinant proteins from insect cells, especially those from the *Spodoptera frugiperda* (Sf9) cell line⁶⁶. The target gene is introduced into insect cells by the baculovirus vector, where it is highly expressed.

Compared to yeast systems, insect systems can accomplish more complex PTMs and produce larger yields. Insect cells can also handle larger, more complex proteins and are typically less expensive to cultivate than mammalian cells⁵⁹. Despite these benefits, insect cell systems typically yield less than bacterial or yeast systems, and their scalability is frequently constrained for large-scale manufacturing. Furthermore, similar to yeast systems, insect cells may produce different glycosylation patterns than mammalian cells, which limits their applicability for some therapeutic applications⁵⁸.

1.3.1.1 Vector selection for recombinant protein expression

The choice of vector family is governed by the host. Plasmid vectors are typically used for the formation of recombinant clones. A promoter, a selectable marker, an origin of replication, and a suitable cloning site for the target gene's insertion are all essential elements of an efficient expression vector⁶⁷. When selecting a vector, copy number and plasmid compatibility should be considered because they affect the organism and protein yields⁶⁸.

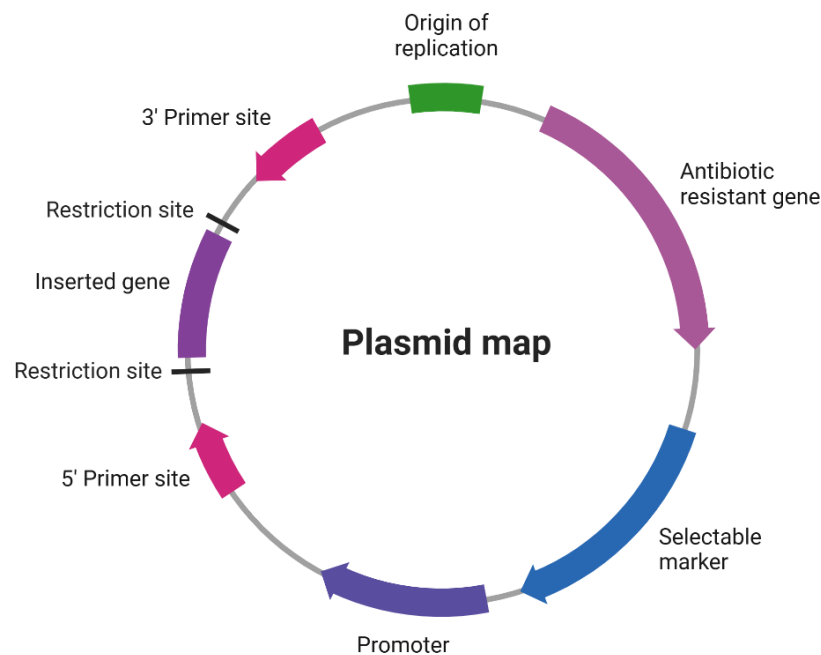


Figure 9. Plasmid map of a typical expression vector. Key elements include a promoter for driving target gene expression, a selectable marker for identifying transformed cells, an origin of replication for plasmid maintenance, and a cloning site for insertion of the target gene. *From Biorender.*

The promoter is one of the most important elements of the expression vector, as it determines the level of transcription of the target gene⁶⁰. To prevent toxicity, ideal promoters should be easily controlled with activators or inhibitors and produce recombinant products up to 30% of the cell's protein composition⁶⁹. Common promoter systems are derived from bacteriophages like T7, T5, and SP6 systems as well as bacteria like lac, tac, and trc⁷⁰.

In *E. coli*, IPTG is commonly used to induce gene expression via the lac operon system, while in yeast and mammalian systems, promoters like AOX1 (in *P. pastoris*) or CMV (in mammalian cells) allow for regulated expression⁶⁰. The ability to control gene expression is particularly important in preventing metabolic overload or toxic effects during the growth phase of the culture.

1.3.2 Tailoring recombinant protein expression in Escherichia Coli

Numerous bacterial hosts have been chosen and examined for effective protein expression. A few strains have undergone modifications to enhance the expression of recombinant proteins. Proteases like Lon or other outer membrane proteases, like OmpT, are typically faulty in these strains.⁷¹

A popular strain for examining *E. coli* protein expression is BL21(DE3). Under the influence of the lacUV5 promoter, it expresses T7 RNA polymerase. The absence of Lon and OmpT proteases in BL21 derivatives may result in leaky or basal expression. Wild-type T7 lysozyme is produced by BL21(DE3) pLysS to lower basal expression^{71,72}.

The expression of the recombinant protein needs to be carefully regulated after it has been introduced into the host organism. Target protein overexpression may overwhelm the host's cellular machinery, resulting in misfolding, inclusion body formation, or proteolytic degradation⁷³. Thus, it is necessary to determine the optimal conditions for protein expression.

A common predictive method for figuring out the production of soluble protein and determining the ideal scale for large-scale growth is small-scale test expression⁷⁴. One of the main issues is that the degree of aeration and the culture conditions affect the solubility and expression level of a recombinant protein, and these factors do not always increase with the volume of the culture. Different sample preparation techniques and protein purification

techniques are employed for each growth scale, which also affects the outcomes of small and large-scale growth^{74,75}.

Recombinant protein expression can be affected by a number of factors, including the composition of the medium, the choice of expression hosts, fusion tags, the rate of protein synthesis, the point of induction, the inducer concentration, the promoter strength, and the copy number of the expression vector⁷⁶. Optimizing the medium composition is crucial for the stability and correct folding of expressed proteins⁷⁷. Another important factor in the production of recombinant proteins is the inducer concentration in the growth medium⁷⁸.

Fusion tags can facilitate protein recovery by enhancing solubility, aiding purification and detection, and promoting proper folding. However, the choice of fusion tag should be based on the target protein, as their presence significantly influences native state interaction, cellular localization, post-translational modifications, and solubility nature⁷⁹.

1.3.2.1 Optimizing recombinant protein purification: key considerations

Recombinant proteins are produced after expression and subjected to purification to remove contaminants and impurities such as cellular proteins, nucleic acids, endotoxins, and viral particles⁵⁸. Protein purification is complex due to their diversity and properties. The procedure depends on the expression system and whether the protein is produced inside or outside the host cells. In yeast, such as *Pichia pastoris*, purification is simple, while in *Escherichia coli*, it is more complex because recombinant proteins are mainly produced inside the cells^{76,80}.

Protein extraction can be accomplished by mechanical, chemical, or enzymatic means in order to extract the material of interest that cells produce. Mechanical methods, such as heat shock, osmotic shock, and high-pressure homogenization, are effective for large-scale production, but can be challenging due to the release of cellular debris and cellular components and the force applied during the process might affect the protein by denaturing or degrading it⁸¹. Enzymatic methods, such as lysozyme, glucanase, mannase, and protease, provide an effective and selective approach to protein extraction. The determination of the enzyme depends on the cell type, its composition, and the desired protein⁸².

Purification can be achieved through a combination of techniques, such as affinity chromatography, size-exclusion chromatography, and ion exchange chromatography.

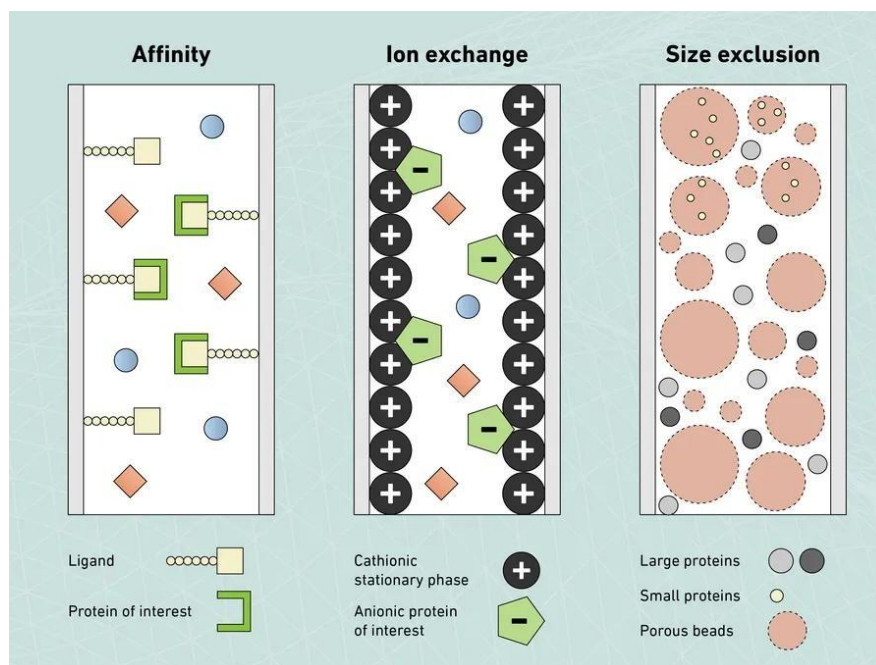


Figure 10⁸³. Schematic representation of column matrices used in common protein purification techniques. (A) Affinity chromatography: A ligand specific to the target protein is immobilized on the matrix. (B) Ion exchange chromatography: The matrix is functionalized with charged groups (e.g., positive for cation exchange, negative for anion exchange) to bind proteins based on their charge. (C) Size exclusion chromatography: The matrix consists of porous beads that separate proteins based on their size. *From the site: <http://www.technologynetworks.com/analysis/articles/an-introduction-to-protein-purification-methods-technologies-and-applications-388443>.*

The most widely used method for purifying recombinant proteins that have been tagged with affinity sequences, like Histidine (His-tags) or Glutathione S-transferase (GST-tags), is affinity chromatography. With this method, the tagged protein can be isolated from other cellular constituents by binding selectively to a ligand (such as glutathione or nickel) affixed to a column matrix⁸⁴. Even though affinity chromatography is very effective, it can occasionally lead to the co-purification of non-specific proteins. Furthermore, in order to prevent interference with the protein's activity, the tag may need to be removed after purification⁸⁵. Common proteases include enterokinase, factor Xa, SUMO protease, and 3C. Cleavage efficiency varies with each fusion protein, and improvements can be achieved by increasing protease concentration or prolonging digestion time⁸⁴.

Size-exclusion chromatography (SEC) is a method used to separate proteins based on size, particularly for removing aggregated proteins or analyzing recombinant protein oligomers. It can be used in combination with other chromatographic methods for higher purity⁸⁶. Ion exchange chromatography (IEC) separates proteins based on charge, particularly for similar-sized proteins with differing net charges. Eluting the purified protein involves adjusting pH or salt concentration in the elution buffer⁸⁷.

In summary, the purification of recombinant proteins involves a combination of techniques that leverage properties such as size, charge, and affinity to isolate the desired protein from cellular contaminants.

1.3.3 Methods for evaluating recombinant protein purity, aggregation, and structure

Prior to conducting any additional experiments, protein purity is an essential quality to assess. Culture media or extraction and purification processes can produce common contaminants in recombinant protein products. SDS-PAGE (denaturing protein gel electrophoresis) is frequently used to assess protein contamination. SDS-PAGE can determine a protein's molecular weight and approximate concentration, making it a fundamental requirement in quality control⁸⁸. It is not always the case that a single band with the anticipated molecular weight indicates that the protein has been homogeneously purified.

Mass spectrometry (MS) enhances protein analysis in several ways. Intact mass measurement confirms protein molecular weight and offers a rapid assessment of PTMs like phosphorylation, acetylation, and glycosylation. Furthermore, tandem mass spectrometry (MS/MS) identifies proteins, verifying the protein of interest and revealing co-purifying host proteins that might otherwise be mistaken for the target recombinant protein⁸⁹.

A homogeneous protein sample consists of molecules that are uniform in size and correctly folded. However, proteins can form soluble aggregates (either low or high molecular weight) through non-covalent interactions between two or more polypeptide chains. These chains may or may not maintain their native conformation within the aggregate. Because these protein complexes can exhibit diminished or absent biological activity, their detection is crucial⁹⁰.

Several techniques can detect protein aggregates, including UV-visible and fluorescence spectroscopy, and dynamic light scattering (DLS). DLS, also known as photon correlation spectroscopy, is a particularly useful method due to its sensitivity, reliability, and versatility. It offers rapid analysis, requires minimal sample preparation and amounts, can analyze a wide range of molecular weights, and can detect even trace amounts of high molecular weight aggregates⁹¹. An intensity distribution analysis with DLS is especially well-suited for aggregation studies, as larger aggregates scatter significantly more light than smaller molecules, facilitating their detection even at low concentrations. Furthermore, DLS can be used to monitor protein stability over time, across varying temperatures, in different buffers, and even in high-throughput formats⁹². Despite these advantages, DLS has limitations. It is sensitive to temperature and solvent viscosity, offers low resolution, is unreliable for precise molecular weight determination, requires transparent samples, and necessitates careful optimization of sample concentration⁸⁸.

Analyzing protein secondary and tertiary structure is crucial for quality control and verifying proper protein folding. Circular dichroism (CD) is particularly prevalent in recombinant protein production for determining secondary structures and folding properties. It is used to assess the impact of mutations, ligands, and fusion tags on protein and polypeptide stability, and routinely monitors protein unfolding and refolding as a function of temperature⁹³. Specifically, in the near-UV region (250-330 nm), CD spectroscopy can characterize the environment surrounding aromatic amino acid residues (Trp, Tyr, and Phe) and the dihedral angles of disulfide bonds⁹⁴. These aromatic groups, frequently found at binding interfaces or sites, serve as natural probes for studying ligand interactions, both qualitatively and quantitatively⁹⁵. CD offers rapid analysis, requires minimal protein, and works best with clear samples free of insoluble aggregates. This technique provides insight into the protein's folding state and can help determine whether the protein adopts the expected α -helical or β -sheet conformations, which are critical for its function⁹⁶.

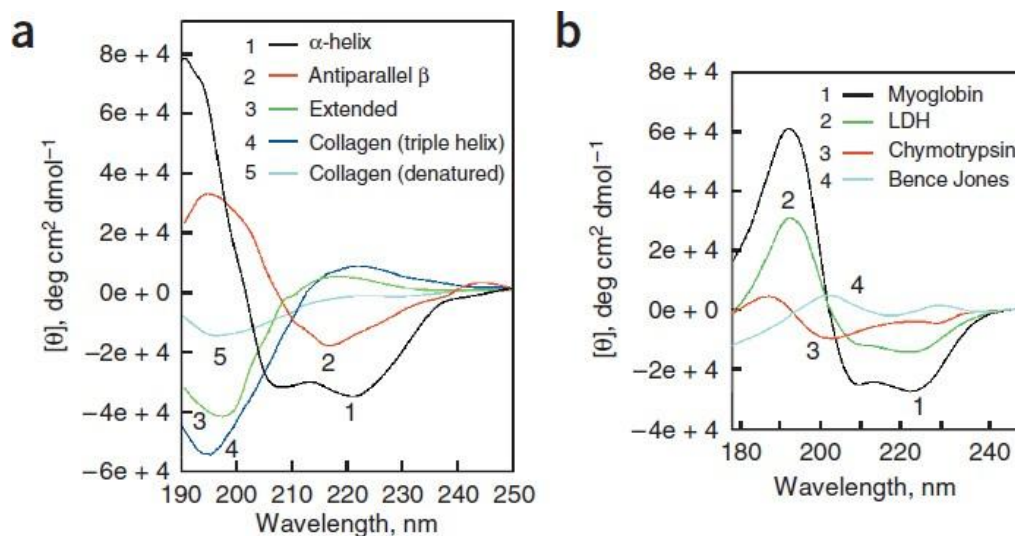


Figure 11⁹⁴. Circular dichroism spectra of polypeptides and proteins with some representative secondary structures. From: “Using circular dichroism spectra to estimate protein secondary structure” Greenfield, N.J. DOI: 10.1038/nprot.2006.202.

Determining a protein's complete three-dimensional fold directly from its sequence remains challenging. X-ray crystallography, the primary technique for determining atomic-resolution 3D structures, utilizes the short wavelength (around 0.1 nm) of X-rays. However, deciphering the 3D structure of a large molecule from the diffraction pattern of its crystal is a complex process, first achieved for a protein in 1960⁹⁷. The resulting diffraction pattern yields a complex 3D electron-density map, which is then interpreted to generate a 3D structure. The resolution of the initial crystallographic data dictates the accuracy of the final atomic model. To date, X-ray crystallography and NMR spectroscopy have revealed the structures of approximately 10,000 different proteins. Interestingly, protein folds often appear more conserved throughout evolution than the underlying amino acid sequences⁹⁸.

Protein structure and function are investigated using a variety of powerful techniques. X-ray diffraction of protein crystals is typically employed to determine the atomic-level 3D structure of larger proteins. For smaller proteins in solution, nuclear magnetic resonance (NMR) spectroscopy is often used. Given that proteins with similar structures frequently share similar functions, sequence homology searches can sometimes predict a protein's biochemical activity by identifying known proteins with comparable amino acid sequences.

2. AIM OF THE PROJECT

This project aims to develop a novel protein replacement therapy for Marinesco-Sjögren syndrome (MSS), a rare genetic disease, using the disease itself as a model system. The underlying hypothesis is that the administration of functional SIL1 protein, deficient in MSS, can alleviate the disease pathology. This has been achieved by creating DNA constructs encoding SIL1 fused to cell-penetrating peptides (CPP), allowing cellular uptake of the therapeutic protein. Since the efficacy of the delivery system depends on factors such as the CPP sequence, its position within the SIL1-CPP chimera and the linker used, the project has focused on systematically optimizing these variables to maximize cellular penetration and restore the wild-type phenotype. A robust and scalable purification protocol for the chosen SIL1-CPP variant has been developed. Rigorous quality control measures, including protein folding and aggregation assessments, ensured the integrity and homogeneity of the purified material prior to *in vitro* and *in vivo* studies. The cellular uptake and potential cytotoxicity of the SIL1-CPP fusion proteins will then be assessed using cellular models of MSS. Finally, the tissue distribution of SIL1-CPP will be assessed *in vivo*. The project is structured into four sequential phases: construct generation, protein purification and characterization, cellular uptake studies, and *in vivo* distribution assessment, each of which contributes to the overall goal of developing a viable protein replacement therapy for MSS.

3. MATERIAL AND METHODS

3.1 Generation of recombinant SIL1 constructs for protein expression

The SIL1-KDEL construct was cloned by GenScript into the pET21a+ vector, therefore the recombinant protein will have an N-terminal His tag and the replacement of the last 4 amino acids “KELR” with the sequence “KDEL” at the C-terminal.

The constructs of SIL1-TAT(YGRKKRRQRRR), SIL1-TATk (YARKAARQARA), SIL1-KDEL and SIL1 wildtype (WT), without the first 31 amino acid at N-terminus, were cloned by GenScript into the pGEX6P1 vector, therefore the recombinant proteins will have a N-terminal GST-tag.

The SIL1 regions and CPPs sequences, along with their calculated molecular weights (MW) and molar extinction coefficients (ϵ) in non-denaturing conditions, are provided in Table 1. Molecular weights and molar extinction coefficients were predicted using the ProtParam tool available on the Expasy server⁹⁹.

Vector	SIL1 regions	CPPs sequence	MW (kDa)	ϵ ($M^{-1} cm^{-1}$)
pET21a+	1-461	KDEL	52.8	33140
pGEX6P-1	31-461	-	48.8	33015
pGEX6P-1	31-461	YGRKKRRQRRR	50.4	34505
pGEX6P-1	31-461	YARKAARQARA	50.1	34505
pGEX6P-1	31-461	KDEL	48.8	33015

Table 1. SIL1 regions and CPPs sequences. Molecular weights (MW), and molar extinction coefficients (ϵ) were calculated with the ProtParam tool.

3.2 BL21(DE3) and BL21(DE3)pLysS Strains: Expression System Details

To ensure precise control over the expression of target protein, we employed the widely utilized T7 RNA polymerase system within *E. coli* cells. Specifically, we utilized two related *E. coli* strains: BL21(DE3) (Sigma-Aldrich #CMC0014 and Novagen # 70235) and BL21(DE3)pLysS (Novagen).

The BL21(DE3) strain is engineered to carry the λ DE3 lysogen, which contains the T7 RNA polymerase gene under the control of the lacUV5 promoter.

The BL21(DE3)pLysS *E. coli* strain expresses also the T7 lysozyme, which inhibits T7 RNA polymerase, minimizing basal expression of the target gene and conferring chloramphenicol resistance.

In both strains, Isopropil- β -D-1-thiogalactopyranoside (IPTG, Santacruz # SC-202185B) induction is used to initiate the T7 RNA polymerase expression cascade, ultimately leading to transcription of the target gene under T7 promoter control.

3.3 Expression and purification of recombinant SIL1-KDEL using the pET21a+ vector in *E. coli* BL21(DE3)pLysS

3.3.1 Preparation of BL21(DE3)pLysS Competent cells

An overnight culture from a single colony was diluted 1:100 in LB medium (LB Broth Miller, Sigma-Aldrich # L3522) with chloramphenicol (CHL) 34 μ g/ml, grown to OD₆₀₀ = 0.4. Cells were harvested and pelleted by centrifuging at 4000 rpm for 10 minutes in the cold. They were then resuspended in 0.1M CaCl₂ solution and incubated 30 minutes on ice. After centrifugation, the bacterial pellet was resuspended in 0.1M CaCl₂ solution with 15% glycerol, divided into aliquots, flash frozen in liquid nitrogen and stored at -80 °C.

3.3.2 Heat-shock transformation of BL21(DE3)pLysS bacterial cells

Aliquots of cells were thawed on ice, 50 ng of plasmid DNA was added, and then incubated on ice for 30 min. Cells were heat shocked at 42°C for 45 s, chilled on ice for 2 min, 900 μ L

of LB was added, and cells were allowed to recover and express antibiotic resistance for 1 h at 37°C. After cell recovery in LB, cells were plated onto LB Agar plates containing the appropriate antibiotics for selection CHL 34 µg/ml and ampicillin (AMP) 50 µg/ml. Plates were then incubated to allow cell growth at 37°C overnight. Single colonies from the transformation plate were grown in LB with the respective antibiotics overnight to create a glycerol stock of single-colony derived cells.

3.3.3 Optimal induction parameters for SIL1-KDEL in BL21(DE3)pLysS

For overnight cultures, 1–5 ml of LB medium supplemented with 50 µg/ml AMP and 34 µg/ml CHL were inoculated from the glycerol stock of cells derived from a single colony and incubated overnight at 37 °C, shaking at 220 rpm. The overnight culture was diluted 1:200 in 40–100 ml to 1 L of fresh LB medium containing antibiotics and was further incubated at 37 °C. Protein expression was induced by the addition of IPTG when OD600 reached 0.4–0.6. Different concentrations of IPTG were tested in combination with different induction times and temperatures in order to find the optimal conditions for protein expression outlined in the table below.

IPTG concentration	Temperature of induction	Time of induction
0.2 mM	Room Temperature	16-18 hours
0.5 mM	Room Temperature	16-18 hours
0.2 mM	4°C	72 hours
0.5 mM	4°C	72 hours
0.5mM	37°C	16-18 hours
0.5mM	16°C	16-18 hours
1mM	16°C	16-18 hours
0.5mM	37°C	1 hour
0.5mM	37°C	2 hours
0.5mM	37°C	3 hours

Table 2. Exploring IPTG Induction. Overview of different IPTG induction concentrations, times and temperatures tested

Bacteria were harvested by centrifugation at 4 °C, 4000 rpm for 20 min. Pellets were stored at -80 °C until further use.

3.3.4 Cell lysis and differential recovery of soluble and insoluble SIL1-KDEL recombinant protein

The bacterial pellet from the expression culture was thawed on ice and resuspended in sonication buffer (50 mM Na₂HPO₄·12H₂O pH7.4, 500 mM NaCl, 10 mM Imidazole, 10% glycerol, protease inhibitor, and 1 mg/ml lysozyme). After incubation for 30 min on ice, the lysate was sonicated at 37% duty cycle with three 10 s bursts with 10 s intervals in a probe sonicator. Soluble proteins were separated from cell debris by centrifugation at 24,000 g for 20–25 min at 4 °C.

For insoluble proteins, the pellet was lysed in inclusion body lysis buffer (50 mM Na₂HPO₄·12H₂O pH7.4, 500 mM NaCl, 10 mM Imidazole, 10% Glycerol, 0.5% SDS and 6M Urea). Sonication and subsequent centrifugation conditions were the same as reported above. The lysate and the insoluble portion were analyzed by the Bradford protein assay (Bio-Rad) and stored at -20 °C until use.

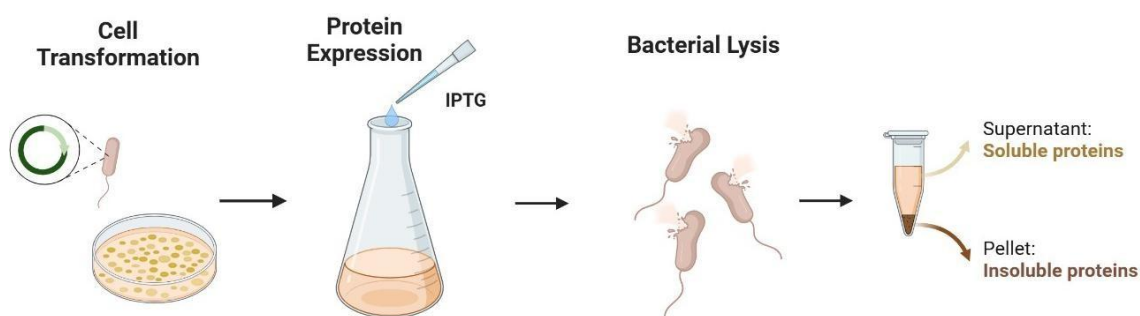


Figure 12. Schematic overview of recombinant protein expression steps. The main steps include: (1) Transformation: Introduction of the expression vector containing the gene of interest into bacterial host cells. (2) Induction: Stimulation of protein expression using an appropriate inducer (e.g., IPTG). (3) Bacterial Lysis: Disruption of bacterial cells to release the expressed recombinant protein. *From Biorender.*

3.3.5 Purification of SIL1-KDEL by AKTA Prime Plus Chromatography system

The bacterial pellet corresponding to 1L of culture was lysed in 50 mM Hepes pH7.4, 400 mM NaCl, 10 mM Imidazole, 10% glycerol, protease inhibitor and 1 mg/ml lysozyme for the soluble part and in 50 mM Hepes pH7.4, 400 mM NaCl, 10% Glycerol, 0.5% SDS and 6M Urea for the extraction from the inclusion bodies following the instructions described in the previous paragraph. The purification steps by affinity chromatography were performed

at room temperature following the manufacturer's instructions¹⁰⁰, using Hepes to exclude the presence of phosphates. All samples and buffers were passed through a 0.22 µm or 0.45 µm filter. Protein elution was followed by monitoring the absorbance at 280 nm. Using AKTA Prime Plus (Cytiva), the 1 mL His Trap HP His tag protein purification column (Cytiva # 29051021) was equilibrated with 5 column volumes of binding buffer (50 mM Hepes pH7.4, 400 mM NaCl, 10 mM Imidazole, 10% glycerol). This chromatographic technique was terminated by immobilized metal ion affinity chromatography (IMAC). After sample loading, in the purification procedure performed under native conditions, elution of histidine-tagged proteins was achieved with a linear imidazole gradient (10mM-1M). In the denaturation protocol, 6 M urea was added to the binding and elution buffers.

3.4 Expression and purification of recombinant SIL1-CPPs using the pGEX6P-1 Vector in E. coli BL21(DE3)

3.4.1 Heat-shock transformation of BL21(DE3) bacterial cells

Aliquots of chemically competent cells (Sigma-Aldrich and Novagen) were thawed on ice, 100 ng of plasmid DNA was added, and then incubated on ice for 30 min. Cells were heat shocked at 42°C for 45 s, chilled on ice for 2 min, 900 µL of LB was added, and cells were allowed to recover and express antibiotic resistance for 1.30 h at 37°C. After cell recovery in LB, cells were plated onto LB Agar plates containing the appropriate antibiotics for selection AMP 50 µg/ml. Plates were then incubated to allow cell growth at 37°C overnight. Single colonies from the transformation plate were grown in LB with the respective antibiotics overnight to create a glycerol stock of single-colony derived cells.

3.4.2 Expression of SIL1-CPPs using pGEX6P-1 in BL21(DE3)

For overnight cultures, 3 ml of LB medium supplemented with 50 µg/ml ampicillin were inoculated from the glycerol stock of cells derived from a single colony and incubated overnight at 37 °C, shaking at 220 rpm. The overnight culture was diluted 1:200 in 20 ml of LB medium containing antibiotics and was further incubated at 37 °C. Protein expression was induced by the addition of IPTG for 1 hour at 37°C when OD600 reached 0.4–0.6 for

the initial expression screening. The bacteria were collected by centrifugation at 4 °C, 4000 rpm for 20 min. The pellets were stored at -80 °C until further use.

3.4.3 Soluble and insoluble recombinant SIL1-CPPs recovery

The bacterial pellet from the expression culture was thawed on ice and resuspended in sonication buffer for soluble proteins (50 mM Hepes pH7.4, 400 mM NaCl, 10% glycerol, protease inhibitor and 1 mg/ml lysozyme). After incubation for 30 min on ice, the lysate was sonicated at 60% duty cycle for 30 min with intervals of 10 seconds on and 10 seconds off in a probe sonicator. Soluble proteins were separated from cell debris by centrifugation at 24,000 g for 20–25 min at 4 °C. For insoluble proteins, the pellet was lysed in inclusion body lysis buffer (50 mM Hepes pH7.4, 400 mM NaCl, 10% Glycerol, 0.5% SDS and 6M Urea). Sonication conditions were the same as above. Insoluble proteins were collected by centrifugation at 24,000 g for 30 min at 4 °C. The soluble lysate and the insoluble portion were analyzed by the Bradford spectrophotometric method and stored at -20 °C until use.

3.4.4 Optimization of purification of GST-tagged SIL1-TAT protein by affinity chromatography

Once the SIL1-TAT protein was chosen based on preliminary data, chimera expression for purifications was induced with 0.5 mM IPTG for 1 h at 37 °C when OD600 reached 0.4-0.6 and for 3 h at 25°C when OD600 reached 0.5-0.6.

The purification protocol for recombinant SIL1-CPP was optimized selecting the most efficient binding buffer, by evaluating the optimal resin-to-protein ratio and identifying strategies for removing bacterial heat shock proteins.

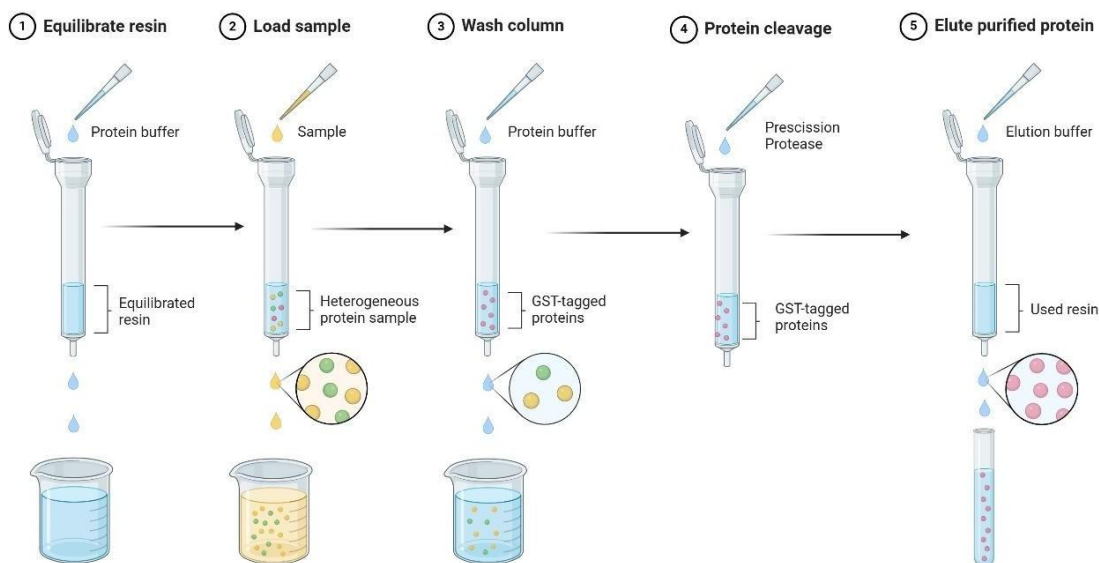


Figure 13. Schematic representation of the affinity chromatography purification protocol. The key steps include: (1) Conditioning the resin with buffer to establish appropriate pH and ionic strength. (2) Applying the sample containing the target protein to the column, allowing it to bind to the affinity ligand. (3) Removing unbound or weakly bound contaminants with wash buffer. (4) On-column cleavage of a tag or fusion protein, if necessary for releasing the target protein from the affinity ligand. (5) Recovering the purified target protein by disrupting its interaction with the ligand using an elution buffer (e.g., containing a competing ligand or altered pH). *Created with Biorender.*

Different buffers were tested for the extraction of GST-tagged proteins:

- 50 mM Hepes pH7.4, 400 mM NaCl, 10% glycerol, protease inhibitor and 1 mg/ml lysozyme;
- PBS without Mg e Ca, protease inhibitor and 1 mg/ml lysozyme;
- 50 mM Hepes pH7.4, 400 mM NaCl, 10% glycerol, 5-10 mM MgCl₂, protease inhibitor and DNase I.

Optimization of protein purification was performed using batch and gravity flow chromatography. The bacterial lysate was loaded onto Glutathione Sepharose 4B agarose resin (Cytiva #GE17-0756-01) equilibrated in the corresponding buffer. Affinity chromatography purification steps were performed at 4 °C and binding was performed for 2 h or overnight under rotation.

The initial resin volume was determined according to manufacturer instructions¹⁰¹, resulting in inefficient binding of the recombinant protein. To optimize the resin-protein ratio, the unbound fraction from the initial binding step was re-incubated with an equal volume of fresh resin.

An additional wash step, consisting of 5 mM ATP and 15 mM MgCl₂ in binding buffer, was implemented to remove potential GroEL contaminants¹⁰². This wash was evaluated at both 4°C and room temperature to optimize contaminant removal.

GST tag removal required cleavage with Prescission Protease (Cytiva #GE27-0843-01). The resin was first equilibrated with the recommended protease working buffer: 50 mM Tris-HCl pH 7.5, 150 mM NaCl, 0.1 mM EDTA, and 0.1 mM DTT. Following equilibration, the resin was incubated overnight with Prescission Protease. The amount of protease used was determined based on the manufacturer's suggested protocol. Elution of the recombinant proteins was performed using one of two buffers: 50 mM Tris-HCl pH 7.5, 150 mM NaCl or 50 mM Tris-HCl pH 7.5, 400 mM NaCl.

3.4.5 Purification of GST-tagged SIL1-TAT and SIL1-WT by Fast Protein Liquid Chromatography

Purification of GST-tagged proteins was performed using pre-packed GSTrap FF columns (Cytiva # GE17-5131-02). Both 1 mL and 5 mL column sizes were used. The column matrix consists of glutathione coupled via a 10-carbon linker to 4% highly cross-linked agarose. Chromatography was carried out on AKTA start and AKTA go systems (Cytiva).

The column was equilibrated with 15 column volumes of ultrapure water (ddH₂O) followed by 10 column volumes of binding buffer (50 mM Hepes pH7.4, 400 mM NaCl). To maximize binding efficiency, the lysate was loaded in two runs at a flow rate of 500 µL to 1 mL per minute. Initially, the lysate was circulated twice through the column. Subsequently, a recirculation loop was established and maintained overnight.

The following day, unbound proteins were removed by washing with 10 column volumes of binding buffer. GroEL was then removed using 3 column volumes of 5 mM ATP and 15 mM MgCl₂ in binding buffer.

GST tag cleavage was achieved using Prescission Protease. Prior to protease addition, the column was equilibrated with 50 mM Tris-HCl pH 7.5, 150 mM NaCl, 0.1 mM EDTA, and 0.1 mM DTT. The resin was subsequently incubated overnight with Prescission Protease at the manufacturer-recommended concentration. Recombinant proteins were eluted using either 50 mM Tris-HCl pH 7.5, 150 mM NaCl or 50 mM Tris-HCl pH 7.5, 400 mM NaCl.

HRV 3C protease (Sigma-Aldrich #SAE0045) was also evaluated for GST tag cleavage. Following removal of GroEL, the resin was incubated directly overnight with HRV 3C protease at the manufacturer-recommended concentration. Recombinant proteins were subsequently eluted with 50 mM NaCl, 400 mM Hepes pH 7.4.

3.5 SDS-PAGE electrophoresis

Protein concentration was assessed using either the Bradford protein assay (Bio-Rad)¹⁰³ or the Pierce™ BCA Protein Assay Kit (Thermo Fisher #23225)¹⁰⁴, according to the manufacturers' instructions. Equal amounts of protein were then resuspended in 4x Laemmli sample buffer (15% β-mercaptoethanol, 0.25 M Tris-HCl pH 6.8, 40% glycerol, 8% SDS, and 0.02% bromophenol blue), boiled for 10 minutes at 100°C, and separated by SDS-PAGE on 8-10% polyacrylamide gels. Electrophoresis was performed at room temperature with an applied voltage of 100-150 V. Low molecular weight SDS markers (PageRuler Plus Prestained Protein Ladder, Thermo Fisher # 26619), were included as molecular weight standards. Following electrophoresis, gels were fixed in a solution of 50:10:40 ethanol: acetic acid: water, stained with a Coomassie brilliant blue solution (0.058 g in 25:8:67 50% ethanol: acetic acid: water), and destained using 25:8:67 50% ethanol: acetic acid: water. In a separate experiment, gels were stained overnight with ReadyBlue Protein Gel Stain (Sigma-Aldrich).

3.6 Protein detection by Western Blotting

After separation by SDS-PAGE described in paragraph 3.5, proteins were transferred to nitrocellulose membrane. The presence of proteins on nitrocellulose membranes and their quantity were assessed by Ponceau S staining (0.5% w/v Ponceau S dissolved in 1% v/v acetic acid). The membranes were incubated for 5 minutes in the Ponceau S solution and then washed once with water, photographed and destained with TBST. The filters were then blocked for 1 hour at room temperature with 5% skimmed milk powder and incubated with the primary antibodies overnight at 4 °C. The secondary antibodies α-rabbit and α-mouse were applied the next day for 1 hour at room temperature and the signal was detected by GE Amersham ECL. The following antibodies were used: SIL1 (Sigma Prestige, #HPA011949),

GAPDH (Santa Cruz #sc-32233), DnaK (Novus Biotechnique, #NBP1-97490) and His Tag (Biolegend, #362602).

3.7 Mass spectrometry analysis of SIL1-TAT recombinant protein

SIL1-TAT protein, eluted after GST tag cleavage, was separated by SDS-PAGE. Excised protein bands from Coomassie-stained gels were subjected to in-gel tryptic digestion using a modified protocol. Briefly, gel pieces were destained (MilliQ water, then 50 mM NH_4HCO_3 /50% ethanol), dehydrated (100% ethanol), and dried. Reduction (10 mM Dithiothreitol in 50 mM NH_4HCO_3 , 56°C, 1 h) and alkylation (55 mM Iodoacetamide, room temperature, 30 min, dark) were performed. Gel pieces were washed (alternating 50 mM NH_4HCO_3 and 100% ethanol), dehydrated (100% ethanol), and dried. Tryptic digestion was carried out overnight at 37°C (1:50 enzyme-to-protein ratio), and the reaction was stopped with formic acid. Peptides were extracted with sequential 50% Acetonitrile and MilliQ water washes, concentrated in SpeedVac, and prepared for nanoLC-ORBITRAP-MS/MS analysis.

Raw data from nanoLC-ORBITRAP-MS/MS were processed using Proteome Discoverer software (version 2.5, Thermo Fisher Scientific) with the SEQUEST algorithm. Database searching was performed against the *Homo sapiens* and *Escherichia coli* UniProtKB/Swiss-Prot databases.

3.8 Characterization of protein stability and folding using CD spectroscopy

The purified SIL1-TAT and SIL1-WT proteins were then concentrated with Amicon devices when necessary (MWCO 10 kDa, Millipore # UFC9010) and the protein concentration was estimated by measuring the absorbance at 280 nm.

After dialysis against a panel of buffers (table 3), the recombinant proteins were analyzed using CD spectroscopy on a Chirascan system (Applied Photophysics, SFR BioSciences). The buffers were chosen to identify optimal conditions for spectroscopic analysis and to investigate the influence of salt composition on chimera stability.

Buffer for CD analysis
1. Hepes 50 mM NaCl 400 mM pH 7.4;
2. Hepes 25 mM NaCl 200 mM pH 7.4;
3. KPi 50 mM 200 mM NaCl 5% v/v glycerol pH 7.4;
4. KPi 25 mM 100 mM NaCl pH 7.4 2.5% v/v glycerol;
5. KPi 25 mM 400 mM NaCl pH 7.4 2.5% v/v glycerol;
6. KPi 25 mM 400 mM NaCl 20% trifluoroethanol pH 7.4;
7. Tris acetate 20 mM NaF 200 mM pH 7.4.

Table 3. Buffers for CD Spectra Analysis. List of different buffers tested for CD spectra analysis

CD spectra were recorded at 15 °C in 1 mm thick quartz cuvette in the range 180 nm to 350 nm with a step size of 0.5 nm, a bandwidth of 1 nm. Four spectra were collected and averaged for each sample, and background and buffer signals were subtracted.

All spectra are reported in terms of mean residue ellipticity (deg cm² dmol⁻¹). Data processing was performed using GraphPad Prism 9.

3.9 Evaluation of recombinant SIL1 protein aggregation and molecular size distribution by dynamic light scattering

The hydrodynamic radius (Rh) of SIL1-TAT and SIL1-WT recombinant proteins was determined by DLS using a Zetasizer Pro-Red instrument (Malvern Panalytical). Measurements were performed in triplicate using 100 µL of each protein sample in the indicated screening buffers (see table above) at a temperature of 15°C. Disposable plastic cuvettes were used for all measurements. The Stokes-Einstein equation was used to calculate Rh from the measured diffusion rates. Data analysis was performed with Zetasizer software (ZS XPLOER 3.2.0.84). Data processing was performed using GraphPad Prism 9.

3.10 X-ray Crystallography of SIL1-TAT protein

The purified protein was dialyzed in 20 mM Tris acetate 400 mM NaCl pH 7.4 and concentrated with Amicon devices (MWCO 10 kDa). The concentration was estimated by measuring the absorbance at 280 nm. Given the complexity of the crystallization process driven by numerous variables (pH, salt, precipitant, temperature and protein concentration), the concentrated protein was subjected to screening tests. These tests systematically explore different conditions, allowing to identify the specific parameters that induce crystal formation. In particular, the following screening tests were used:

- JBScreen JCSG++ HTS (Jena bioscience, #CS-206L)
- Stura FootPrint Combination HT-96 (Molecular dimensions, #MD1-43)
- JBScreen Basic HTS (Jena bioscience, #CS-203L)
- Clear strategy screen II HT96 (Molecular dimensions, #MD1-32-ECO)

Long needle-shaped crystals were obtained by the sitting drop vapor diffusion method at room temperature (20 °C). The well solution consisted of 80 ul of 20% w/v polyethylene glycol monomethyl ether 2000, 100 mM TRIS pH 8.5, 200 mM Trimethylamine N-oxide. The crystals were flash frozen at 100 K in a flow of N₂ gas in PEG-200 cryoprotectant. The crystals diffracted X-rays at 3 Å at the European Synchrotron Radiation Facility beamline.

3.11 Determination of Endotoxin Levels in Purified SIL1-TAT and SIL1-WT

Endotoxin quantification was performed using the Thermo Scientific™ Pierce™ Chromogenic Endotoxin Quant kit (# A39552) , following the manufacturer's protocol¹⁰⁵. This assay employs lysates derived from horseshoe crab blood amoebocytes to quantitatively measure endotoxin levels in protein, peptide, antibody, or nucleic acid samples. The protein was resuspended in 50 mM Tris-HCl, 150 mM NaCl, pH 7.5. A standard curve, ranging from 0.01 to 0.1 EU/mL, was used to calculate endotoxin concentrations based on the intensity of the developed colorimetric signal.

3.12 Cells culture

Primary Human Dermal Fibroblast (HF, NDHF Promo Cell #FB60C12350) supplied by Carlo Erba were used as the control cell line and compared with primary dermal fibroblast from a young Marinesco–Sjögren syndrome patient, supplied by Telethon Network of Genetic Biobanks—TNGB (HF-MSS). Cells were cultured in Dulbecco's modified Eagle's medium + GlutaMAX (GIBCO, 61965-026), supplemented with 10% of Fetal Bovine Serum (FBS) and 1% penicillin/streptomycin PenStrep (GIBCO, 15070-063). Cells were maintained at 37 °C, 5% CO₂, and detached by Trypsin-EDTA 0.5% (GIBCO, 15400-054).

3.13 Analysis of extracellular matrix remodelling in Sil1-deficient fibroblasts

Transcriptomic analysis was performed using RNA-Seq on total RNA extracted from control and patient fibroblasts. Data were subsequently analyzed using the String database and the Ingenuity Pathway Analysis platform and validated using real-time quantitative PCR (q-PCR). Extracellular matrix (ECM) remodelling was visualized using confocal fluorescence microscopy. ECM morphology was examined both *in vitro* and in mouse tendon tissue using transmission electron microscopy (TEM). Details of the procedures used are explained in the Materials and Methods section of the article “Sil1-deficient fibroblasts generate an aberrant extracellular matrix leading to tendon disorganization in Marinesco-Sjögren syndrome¹⁰⁶” Amodei et al. DOI: <https://doi.org/10.1186/s12967-024-05582-0>

3.14 Evaluation of *in vitro* cytotoxicity of SIL1-TAT

For cytotoxicity assay, HF and HF-MSS were plated in 96-well plates at a density of 5×10^3 /well under standard growth conditions. After 24 h, cells were treated overnight with increasing doses of SIL1-TAT in 50 mM Tris-HCl 400 mM NaCl, pH 7.5. At the end of the treatment period, cells were incubated with 3-(4,5-dimethylthiazol-2-yl)-2,5-diphenyltetrazolium bromide (MTT) solution (serum-free medium with 0.5 mg/mL MTT) for 3 h.

After removal of the MTT solution, dimethyl sulfoxide (DMSO) was added to each well and then the absorbance value at 570 nm was measured using a spectrophotometer (Synergy). All experiments were performed in duplicate and as independent experiments.

3.15 Immunofluorescence to evaluate SIL1-TAT cellular uptake

HF-MSS cells were plated on coverslips at a density of 6x10⁵ cells/well. After 24 hours, cells were treated for 4 hours with increasing doses of SIL1-TAT in 50 mM Tris-HCl 400 mM NaCl, pH 7.5 and subsequently fixed in cold methanol for 5 minutes at -20°C.

Cells were then incubated for nonspecific site blocking and permeabilization in blocking solution [0.05% saponin, 0.5% bovine serum albumin and 50 mM NH₄Cl in PBS] for 1 hour at room temperature. Incubation of the primary antibody SIL1 (Sigma Prestige, #HPA011949) and HSP47 (SantaCruz, #sc-5293), was performed overnight at +4 °C. The next day, the secondary fluorescent anti-IgG Alexa Fluor antibody was incubated for 1 h at room temperature together with DAPI to counterstain the nuclei. The signal was then detected with a confocal microscope (Zeiss LSM 800) and the images were post-processed and analyzed by Zen 2.3 software (Zeiss, Jena, Germany).

3.16 Mouse model of Marinesco-Sjogren syndrome

Sil1 homozygous (woozy) mice (Sil1wz) were generated by crossing heterozygous Sil1 mice (Sil1ht) (CXB5/By-Sil1wz/J, JAX stock #003777). Genotyping was performed via PCR, following Jackson Laboratories Protocol #22452, using DNA extracted from ear punch samples. Agarose gel electrophoresis was used to determine genotypes: *Sil1wz* (758 bp), *Sil1ht* (626 bp and 758 bp), and wild type (626 bp). All procedures involving the mice were performed according to the protocol approved by the Institutional Animal Care and Use Committee of the Italian Ministry of Health (authorization n° 596/2021-PR, issued pursuant to art. 31 of Legislative Decree 26/2014).

3.16.1 Rotarod Acceleration Test

Motor function and neuromuscular coordination were assessed using a Rotarod 7650 acceleration model (Ugo Basile, Biological Research Apparatus, Varese, Italy) using, twelve *Sil1wz* and twelve *Sil1ht* mice. Prior to testing, mice underwent three 60-second training sessions at a constant speed of 4 rpm. For the acceleration test, mice were placed on the rotating bar (7 rpm initial speed) and allowed a 30-second acclimatization period. The bar then accelerated from 7 to 40 rpm at a constant rate of 0.11 rpm/s (maximum duration 300 seconds). Latency to fall was recorded as the primary endpoint. Three trials were performed with 30-minute inter-trial rest periods, and the mean latency to fall was used for analysis. Testing was conducted weekly from weeks 5 to 14. Data are presented as a percentage of each individual's maximum time on the apparatus, rather than as raw latency times.

3.16.2 Beam Walking Test

For the test, twelve *Sil1wz* and twelve *Sil1ht* mice were used. Mice were trained for two days prior to testing on a custom-made 0.8 cm wide, 80 cm long round PVC beam suspended 40 cm above bedding. On the test day, mice performed three trials with at least 5 minutes of rest between trials. Trials were video-recorded, and the following parameters were scored: time to traverse the beam, number of hindfoot missteps (defined as hind limb slippage along the beam), and falls. A maximum time of 60 seconds was assigned if the animal fell or exceeded this time to traverse the beam. The mean time to traverse and the mean number of missteps across the three trials were used for statistical analysis. Testing was conducted weekly from weeks 5 to 12.

3.16.3 *In vivo* SIL1-TAT administration and tissue collection

A 31-week-old *Sil1wz* mice received a single intraperitoneal injection of SIL1-TAT at a dose of 3 µg/g body weight. The mouse was monitored for signs of distress both during and after the injection. Six hours post-administration, tissue samples (cerebellum, spleen, and quadriceps) were collected from the treated mouse and from an age-matched untreated control mouse for subsequent western blot analysis to verify their tissue distribution.

4. RESULTS

The results of this thesis were obtained thanks to the collaboration with the laboratories of Biotech “Fibrosys” and the Institut des Sciences Analytiques of the University Claude Bernard Lyon 1.

4.1 Selection of TAT and KDEL sequences for SIL1 delivery

Prior to the initiation of this study, preliminary experiments within our laboratory evaluated several CPPs sequences for their potential to facilitate SIL1 protein delivery. These initial investigations involved the design of eukaryotic expression vectors encoding SIL1 fused to various CPPs at either the N- or C-terminus. Transient transfection of these constructs into eukaryotic cells resulted in variable levels of SIL1 protein expression. The CPP TAT and the ER retention signal KDEL were selected for further investigation. Due to the expected large-scale protein requirements and considering the preliminary expression data obtained from the eukaryotic system, a prokaryotic expression system was chosen to improve protein yield. Consequently, DNA constructs were designed for the expression of the SIL1-CPP fusion protein in *E. coli*.

4.2 Impact of induction parameters on SIL1-KDEL production in BL21(DE3)pLysS

To optimize the expression of recombinant SIL1-KDEL protein in *E. coli* BL21(DE3)pLysS cells, we studied the effects of different induction parameters, including IPTG concentration, temperature, and induction time. *E. coli* BL21(DE3)pLysS cells were selected as the expression host for their intrinsic ability to minimize basal transcription, thus providing tight control over gene expression. Chemically competent BL21(DE3)pLysS cells were transformed with the pET21a+ expression vector hosting the SIL1-KDEL gene. This construct facilitated the production of SIL1-KDEL protein with an N-terminal hexahistidine (His) tag, an established affinity domain for metal affinity purification. Total protein lysates were analyzed by Western blotting using an anti-SIL1 antibody to assess SIL1-KDEL protein levels (Figure 14).

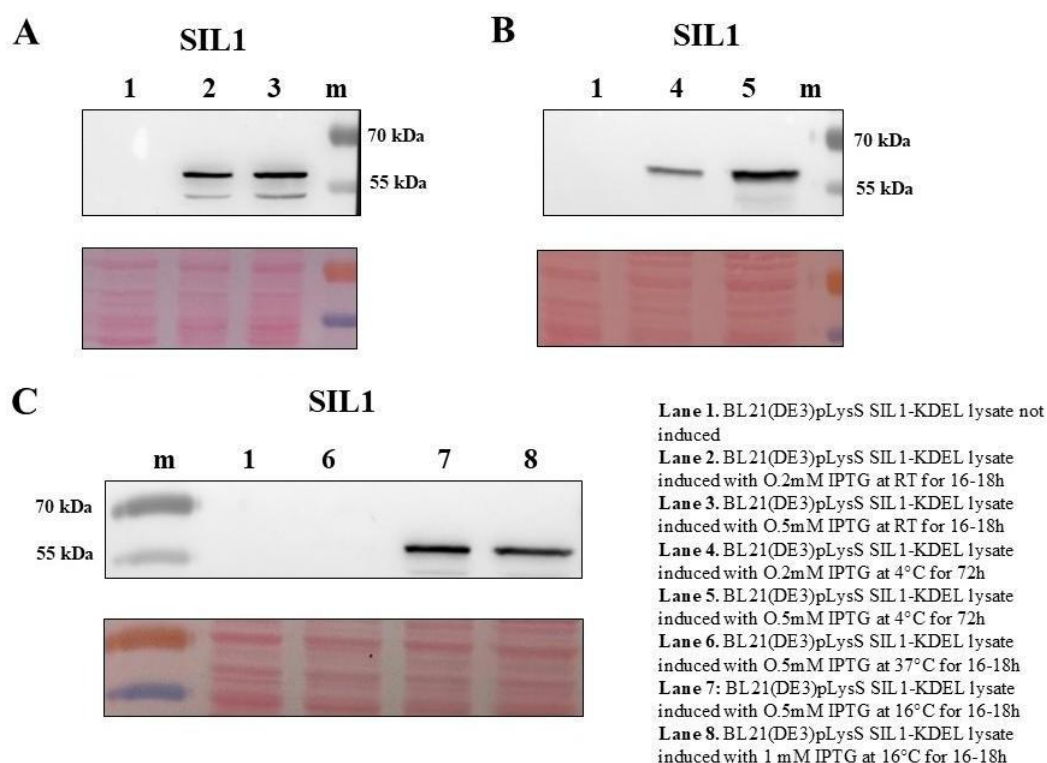


Figure 14. Analysis of SIL1-KDEL expression in BL21(DE3)pLysS. Western blot and relative ponceau staining of SIL1-KDEL expression after (A) overnight induction at room temperature with 0.2 mM and 0.5 mM IPTG; (B) 72-hour induction at 4°C with 0.2 mM and 0.5 mM IPTG; and (C) overnight induction at 37°C and 16°C with 0.2 mM, 0.5 mM IPTG and 1mM concentrations.

Specifically, the effect of IPTG concentration at room temperature (RT) was tested under an induction time of 16–18 h (Figure 14, panel A). Comparison of lysates induced with 0.2 mM (Figure 14, panel A, lane 2) and 0.5 mM IPTG (Figure 14, panel A, lane 3) revealed a clear dose-dependent increase in SIL1-KDEL protein levels. Indeed, a significant increase in expression was observed with 0.5 mM IPTG, suggesting that this is a more effective concentration to induce SIL1-KDEL transcription under these conditions. Given the benefits of low-temperature induction to minimize protein degradation and allow proper folding, we compared SIL1-KDEL expression at 4 °C for an extended period of 72 h (panel B, Figure 14). Comparison of 0.2 mM IPTG-induced lysates (Figure 14, panel B, lane 4) and 0.5 mM IPTG-induced lysates (Figure 14, panel B, lane 5) at 4°C with the corresponding RT inductions (Figure 14, panel A) showed a significant reduction in SIL1-KDEL protein levels. This observation indicates that prolonged low-temperature induction is not favourable for SIL1-KDEL expression, perhaps due to reduced translational efficiency or other temperature-dependent reasons. To clarify more the temperature effect on SIL1-KDEL expression, we induced at 37°C and 16°C using 0.5 mM IPTG for 16–18 h (panel C, figure

14). Induction at 37°C (Figure 14, panel C, lane 6) was found to reduce SIL1-KDEL expression compared to RT control (Figure 14, panel A, lane 3), suggesting that higher temperatures combined with long time adversely influence protein stability or translational efficiency. In contrast, induction at 16°C (Figure 14, panel C, lane 7) resulted in SIL1-KDEL expression that was detectable but lower than in RT control. In an attempt to clarify whether a concentration increase of IPTG would rescue SIL1-KDEL expression at 16°C, we induced with 1 mM IPTG (Figure 14, panel C, lane 8). Elevated concentration did not significantly increase levels of SIL1-KDEL more than 0.5 mM induction at 16°C (Figure 14, panel C, lane 7), indicating that IPTG concentration alone is not sufficient to overcome the inhibitory effect of low temperature on expression.

To identify the impact of induction conditions on SIL1-KDEL solubility, experiments were performed using a concentration of 0.5 mM IPTG. Total protein lysates and inclusion body fractions were analyzed by Western blotting with an anti-SIL1 antibody (Figure 15). Inclusion bodies are insoluble, dense aggregates of misfolded or partially folded proteins that are prone to occur in bacterial cells, especially *E. coli*, during overexpression of recombinant proteins.

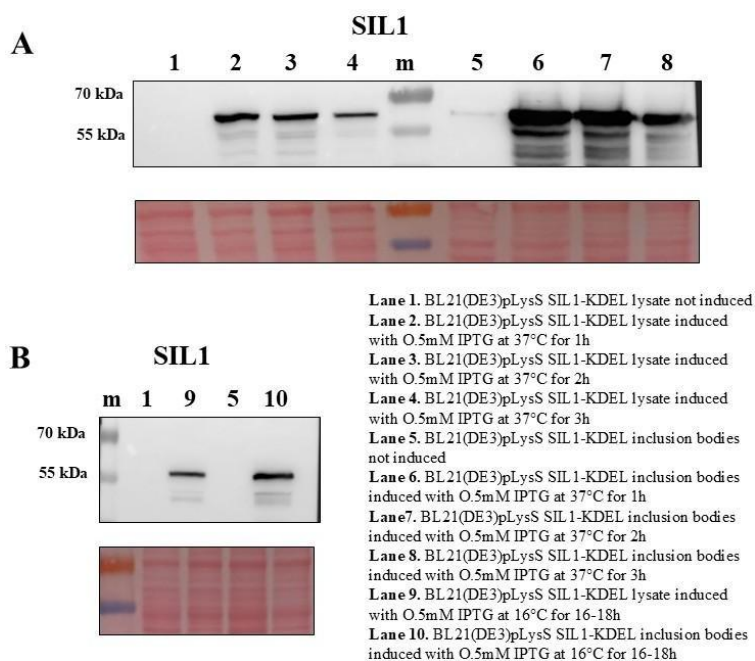


Figure 15. Analysis of SIL1-KDEL expression in BL21(DE3)pLysS. Western blot and relative ponceau staining of SIL1-KDEL expression and inclusion body formation were assessed using 0.5 mM IPTG under different conditions. (A) Acute induction at 37°C for up to 3 hours. (B) Overnight induction at 16°C.

The time profile of SIL1-KDEL expression at 37 °C is shown in panel A of Figure 15. Lysates induced with 0.5 mM IPTG for 1 h (Figure 15, panel A, lane 2), 2 h (Figure 15, panel A, lane 3), and 3 h (Figure 15, panel A, lane 4) are compared, and a clear time-dependent decrease in SIL1-KDEL protein levels is observed, once again demonstrating that increased induction times at 37 °C reduce protein expression. The respective inclusion body yields induced for 1 h (Figure 15, panel A, lane 6), 2 h (Figure 15, panel A, lane 7), and 3 h (Figure 15, panel A, lane 8) indicate that the majority of protein synthesized is in the form of precipitated inclusion bodies at 37 °C. Once more, SIL1-KDEL protein is compromised with longer induction time. This means that the observed reduction in protein content is not due to reduced synthesis alone but is also accompanied by degradation or other loss processes in soluble and aggregate proteins. SIL1-KDEL amounts in the inclusion body and lysate fractions after 16-18 h of induction at 16 °C are shown in panel B of Figure 15. Although detectable SIL1-KDEL expression is observed for the lysate fraction (Figure 15, panel B, lane 9), it is very low compared with inductions at 37 °C (Figure 15, panel A). In the respective inclusion body fraction (lane 10), SIL1-KDEL expression is relatively higher than in the lysate. These results together suggest that while induction at 37°C leads to overall increased expression of SIL1-KDEL in the soluble fraction, it also induces massive inclusion body formation.

To investigate the integrity of N-terminal His-tag of SIL1-KDEL and to better understand how induction conditions affect protein expression, Western blot analysis was conducted using an anti-His tag antibody in the lysates and inclusion body fractions from *E. coli* BL21(DE3)pLysS cells (Figure 16). The study also aimed to determine whether the observed differences in SIL1-KDEL expression were the result of loss of the His tag.

Consistent with previous observations, the 0.5 mM IPTG induced lysate for 72 h at 4°C (Figure 16, lane 2) gave a weak signal in support of the fact that the expression of SIL1-KDEL is extremely low under the condition of increased induction time and reduced temperature.

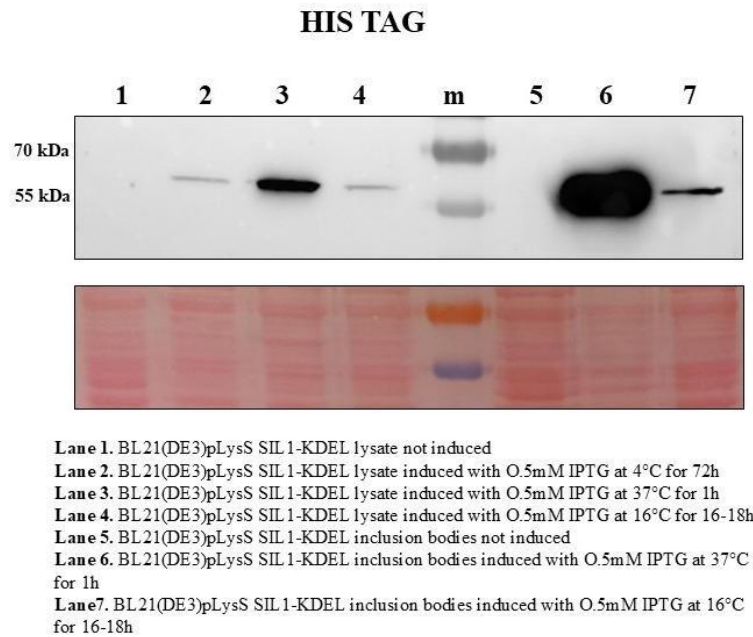


Figure 16. Comparative analysis of soluble and insoluble protein expression of SIL1-KDEL. Western blot and relative ponceau staining of SIL1-KDEL expression and inclusion body formation were assessed using 0.5 mM IPTG under different conditions.

Similarly, lysate induced for 16–18 h at 16°C with 0.5 mM IPTG (Figure 16, lane 4) yielded a moderate signal consistent with our previous results showing reduced expression under this temperature regimen. In comparison, 0.5 mM IPTG induction at 37°C for 1 h (Figure 16, lane 3) showed high SIL1-KDEL expression, in agreement with our preliminary findings that 37°C induction yields higher total protein levels. Appearance of His-tagged protein in all lysate fractions, although to different extents, suggests that variation in expression is not primarily the result of His-tag loss. Inclusion body fractions showed a fraction similar to lysate fractions. The inclusion body fraction induced by 0.5 mM IPTG at 37°C for 1 h (Figure 16, lane 6) shows that a significant fraction of SIL1-KDEL is accumulated into inclusion bodies at 37°C, as compared with the inclusion body fraction induced by 0.5 mM IPTG at 16°C for 16–18 h (Figure 16, lane 7).

The capacity to detect His-tagged SIL1-KDEL protein in all conditions studied, including inclusion body fractions, suggests that variations in protein abundance seen are likely a result of His-tag cleavage. Instead, these changes appear to represent actual variations in SIL1-KDEL expression and localization, subject to the induction conditions.

4.2.1 Purification of recombinant SIL1-KDEL from inclusion bodies

Initial purification efforts of the SIL1-KDEL fusion protein included purification of the protein from the soluble fraction by IMAC on an ÄKTA Prime Plus system. The soluble yield of SIL1-KDEL was not sufficient to meet the needs of the application (*data not shown*). Purification strategies were therefore directed to the insoluble fraction. To evaluate the effectiveness of IMAC in purifying SIL1-KDEL from the inclusion body, we performed Western blot analysis of various fractions obtained during purification against an anti-SIL1 antibody (Figure 17).

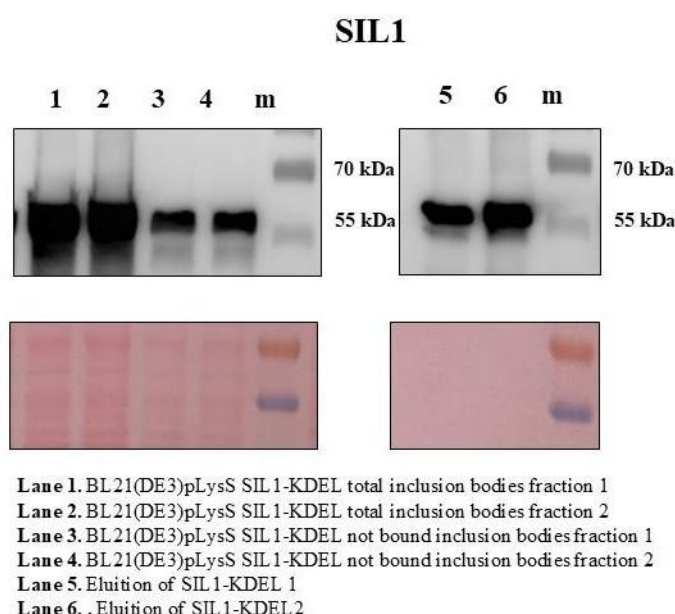


Figure 17. SIL1-KDEL purification from inclusion bodies. Western blot and relative ponceau staining of SIL1-KDEL purification from inclusion bodies using AKTA Prime Plus system.

Total inclusion body fractions (loaded in duplicate, Figure 17, lanes 1 and 2) show strong signals at approximately 70 kDa strengthening the presence of SIL1-KDEL in the inclusion bodies. The "unbound" fraction (loaded in duplicate, Figure 17, lanes 3 and 4) allows us to know the amount of protein not bound to the IMAC column. A strong marked decrease in signal intensity is evident compared to the total inclusion body fractions (Figure 17, lanes 1 and 2). The presence of residual signal shows that binding was not complete or that some parts of the SIL1-KDEL protein were lost during sample preparation. The elution fractions (lanes 5 and 6) have strong signals at around 70 kDa confirming the successful elution of the

SIL1-KDEL protein from the IMAC column. Despite the efforts, the yield of purified protein recovered after elution from the inclusion bodies remained insufficient.

4.3 Development of a pGEX6P-1-based expression system for SIL1-CPP fusion proteins

To address limitations encountered during purification of the His-tagged SIL1-CPP fusion protein, we implemented a strategic redesign of the expression construct. Specifically, the SIL1-CPPs gene, lacking its native signal peptide sequence, was subcloned into the pGEX6P-1 expression vector. This modification resulted in the generation of a recombinant protein carrying an N-terminal glutathione S-transferase (GST) tag, separated from the SIL1-CPP sequence by an HRV 3C protease cleavage site (LEVLFQGP).

This transition to a GST-tagged system was based on the established advantages of larger affinity tags in purifying recombinant proteins. GST tags are known to improve the solubility, stability, and binding affinity of proteins to affinity matrices, thereby improving overall purification yields. These attributes are particularly advantageous for proteins that exhibit aggregation tendencies or low levels of expression, as observed with the initial His-tagged SIL1-CPP construct. In addition, deletion of the native signal peptide sequence was found to be crucial. Native signal peptides, primarily designed for eukaryotic cellular trafficking, are nonfunctional in *E. coli*. Bacterial cells lack the complex membrane-bound organelles and associated trafficking machinery required for proper signal peptide processing.

Due to their potential to facilitate intracellular distribution and ER localization of the SIL1 protein, the CPP TAT, along with its homolog TAT κ and the ER retention signal KDEL were selected for inclusion in the SIL1-CPP fusion construct.

In Figure 18, a schematic representation of the pGEX6P-1 expression vector (left) and the resulting SIL1-CPP fusion protein constructs (right) is provided. The schematic highlights the key features of the construct, including the GST tag, the HRV 3C cleavage site, and selected CPP and ER retention signal components.

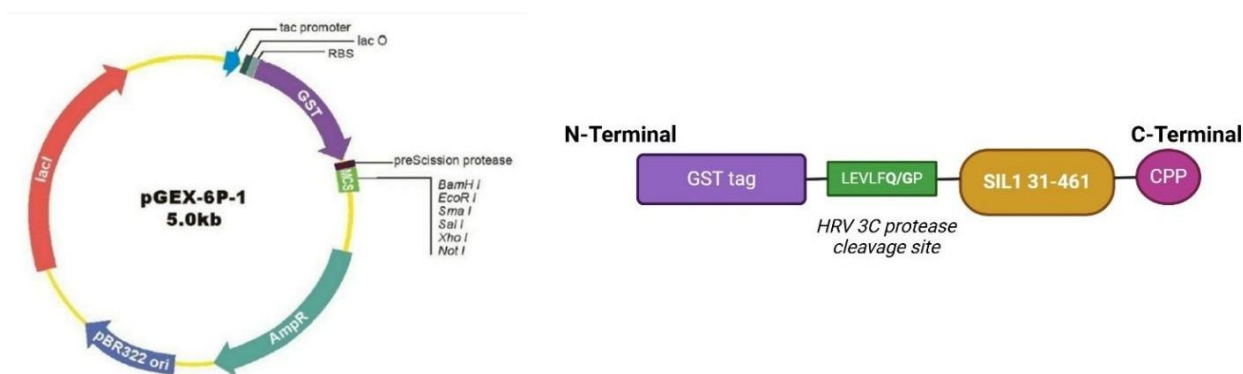


Figure 18. Strategies to improve purification of SIL1-CPPs fusion protein. pGEX6P-1 plasmid vector and SIL1-CPPs fusion proteins constructs. *Created with Biorender.*

4.3.1 Expression of GST-tagged SIL1-CPP constructs in E. coli BL21(DE3)

For optimal recombinant protein expression, E. coli BL21(DE3) cells were used as the expression host to utilize their capacity for greater expression levels than BL21(DE3)pLysS. Plasmid DNA expressing SIL1 fusion constructs containing TAT, TAT κ , or KDEL sequences was used to transform bacterial cells, and protein expression was examined using Western blotting. Supernatant (soluble) and pellet (insoluble) fractions were both probed using an anti-SIL1 antibody (Figure 19).

Analysis of the supernatant fractions revealed distinct bands at approximately 70 kDa in lysates from two independent colonies expressing SIL1-TAT (Figure 19, lanes 2 and 3), SIL1-TAT κ (Figure 19, lanes 4 and 5), and SIL1-KDEL (Figure 19, lanes 6 and 7). These bands confirm successful expression of all three SIL1-CPP constructs in the soluble fraction.

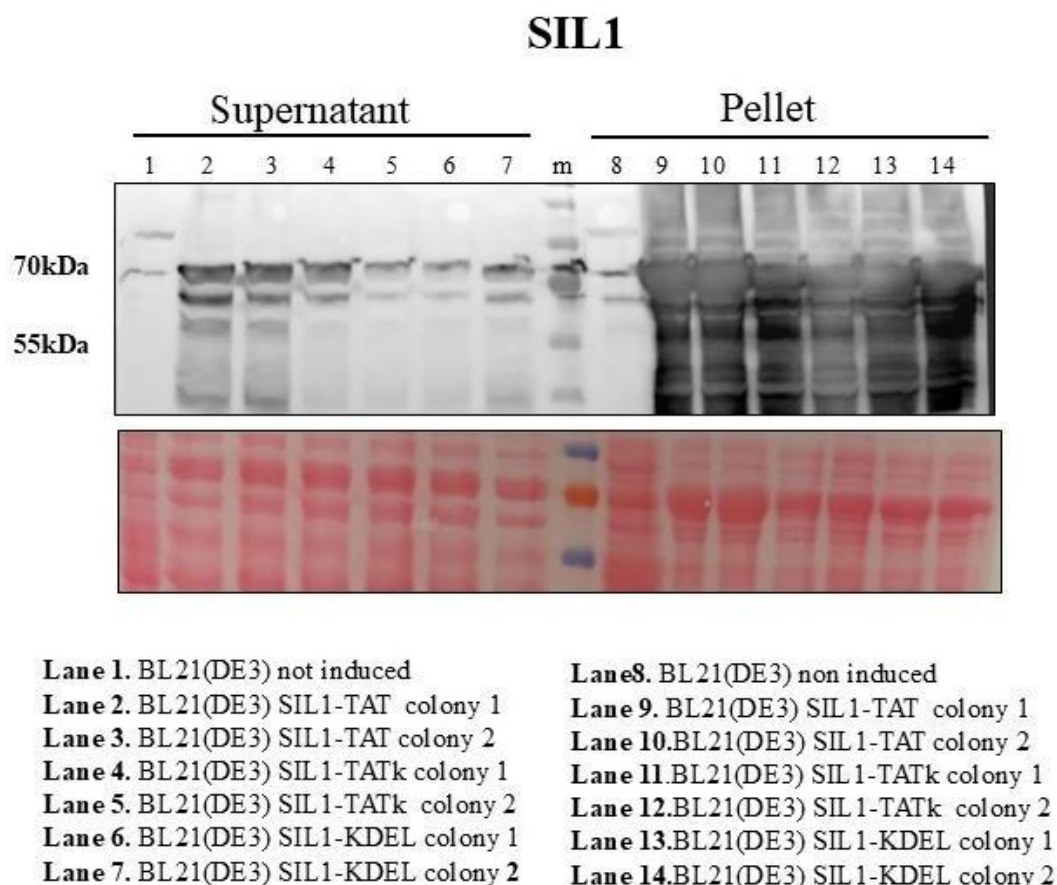


Figure 19. Evaluation of the expression of GST-tagged SIL1-CPPs in BL21(DE3). Western blot and relative ponceau staining of GST-tagged SIL1-CPPs expression in BL21(DE3) cells in the supernatant and pellet.

In parallel, analysis of the pellet fractions showed the presence of SIL1-CPP proteins in the insoluble fraction for the three constructs (Figure 19, SIL1-TAT: lanes 9 and 10; SIL1-TATk: lanes 11 and 12; SIL1-KDEL: lanes 13 and 14). While this indicates the presence of inclusion bodies, the signal intensity in the pellet fractions was lower relative to the earlier reported inclusion body formation with the histidine-tagged SIL1-KDEL protein (Figure 16). This shows that the substitution to the pGEX6P-1 vector and the GST tag, as well as the deletion of the signal peptide, resulted in a relative increase in soluble SIL1-CPP protein. Additionally, comparison of expression between the three constructs indicated that the greatest soluble protein expression was of SIL1-TAT that was selected for further investigations.

4.3.2 Optimization of GST-tagged SIL1-TAT purification: key parameters and scale-up

To optimize the affinity chromatography purification process for GST-tagged recombinant SIL1-TAT protein, we focused on two critical parameters: resin-to-protein ratio and the binding buffer composition.

4.3.2.1 Impact of binding buffer composition on SIL1-TAT purification

The effect of binding buffer composition on protein-ligand interaction was analyzed by Western blot analysis using an anti-SIL1 antibody (Figure 20, Panels A and B) and Coomassie staining (Figure 20, Panel C).

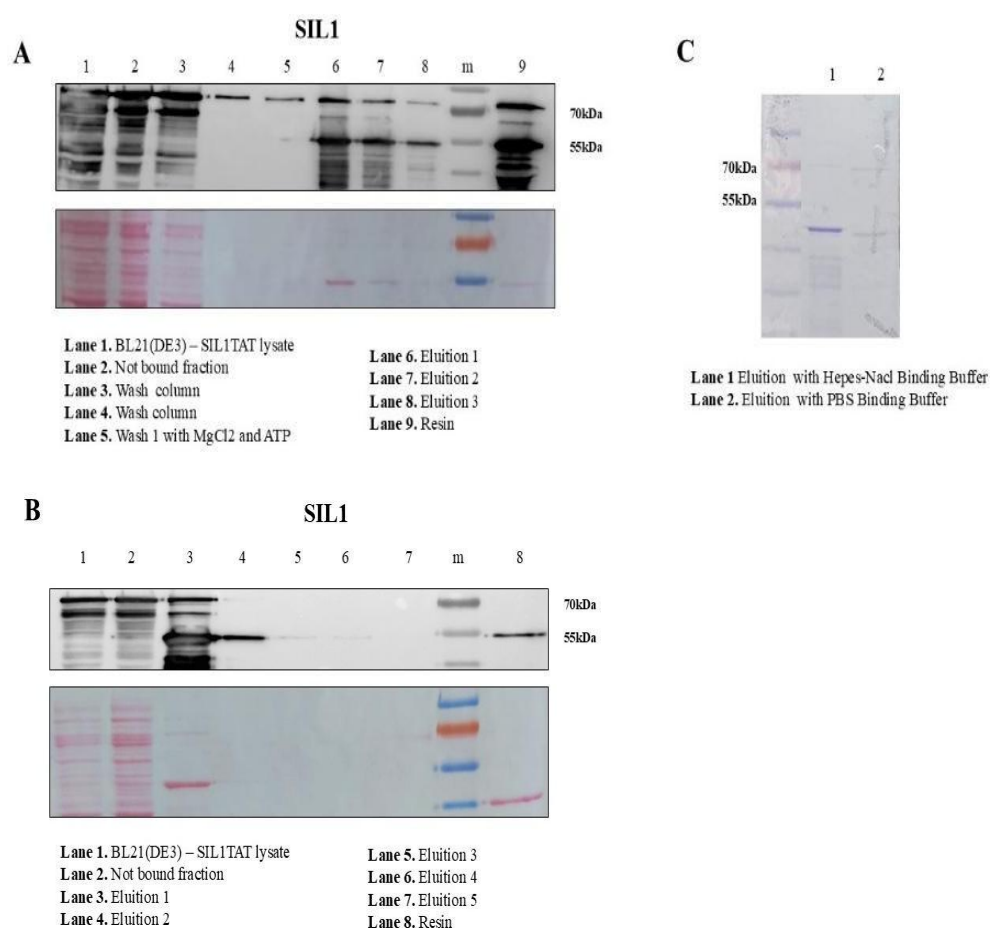


Figure 20. Impact of Binding Buffer on SIL1-TAT Purification. (A) Western blot and relative ponceau staining of SIL1-TAT purified using 50 mM Hepes, 400 mM NaCl, pH 7.4 as the binding buffer. (B) Western blot and relative ponceau staining of SIL1-TAT purified using PBS (without Mg²⁺ and Ca²⁺) as the binding buffer. (C) Coomassie-stained SDS-page comparing SIL1-TAT yields from both purification series.

Initial purification was performed with 50 mM Hepes, 400 mM NaCl, pH 7.4 as binding buffer, in addition to on-column cleavage of the GST tag. Western blot analysis (Figure 20, Panel A) showed effective expression of SIL1-TAT in the total lysate (lane 1, panel A) and a slight reduction in signal intensity in the unbound fraction (lane 2, Figure 20, panel A) illustrating protein binding to the affinity resin. Resin washes (lanes 3 and 4, Figure 20, panel A) had minimal signal, demonstrating that unbound protein had been successfully washed away. The wash step used to remove GroEL contamination ¹⁰²(lane 5, Figure 20, panel A) had a weak signal, suggesting minimal loss of SIL1-TAT. The elution fractions (lanes 6, 7 and 8, Figure 20, panel A) had variable signal intensities, indicating that the majority of SIL1-TAT was eluted in the initial fraction (Figure 20, lane 6, panel A). To further monitor the elution efficiency, a fraction of the resin remaining after the elution process was subjected to analysis. The resin (lane 9, Figure 20, panel A) showed a very weak signal, which confirmed the successful elution of the bound protein.

A second purification was performed using PBS (without Mg²⁺ and Ca²⁺) as binding buffer. Western blot analysis (Figure 20, panel B) again confirmed the successful expression of SIL1-TAT in the total lysate (lane 1, Figure 20, panel B), whereas, in the unbound fraction (lane 2, Figure 20, panel B), the residual signal indicates possible incomplete binding. The elution fractions (lanes 3, 4, 5, 6, and 7, Figure 2, panel B) showed variable signal intensities, with most of the SIL1-TAT eluted in the first two fractions (lanes 3 and 4, panel B). The resin (lane 8, Figure 20, panel B) showed a very weak signal, confirming efficient elution.

Coomassie staining (Figure 20, panel C) was used to compare the elution profiles obtained with Hepes-NaCl and PBS binding buffers. Elution with PBS as binding buffer, appeared to produce a purer band pattern with less co-purification of contaminants. However, the overall yield of SIL1-TAT appeared to be lower than with Hepes-NaCl binding buffer. These results demonstrate that the choice of buffer has a significant impact on both protein purity and yield.

4.3.2.2 Impact of resin-to-protein ratio on SIL1-TAT purification

The presence of residual protein in the unbound fractions suggested that binding conditions may require optimization. To ensure maximal binding efficiency of SIL1-TAT protein to Glutathione Sepharose 4B resin, we conducted a series of experiments with particular attention to incubation time and rebinding of free protein. Western blot analysis using an

anti-SIL1 antibody was performed for various fractions obtained during purification (Figure 21).

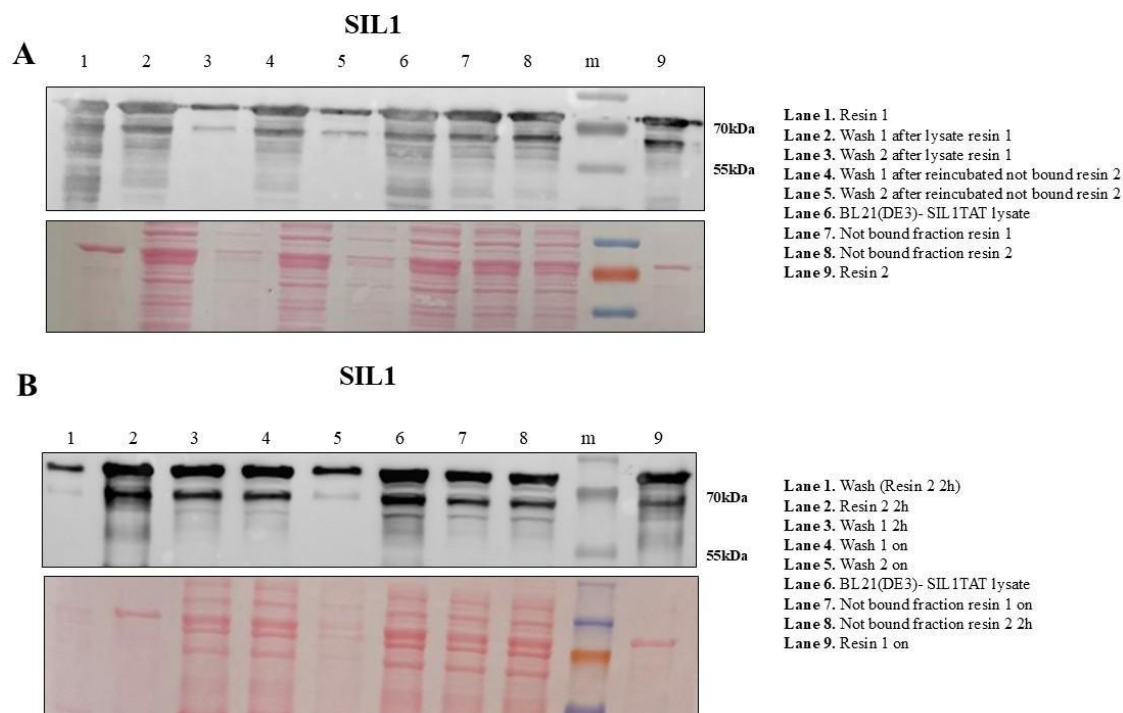


Figure 21. Optimization of SIL1-TAT binding to Glutathione Sepharose 4B resin. (A) Re-incubation of the initial unbound fraction with fresh resin. (B) Comparison of 2-hour and overnight incubation times.

We tested the influence of increasing the amount of resin to increase the total number of available binding sites and improve the capture of the protein of interest. In particular, we reincubated the initial unbound fraction (lane 7, Figure 21, panel A) for 2 h with an equal volume of fresh Glutathione Sepharose 4B resin. Western blotting (Figure 21, panel A) showed a further decrease of SIL1-TAT in the unbound fraction after reincubation (lane 8, panel A) compared to the initial unbound fraction (lane 7, Figure 21, panel A). The amount of SIL1-TAT bound to the second resin (lane 9, Figure 21, panel A) showed that reincubation could be helpful in more efficient capture of the protein of interest.

In parallel, to verify which incubation time was optimal for the binding of SIL1-TAT to the affinity resin, we compared two incubation times: a 2-h incubation and an overnight incubation. Western blotting (Figure 21, Panel B) revealed that the percentage of SIL1-TAT protein that bound to resin after overnight incubation (lane 9, Figure 21, Panel B) was significantly higher than the percentage bound after 2 h (lane 2, Figure 21, Panel B).

The observation that overnight incubation increases binding is perhaps a consequence of increasing the contact time between protein and resin to promote better interaction. These results demonstrate that extended overnight incubation with increased amount of resin enhances the binding of SIL1-TAT to glutathione Sepharose 4B resin.

4.3.2.3 Enhanced purification of SIL1-TAT via automated FPLC

To improve the purity and reproducibility of SIL1-TAT purification, the previously optimized purification protocol was translated from batch purification to fast protein liquid chromatography (FPLC) using an ÄKTA start system. To evaluate the purification efficiency of SIL1-TAT protein using affinity chromatography, we performed Coomassie staining on various fractions obtained during the purification process (Figure 22).

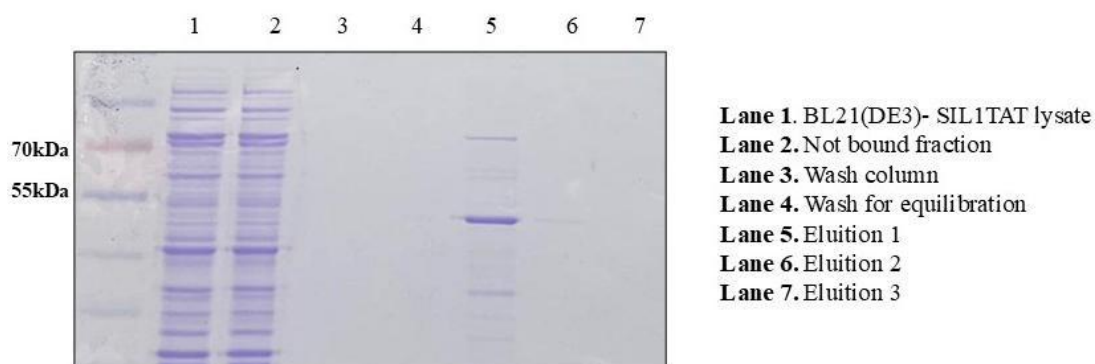


Figure 22. SIL1-TAT purification by affinity chromatography. Coomassie staining SDS-page of SIL1-TAT purification by Fast Protein Liquid Chromatography.

In the total lysate of *E. coli* BL21(DE3) cells expressing SIL1-TAT(lane 1), the most prominent band around 70 kDa confirms the successful expression of SIL1-TAT, although it is difficult to specifically identify the protein of interest due to the complexity of the lysate. In the "unbound" fraction (lane 2), a significant reduction in overall protein staining intensity is observed compared to the total lysate (lane 1). This indicates that a substantial portion of the proteins, including SIL1-TAT, has successfully bound to the affinity resin. Successful removal of unbound proteins from the wash steps (lanes 3 and 4) confirms the specificity of the affinity purification through the evidence of low intensity of staining. The first elution (lane 5) shows a prominent band at approximately 55 kDa, suggesting successful elution of

the SIL1-TAT protein, after cleavage of the GST tag. The next elutions (lanes 6 and 7) show extremely low intensity of staining, which suggests that most of the SIL1-TAT protein is eluted in the first fraction of elution. The presence of the 70 kDa band and a less intense band in the region below 55 kDa (lane 5) suggests possible presence of contaminants, partial cleavage of the protein in-column, or possible presence of degradation products.

The transition to an automated system not only resulted in an improvement in purity, but also a significant increase in the overall yield of purified SIL1-TAT protein. Further tests are needed to identify the nature of the less intense bands present in the elution product.

4.3.3 Evaluation of SIL1-TAT purity by mass spectrometry

To fully characterize the recombinant SIL1-TAT protein and evaluate its purity, MS was used. In fact, SDS-PAGE analysis of the purified SIL1-TAT protein (Figure 22, lane 5) revealed the presence of additional bands beyond the expected ~50 kDa band of the target protein, so the SDS-PAGE gel regions corresponding to the SIL1-TAT band and contaminant bands (Figure 23, panel A) were excised and subjected to trypsin in-gel digestion. This involved enzymatic cleavage of the separated proteins into smaller peptides, followed by extraction of the peptides and preparation for MS.

The resulting peptide samples were analysed by nanoLC-ORBITRAP-MS/MS coupled with a Thermo Scientific Orbitrap Fusion Tribrid mass spectrometer. The acquired mass spectra were searched against the *Homo Sapiens* and *E. coli* UniProtKB/Swiss-Prot database, using the Sequest HT search engine within Proteome Discoverer software.

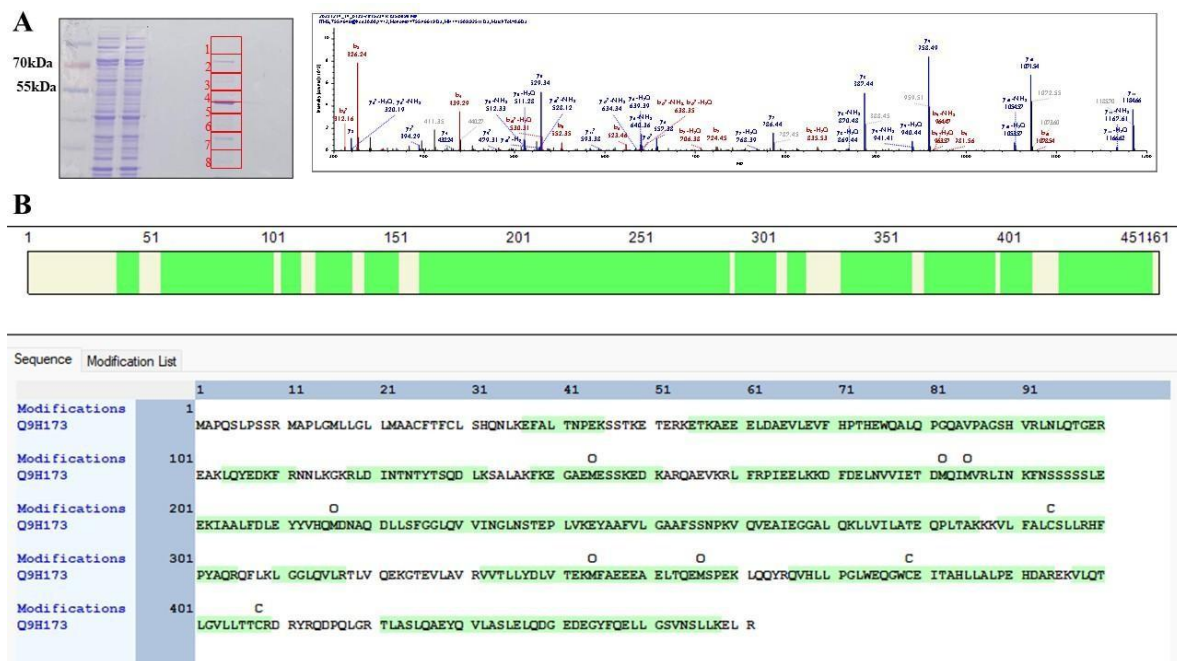


Figure 23. Analysis of SIL1-TAT purification and peptide identification. (A) Coomassie-stained SDS-PAGE gel of purified SIL1-TAT, showing representative gel bands excised for mass spectrometry (MS) analysis (left). Corresponding representative mass-to-charge (m/z) spectra obtained from MS analysis are also shown (right). (B) Illustration of peptide-spectrum matching (PSM). The fragmentation pattern observed by MS/MS is compared to theoretical spectra derived from protein databases. This process generates PSMs, linking the observed spectra to specific peptide sequences of SIL1.

In Figure 23, panel A, a representative mass spectrum (right) of the SDS-PAGE gel-excised band for SIL1-TAT fusion protein is shown with specific peaks identified as corresponding peptide fragments that were used to identify the protein. In Figure 23, panel B, green bars indicate regions of the SIL1 protein sequence determined by mass spectrometry and providing sequence coverage information and confirming its identity. The combination of SDS-PAGE, mass spectrometry, and sequence analysis allowed for the identification of the protein and determination of molecular weight and clarification of potential post-translational modifications. In particular, analysis of the gel slice corresponding to the expected SIL1-TAT band (~50 kDa, band 5) confirmed the presence of the human SIL1 sequence with 77% sequence coverage and 24 unique peptides identified. The band migrating at ~70 kDa (band 2) also showed SIL1 as the top hit, but with a slightly lower coverage of 74%, indicating, as suspected, the presence of a portion of uncut SIL1-TAT. This suggests that some of the SIL1-TAT protein may not have been fully processed.

MS allowed us to identify bacterial contaminants present. Additionally, the first protein identifications for each excised gel band against the E. coli UniProtKB/Swiss-Prot database using the Sequest HT search engine within Proteome Discoverer are summarized in Table 4. Interestingly, the beta subunit of E. coli ATP synthase was also identified as a significant component in band 5, suggesting minimal co-purification with SIL1-TAT. The analysis also showed identification of DnaK, the bacterial homologue of BiP, in the band migrating at ~70 kDa (band 2). In band 3 (~60 kDa), Chaperonin GroEL, an E. coli chaperone, was identified with 66% coverage, suggesting potential interaction or co-purification with SIL1-TAT. A number of bacterial proteins associated with membrane components were identified in the lower molecular weight bands (bands 6, 7, 8 etc.), consistent with the presence of small amounts of E. coli contaminants. These results demonstrate that while the purified SIL1-TAT preparation contained the target protein as the major component, significant contamination with E. coli proteins was observed, including ATP synthase beta subunit, GroEL and membrane-associated proteins. These results highlight the need for additional purification steps to remove these contaminating species before SIL1-TAT can be used in downstream applications.

Accession	Description	Sum PEP Score	Coverage [%]	# Peptides	# PSMs	# Unique Peptides	MW [kDa]
First gel band							
P0A6M8	Elongation factor G OS=Escherichia coli (strain K12)	26.162	27	11	14	11	77.5
P05055	Polyribonucleotide nucleotidyltransferase OS=Escherichia coli (strain K12)	25.539	24	11	14	11	77.1
P0A940	Outer membrane protein assembly factor BamA OS=Escherichia coli (strain K12)	24.314	23	14	14	14	90.5
P31554	LPS-assembly protein LptD OS=Escherichia coli (strain K12)	18.85	23	13	15	13	89.6
P63284	Chaperone protein ClpB OS=Escherichia coli (strain K12)	18.317	21	11	12	11	95.5

Second gel band							
P05055	Polyribonucleotide nucleotidyltransferase OS=Escherichia coli (strain K12)	42.9	35	16	21	16	77.1
P0AAI3	ATP-dependent zinc metalloprotease FtsH OS=Escherichia coli (strain K12)	41.093	34	17	21	17	70.7
P0A6Y8	Chaperone protein DnaK OS=Escherichia coli (strain K12)	23.048	33	12	14	12	69.1
W1F6E5	Elongation factor 4 OS=Escherichia coli ISC7	29.742	32	15	16	15	66.8
P0AG67	Small ribosomal subunit protein bS1 OS=Escherichia coli (strain K12)	29.155	32	12	13	12	61.1
Third gel band							
P0A6F5	Chaperonin GroEL OS=Escherichia coli (strain K12)	77.364	66	21	42	19	57.3
P0AC41	Succinate dehydrogenase flavoprotein subunit OS=Escherichia coli (strain K12)	38.646	35	13	16	13	64.4
P0ADY1	Periplasmic chaperone PpiD OS=Escherichia coli (strain K12)	31.81	33	15	17	15	68.1
P33599	NADH-quinone oxidoreductase subunit C/D OS=Escherichia coli (strain K12)	16.117	19	9	10	5	68.2
P0A6Z3	Chaperone protein HtpG OS=Escherichia coli (strain K12)	11.913	15	7	7	7	71.4
Fourth gel band							
P0ABB0	ATP synthase subunit alpha OS=Escherichia coli (strain K12)	25.068	30	11	12	11	55.2
P0A6F5	Chaperonin GroEL OS=Escherichia coli (strain K12)	18.786	24	6	7	6	57.3
Fifth gel band							
P0ABB4	ATP synthase subunit beta OS=Escherichia coli (strain K12)	58.205	67	20	38	20	50.3

P0A8J8	ATP-dependent RNA helicase RhlB OS=Escherichia coli (strain K12)	18.038	35	9	12	9	47.1
P02930	Outer membrane protein TolC OS=Escherichia coli (strain K12)	14.622	26	8	9	8	53.7
P0A6H5	ATP-dependent protease ATPase subunit HslU OS=Escherichia coli (strain K12) O	16.99	25	8	9	8	49.6
A0A140NF01	Transcription termination factor Rho OS=Escherichia coli (strain B / BL21- DE3)	13.238	21	7	8	7	47
Sixth gel band							
P0AE06	Multidrug efflux pump subunit AcrA OS=Escherichia coli (strain K12)	35.889	66	16	20	16	42.2
P0A9A6	Cell division protein FtsZ OS=Escherichia coli (strain K12)	29.537	62	14	16	14	40.3
P0A7G6	Protein RecA OS=Escherichia coli (strain K12)	21.305	49	10	11	10	38
D7Y694	Elongation factor Tu OS=Escherichia coli (strain MS 115-1)	32.552	46	12	17	12	44.9
P02931	Outer membrane porin F OS=Escherichia coli (strain K12)	22.273	46	9	17	9	39.3
Seventh gel band							
P02931	Outer membrane porin F OS=Escherichia coli (strain K12)	48.423	76	18	62	18	39.3
P0A910	Outer membrane protein A OS=Escherichia coli (strain K12)	23.592	53	11	18	5	37.2
P0ABB4	ATP synthase subunit beta OS=Escherichia coli (strain K12)	21.989	29	10	15	10	50.3
P0A903	Outer membrane protein assembly factor BamC OS=Escherichia coli (strain K12)	16.099	38	8	9	8	36.8

W1EVD7	Outer membrane protein assembly factor BamB OS=Escherichia coli ISC7	15.538	48	10	10	10	33.5
Eighth gel band							
A0A0H3EUC9	Beta-lactamase OS=Escherichia coli O83:H1 (strain NRG 857C / AIEC)	73.964	85	25	94	25	31.5
P02931	Outer membrane porin F OS=Escherichia coli (strain K12)	31.013	69	13	23	13	39.3
P0A910	Outer membrane protein A OS=Escherichia coli (strain K12)	22.376	43	9	15	4	37.2
P0ABB4	ATP synthase subunit beta OS=Escherichia coli (strain K12)	18.151	17	6	9	6	50.3
P0A7L0	Large ribosomal subunit protein uL1 OS=Escherichia coli (strain K12)	15.254	34	6	7	6	24.7

Table 4. Top Protein Identifications in Excised Gel Bands. Proteins were identified by searching acquired mass spectra against the *E. coli* UniProtKB/Swiss-Prot database using the Sequest HT search engine within Proteome Discoverer.

4.3.3.1 Addressing GroEL contamination in SIL1-TAT purification

Lowering the *E. coli* expression temperature to 25°C was implemented to reduce bacterial contamination and in particular chaperonin contamination¹⁰⁷. This strategy minimizes cellular stress during recombinant protein production by slowing protein synthesis, improving folding, enhancing solubility, and ultimately reducing the production of stress-induced chaperones like GroEL.

To assess the impact of this change on the purification process of the recombinant SIL1-TAT protein, a 2-liter bacterial culture purification was performed using the automated ÄKTA go system (Cytiva) at Claude Bernard University of Lyon. Success of this lowering temperature strategy was confirmed by SDS-PAGE analysis of various fractions achieved during the purification process of SIL1-TAT (Figure 24).

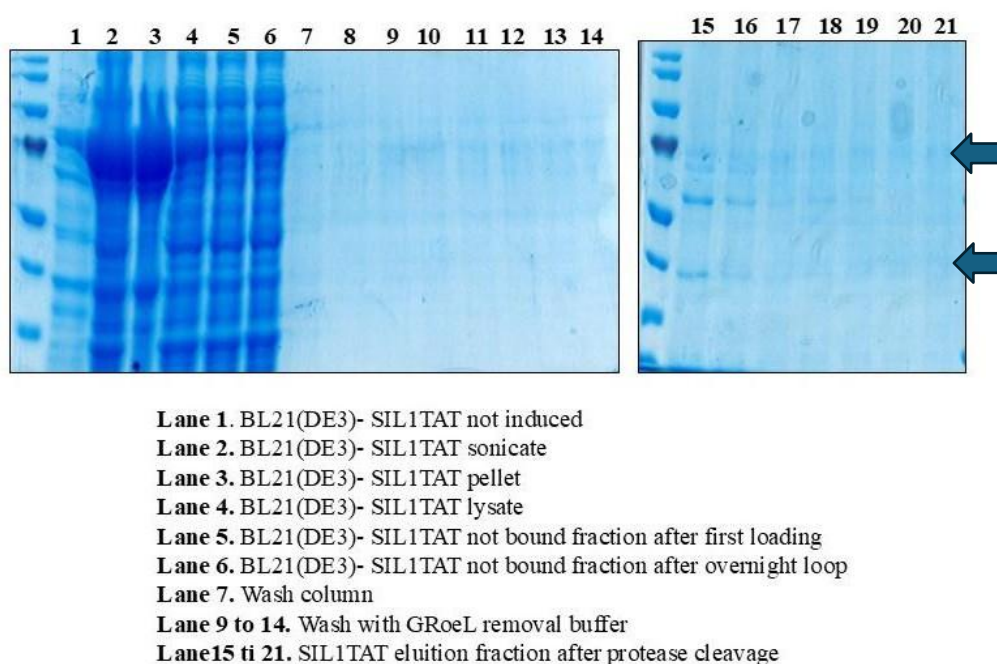


Figure 24. Reduced GroEL contamination and improved SIL1-TAT purification. Coomassie staining SDS-page of the 2-liter purification of SIL1-TAT purification via AKTA go system.

The gel exhibited successful expression of SIL1-TAT, indicated by a dominant band in the total lysate (Figure 24, lane 4) comparable to expression levels observed previously. This indicated that lowering the expression temperature had no impact on the overall yield of the recombinant protein. The presence of bands in the GroEL removal wash fractions (Figure 24, lanes 9-14) leads us to assume that some GroEL contamination was partially removed. The elution fractions (Figure 24, lanes 15-21) show a band corresponding to SIL1-TAT after protease cleavage, furthermore the bands corresponding to the GroEL contaminants (arrow in Figure 24) are less evident. The findings show that lowering the induction temperature had no effect on the yield of purification of SIL1-TAT but contributed to reduce the GroEL contamination.

4.3.4 Structural and Biophysical Characterization of SIL1-TAT

To define the structural and biophysical properties of the SIL1-TAT protein, we performed 3D structural modelling, CD spectroscopy, and DLS analysis (Figure 25). This combined DLS and CD analysis provided a comprehensive assessment of the protein's size, aggregation propensity, and structural characteristics.

Prior to biophysical characterization, the SIL1-TAT purified protein was concentrated and dialyzed against 20 mM Tris-acetate, 200 mM NaF, pH 7.4. The chosen buffer, in the absence of interfering detergents, provided an optimal environment for both protein integrity and accurate CD analysis. Protein concentration was determined by measuring absorbance at 280 nm using a calculated molar extinction coefficient (ϵ) of 34500 M⁻¹cm⁻¹ (Table 1).

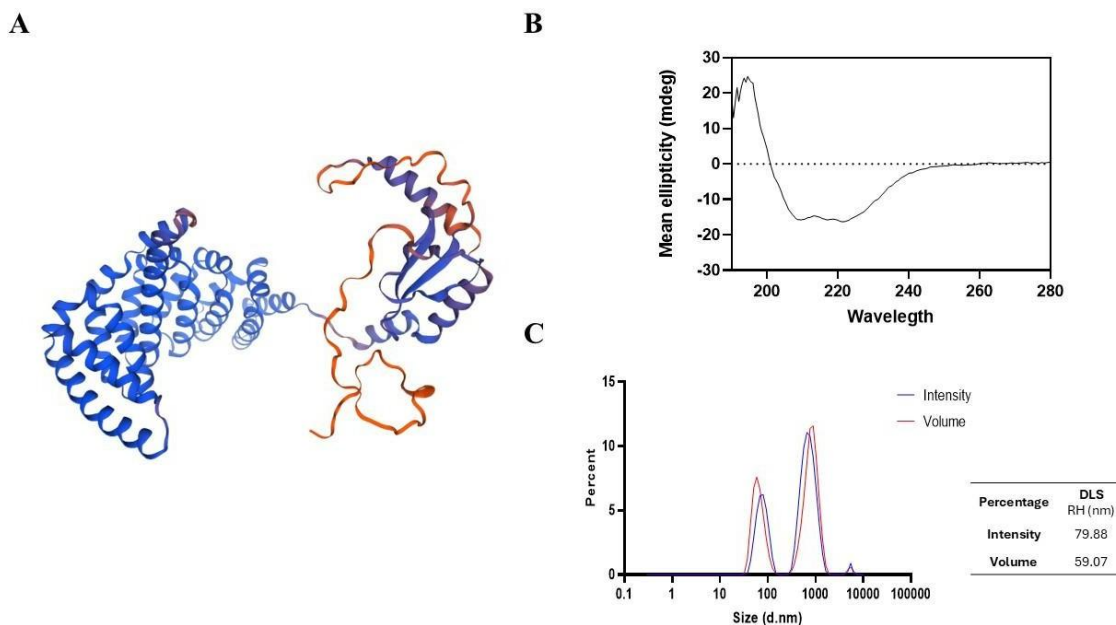


Figure 25. Structural Analysis of SIL1-TAT (A) In silico predicted three-dimensional structure of SIL1-TAT generated using SWISS-MODEL, revealing a predominantly α -helical fold. (B) CD spectrum of purified SIL1-TAT 14 μ M in 20 mM Tris-acetate, 200 mM NaF, pH 7.4 and relative (C) DLS analysis.

The secondary structure of the purified SIL1-TAT fusion protein was assessed using CD spectroscopy (Figure 25, panel B). The observed CD spectrum showed a characteristic profile of a predominantly α -helical structure with the distinctive peaks at 190 nm (positive) and 208 nm and 220 nm (negative). These data indicate the correct folding of the protein and are in agreement with our in silico prediction (Figure 25, panel A) and are further corroborated by the structural homology observed with the deposited crystal structure of the yeast homologue of SIL1 (PDB code: 3QML)¹⁵. The oligomeric state of SIL1-TAT was then investigated using DLS (Figure 25, panel C). The DLS data yielded a hydrodynamic radius (Rh) of 59.07 nm. Based on previously reported empirical predictions¹⁰⁸, this Rh value is consistent with the protein existing as a tetramer in solution. DLS analysis also revealed a minor population of aggregated protein with a larger Rh of 79.88 nm. These results provide

valuable insights into the structure and biophysical properties of SIL1-TAT, suggesting that it is a folded, stable, and potentially tetrameric protein.

4.3.4.1 Investigation of SIL1-TAT folding stability: buffer, thermal and cryopreservation effects

In order to evaluate the impact of buffer composition on secondary structure of SIL1-TAT, we performed CD spectroscopy in varied buffers (Figure 26). With Hepes-NaCl buffers, a significant variation in CD spectra was observed based on protein and salt concentration. Higher Hepes and NaCl concentrations showed a more intense α -helical signature of SIL1-TAT (Figure 26, panel A). This effect was particularly significant at 220 nm, the negative peak wavelength of the α -helical secondary structure. CD spectra measured directly after dilution of the protein in KPi buffer (Figure 26, panel B), showed a dramatic loss of secondary structure, as expected for rapid unfolding (blue). Subsequent efforts to restore native folding of SIL1-TAT after exposure to suboptimal buffer conditions were unsuccessful. Reincorporation of the 400 mM NaCl (green) and adding 20% trifluoroethanol (TFE) (purple), a known α -helix stabilizer¹⁰⁹, were not sufficient to restore the characteristic CD spectrum.

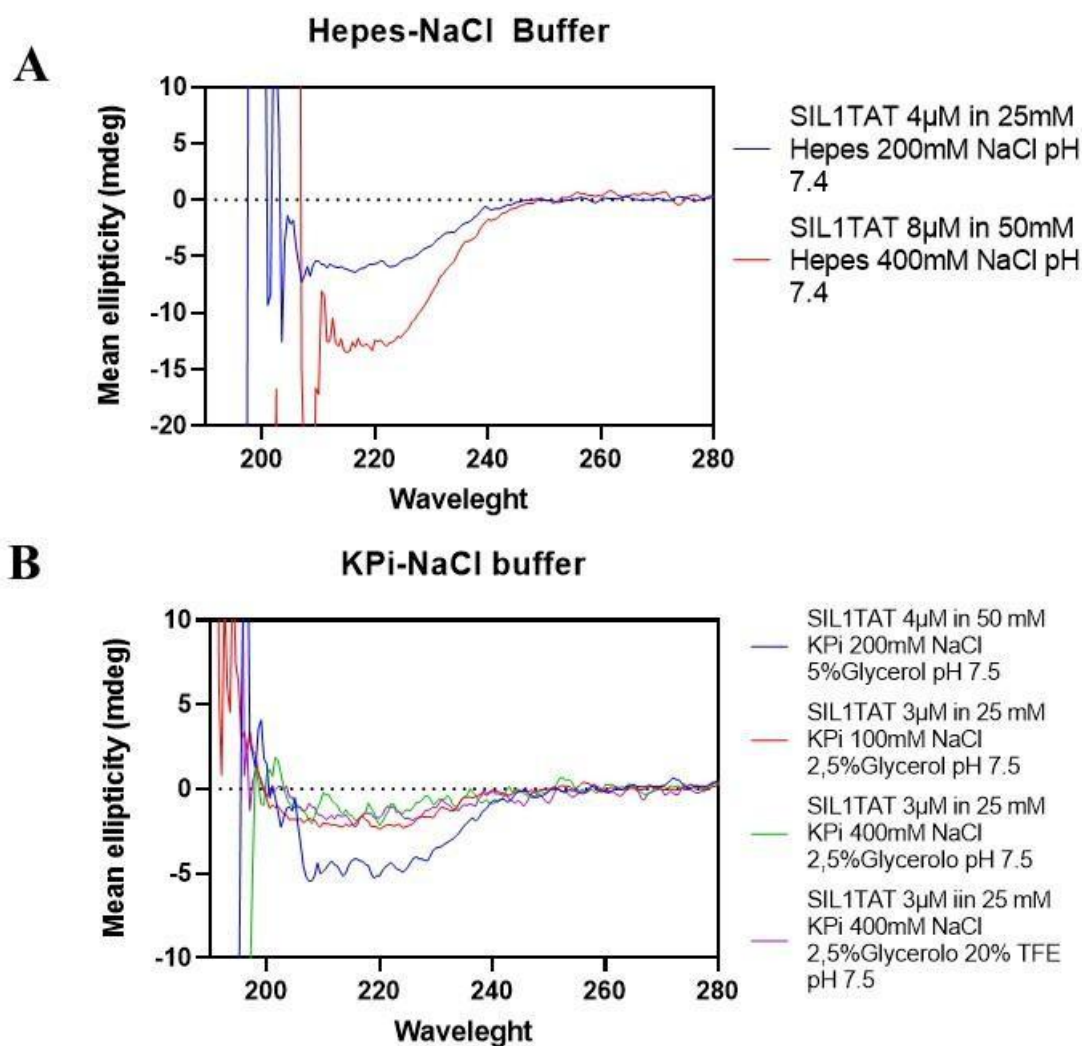


Figure 26. Evaluation of the influence of the buffer on the secondary structure of SIL 1-TAT. (A) CD spectra of SIL1-TAT in buffers containing 50mM Hepes 400 mM NaCl pH 7.4 (red) and 25 mM Hepes 200 mM NaCl pH 7.4 (blue). (B) CD spectra of SIL1-TAT in 50 mM KPi, 200 mM NaCl, 5% glycerol, pH 7.5 (blue), in 25 mM KPi, 100 mM NaCl, 2.5% glycerol, pH 7.5 (red), 25 mM KPi, 400 mM NaCl, 5% glycerol, pH 7.5 (green) and 25 mM KPi, 400 mM NaCl, 5% glycerol, 20% TFE pH 7.5.

To assess protein stability and aggregation, thermal denaturation experiments were also performed, monitoring changes in secondary structure at a fixed wavelength with increasing temperature. These experiments not only followed the unfolding of the protein but also assessed the reversibility of any observed conformational changes. Thermal melting of SIL1-TAT, performed from 5°C to 60°C and then to 5°C, showed excellent thermal stability. This reversibility, as shown by the CD melting curve in Figure 27, reveals that the secondary structure of the protein is not significantly affected by thermal denaturation and subsequent cooling. Minimal precipitation of SIL1-TAT in solution was observed after the thermal denaturation experiment.

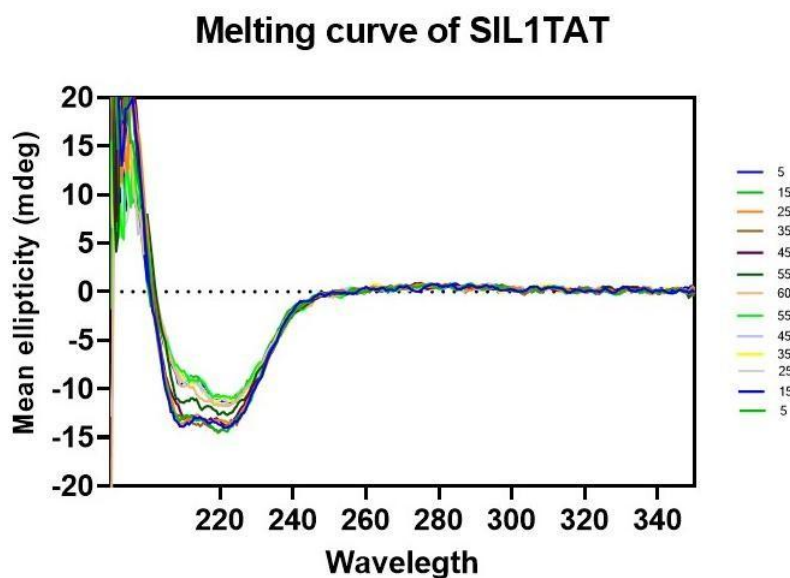


Figure 27. Characterization of SIL1-TAT thermal stability by CD spectroscopy. (A) Melting curve CD spectra of SIL1-TAT in 20 mM Tris-acetate, 200 mM NaF, pH 7.4 from 5°C to 60°C and back to 5°C.

To determine the optimal storage conditions that preserve the structural stability and activity of recombinant SIL1-TAT, we examined the effects of various temperature and cryoprotection conditions with circular CD monitoring the folding of the protein. Flash freezing and storage at -20°C and -80°C caused partial unfolding and protein precipitation (Figure 28, panel A). The unfrozen sample (blue, panel A) spectrum has normal minima at 208 nm and 222 nm, indicative of α -helical structure. The frozen samples' spectra have lost the minima significantly, indicative of loss of α -helical content. The sample frozen at -20°C (red, figure 28, panel A) shows a more pronounced loss of secondary structure than the sample frozen at -80°C (green, figure 28, panel A). To mitigate freezing-induced damage, glycerol, a cryoprotectant, was studied (Figure 28, panel B). However, addition of even 20% glycerol to SIL1-TAT solution disrupted the native folding of the protein (red, figure 28, panel B), and subsequent freezing at -20°C resulted in complete unfolding (green, figure 28, panel B). This indicates that glycerol, while often used as a cryoprotectant, may not be sufficient to protect SIL1-TAT from freezing-induced damage under these conditions.

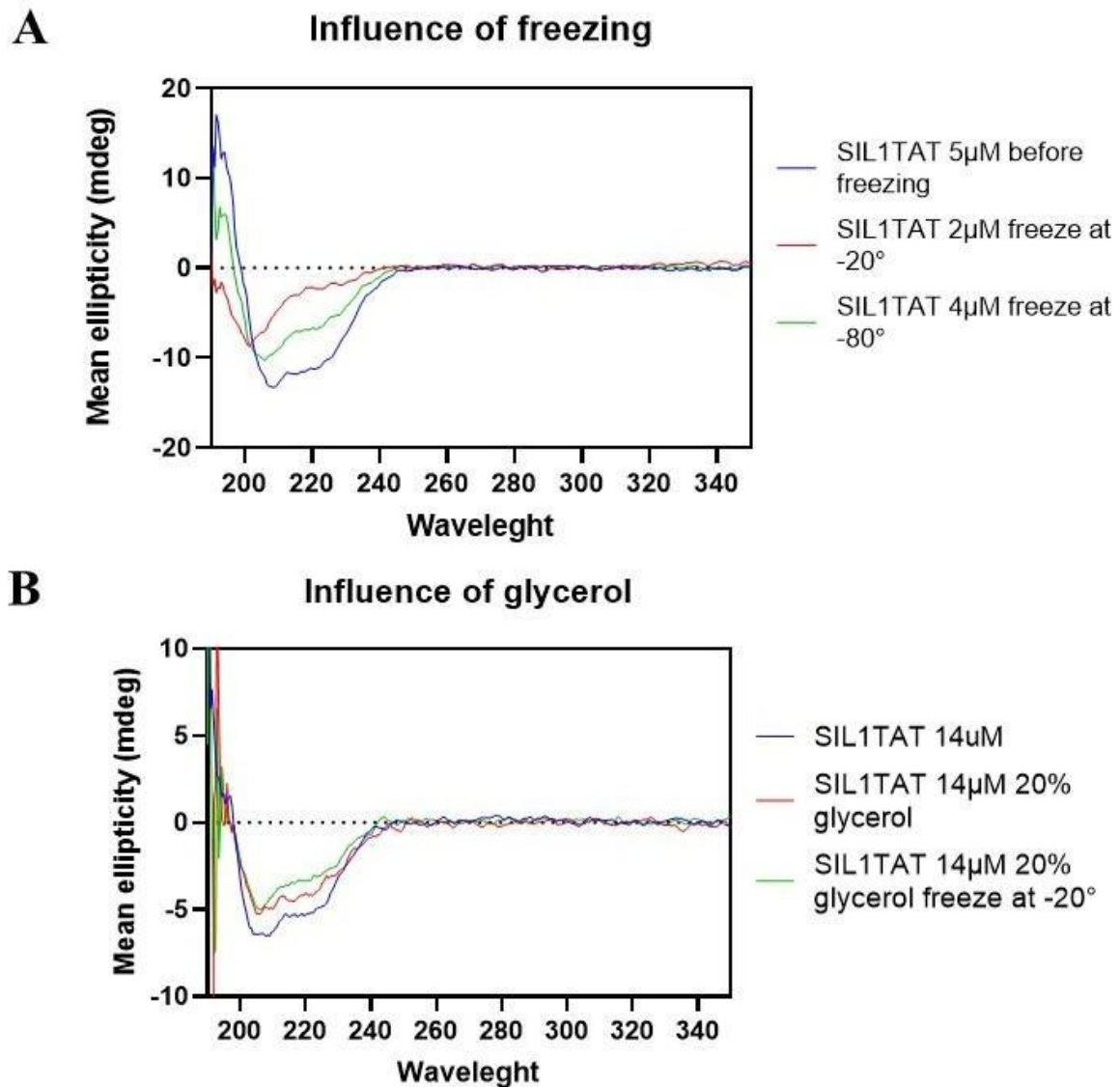


Figure 28. Impact of freezing and glycerol on SIL1-TAT structural stability. CD spectrum of SIL1-TAT in 20 mM Tris acetate 200 mM NaF pH 7.4 (blue) A) snapfrozen and then stored at -20° (red), snapfrozen and then stored at -80° (green) D) immediately after the addition of 20% glycerol (red) and then frozen at -20°C (green).

In contrast, storage of SIL1-TAT at 4°C for 72 h did not negatively impact protein folding (Figure 29). However, a quantifiable decrease in protein concentration was observed, suggesting a minor loss of soluble protein. We can therefore infer that while the overall secondary structure of SIL1-TAT is stable at 4°C for 72 h, prolonged exposure to this temperature leads to a partial loss of solubility, likely due to aggregation or precipitation. These experiments together identified the optimal conditions for preserving the SIL1-TAT fusion protein.

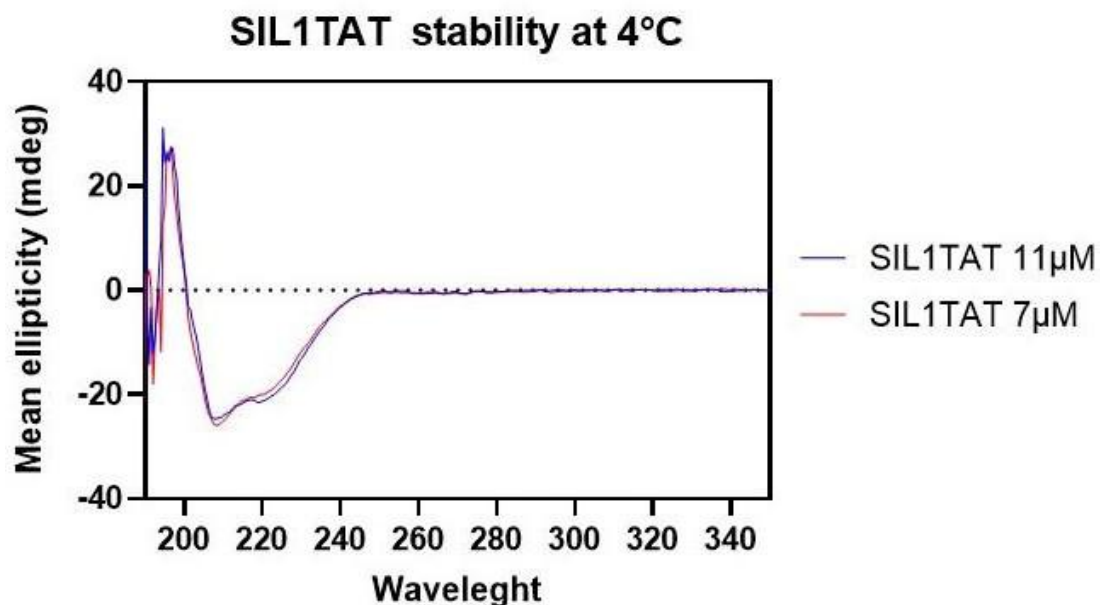


Figure 29. SIL1-TAT fusion protein stability at 4°C. CD spectrum of SIL1-TAT in 20 mM Tris acetate 200 mM NaF pH 7.4 (blue) store at 4°C for 72 hours (red).

4.3.4.2 Structure determination of SIL1-TAT via X-ray crystallography

To determine the three-dimensional structure of SIL1-TAT, crystallization trials were conducted using X-ray crystallography. Purified SIL1-TAT, dialyzed into 20 mM Tris-acetate, 400 mM NaCl, pH 7.4, and concentrated to 12 mg/mL (250 µM), was subjected to a systematic screening of crystallization conditions. Four commercially available crystallization screening kits (as detailed in the Materials and Methods section) were employed. Despite extensive screening, only six wells yielded crystal were obtained via the sitting-drop vapor diffusion method at room temperature. Of the six crystallization hits, only one condition produced crystals suitable for X-ray diffraction analysis. Needle-shaped crystals were obtained from a reservoir solution containing 20% (w/v) polyethylene glycol monomethyl ether 2000, 100 mM Tris pH 8.5, and 200 mM trimethylamine N-oxide. These crystals were subsequently subjected to X-ray diffraction analysis and diffracted X-rays to a resolution of 3 Å (Figure 30).

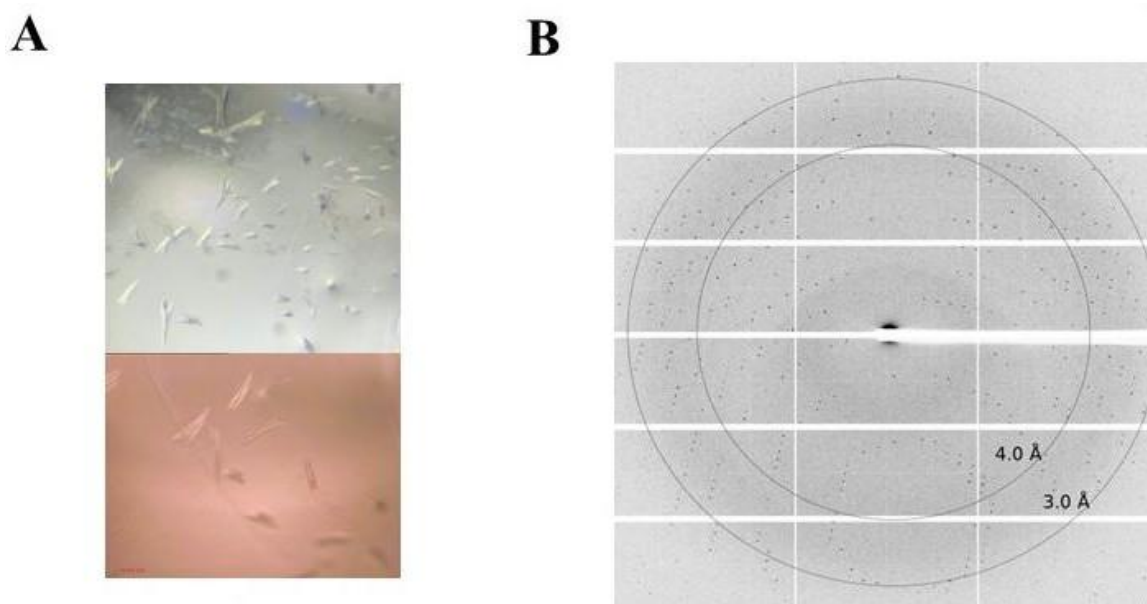


Figure 30. Crystallization trials of SIL1-TAT. (A) SIL1-TAT crystals obtain with 20 % w/v Polyethylene glycol monomethyl ether 2,000 100 mM Tris; pH 8.5 200 mM Trimethylamine N-oxide as precipitant (B) X-ray crystallography pattern of SIL1-TAT

The crystals were notably small, resulting in weak diffraction patterns characterized by limited high-resolution data. Consequently, the acquired diffraction data proved insufficient for de novo structure determination of SIL1-TAT. The 3 Å resolution data did not allow for unambiguous electron density map calculation required for accurate atomic modelling. Given the limitations in crystal size and diffraction quality, further optimization of SIL1-TAT crystallization conditions is currently underway. This optimization aims to improve crystal size and morphology, ultimately enabling the collection of higher-resolution diffraction data suitable for three-dimensional structure determination.

4.4 Structural integrity and stability studies of purified SIL1-WT protein

To provide a comprehensive and rigorous assessment and to rate the possible impact of the TAT CPP on the folding of the protein, we purified the SIL1 WT protein using the same production conditions as the fusion protein. The inclusion of the wild-type control was essential to establish a baseline and discern the specific effects attributable to the experimental modifications.

SDS-PAGE analysis of various fractions obtained during the purification process (Figure 31), demonstrates a successful expression of SIL1 WT, as evidenced by the presence of a prominent band in the elution fractions with the presence of minimal contamination bands (Figure 31, lanes 13-19). However, the production yield was lower than that of the SIL1-TAT protein.

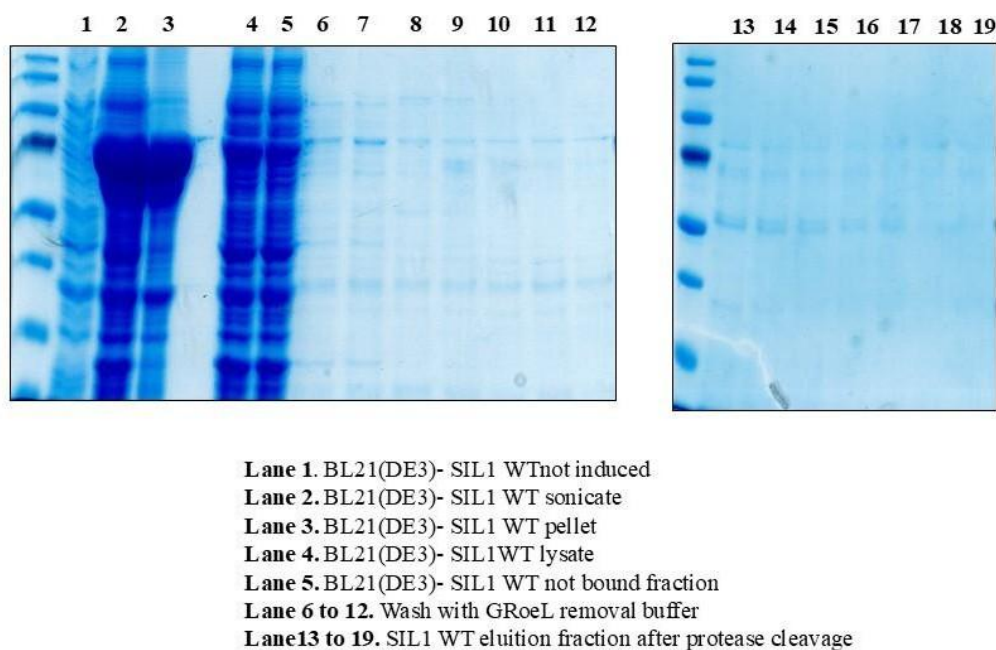


Figure 31. Evaluation of SIL1-WT expression in BL21(DE3). Coomassie staining SDS-page of SIL1-WT purification by affinity chromatography.

Prior to biophysical characterization, the SIL1-WT purified protein was concentrated and dialyzed against 20 mM Tris-acetate, 200 mM NaF, pH 7.4. Protein concentration was determined by measuring absorbance at 280 nm using a calculated molar extinction coefficient (ϵ) of 33015 M⁻¹cm⁻¹ (Table 1).

To characterize the structural and biophysical properties of SIL1 WT, CD spectroscopy and DLS analysis were performed (Figure 32). The resulting CD spectrum (Figure 32, Panel A) showed the typical features of a mixed α - β folding, revealing correct folding of the protein. DLS analysis revealed a major population with a Rh of 43.68 nm, consistent with a tetrameric oligomerization state. However, a minor population, representing less than 0.5% of the total protein, with a Rh > 92.89 nm, which is indicative of the presence of aggregated species, was observed. Besides, the proportion of this aggregated population was greater

over time, indicating temporal instability of the purified SIL1-WT protein (Figure 32, Panel B).

Thermal stability of SIL1-WT was assessed by thermal denaturation experiments followed by CD spectroscopy. The protein was subjected to a temperature gradient from 10 °C to 60 °C and a subsequent cooling cycle to 10 °C. CD melting curve analysis (Figure 32, panel C) showed relatively little irreversible unfolding of the protein. Similar to SIL1-TAT protein, flash freezing and storage at -20 °C and -80 °C led to partial unfolding and precipitation of SIL1-WT (Figure 32, panel D). The reduced production yield and increased thermal instability of SIL1 WT suggest a potentially lower intrinsic stability compared to the SIL1-TAT fusion protein. This comparative analysis led us to speculate that the presence of CPP TAT within the fusion construct may exert a stabilizing influence on the recombinant SIL1 protein.

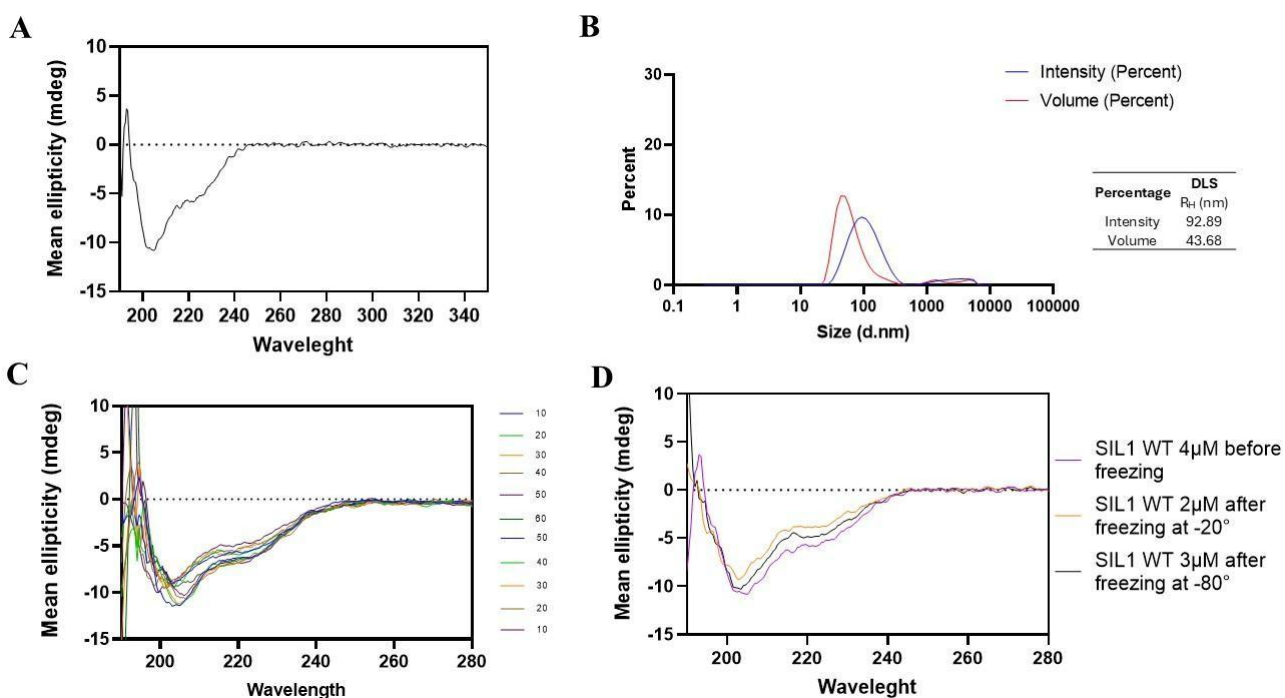


Figure 32. Structural and biophysical characterization of SIL1 WT. (A) CD spectrum of purified SIL1-WT 4 μ M in 20 mM Tris-acetate, 200 mM NaF, pH 7.4 and relative (B) DLS analysis. (C) Melting curve CD spectra of SIL1-WT in 20 mM Tris-acetate, 200 mM NaF, pH 7.4 from 10°C to 60°C and back to 10°C. (D) CD spectrum of SIL1-TAT in 20 mM Tris acetate 200 mM NaF pH 7.4 (blue), snapfrozen and then stored at -20° (orange), snapfrozen and then stored at -80° (black).

4.5 Assessing the therapeutic potential of SIL1-TAT: *in vivo* and *in vitro* studies

Developing effective therapies often requires a multifaceted approach, including both *in vivo* and *in vitro* investigations to thoroughly assess safety and efficacy. This section presents the results of key experiments designed to assess the therapeutic potential of SIL1-TAT, focusing on its biocompatibility, tissue distribution, and cellular uptake.

4.5.1 Preliminary tissue distribution of SIL1-TAT recombinant protein in mice

To ensure the safety and efficacy of SIL1-TAT, particularly for *in vivo* applications, endotoxin content in the purified SIL1-TAT eluate was measured by a chromogenic Limulus Amoebocyte Lysate (LAL) assay, as described in the Materials and Methods. SIL1-TAT elution contained 0.1 EU/mL of endotoxin. This is acceptable since the maximum endotoxin concentration that is safe for a 20g mouse is 0.1 EU administered over a 1-hour period¹¹⁰.

Following endotoxin testing, the first *in vivo* study was conducted to establish SIL1-TAT tissue distribution. A 31-week-old woozy mouse was given a single intraperitoneal injection of SIL1-TAT at a dose of 3 µg/g body weight as treatment. The mouse was monitored during and after treatment for any sign of distress. Six hours post-administration, tissue samples (quadriceps, spleen, and cerebellum) were collected from the treated mouse as well as from a control (untreated) mouse. Western blotting of the tissue samples confirmed SIL1-TAT penetration in all the tissues analysed (Figure 33).

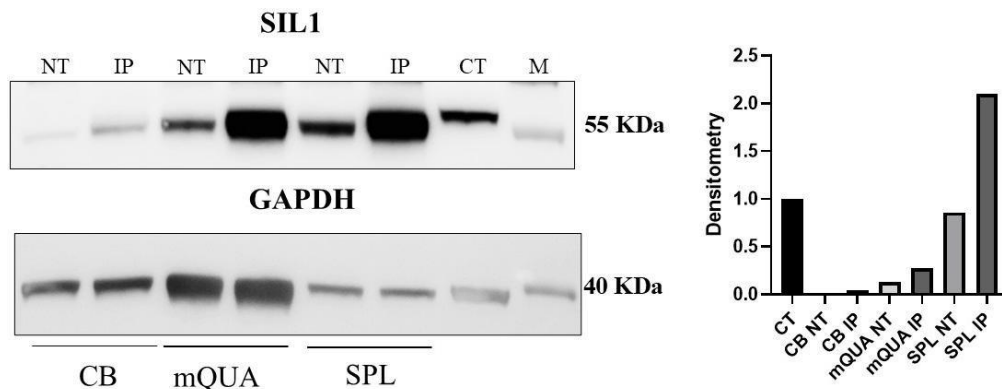


Figure 33. Western blot analysis of SIL1-TAT tissue distribution. Western blot analysis and relative quantification of SIL1-TAT in tissue lysates (cerebellum, spleen, quadriceps) from a 31-week-old *woozy* mouse, 6 hours post-intraperitoneal injection, compared to untreated control mouse tissues and HF cell lysate (control).

Western blotting revealed heterogeneous levels of SIL1 protein among different samples, with the most abundant SIL1 in spleen (top panel, figure 33). The housekeeping protein GAPDH was used as a loading control, depicting equal protein loading across lanes (lower panel, figure 33). Densitometry of SIL1 bands enabled quantitative comparison of SIL1 expression levels. This initial *in vivo* experiment provided valuable information regarding the tissue distribution of SIL1-TAT following intraperitoneal administration and offered preliminary insights into the *in vivo* stability of the fusion protein.

4.5.2 *In vitro* studies on cytotoxicity and cellular uptake of SIL1-TAT

Along with the *in vivo* experiments, the cytotoxicity of recombinant SIL1-TAT was evaluated *in vitro* by MTT assay. We evaluated the viability on patient cells (HF-MSS) and control cells (HFCTRL), at different concentrations of SIL1-TAT (Figure 34). Cell viability remained high (approximately 80-100%) at all concentrations of SIL1-TAT (0-1000 ng). No dose-dependent decline in cell viability was observed, implying that SIL1-TAT is not toxic to cells at the concentrations tested. In addition, no difference in cell viability was observed between HF CTRL and HF-MSS. These results indicate that SIL1-TAT does not induce significant cell toxicity at the concentrations tested.

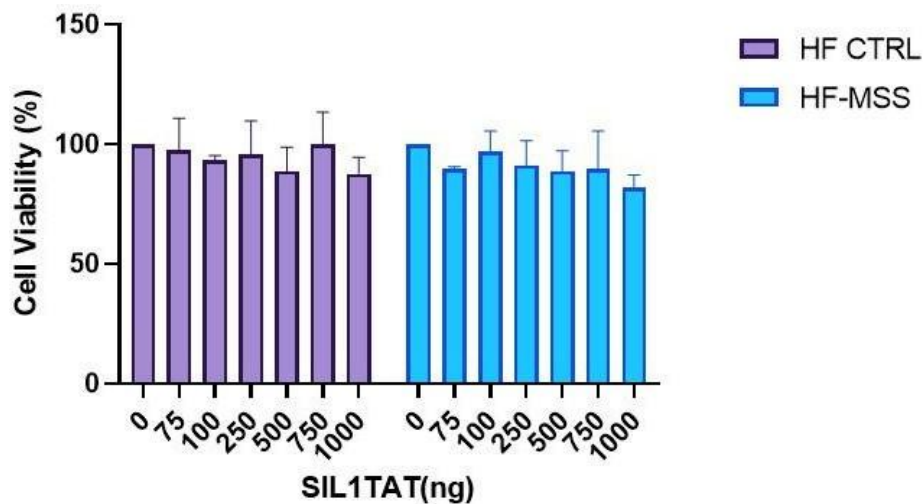


Figure 34. *In Vitro* Cytotoxicity of SIL1-TAT in HF and HF-MSS Cells. HF and HF-MSS cells were seeded in 96-well plates (5×10^3 cells/well) and treated overnight with increasing concentrations of SIL1-TAT in 50 mM Tris-HCl, 400 mM NaCl, pH 7.5. Data represent cell viability *mean* $\pm$ *SD*, of duplicate experiments.

Intracellular internalization of SIL1-TAT fusion protein in HF-MSS cells was analyzed by immunofluorescence microscopy (Figure 35). SIL1-TAT was incubated for 4 hours with HF-MSS cells at concentrations equal to those used in cytotoxicity assays. Localization of SIL1-TAT (green, figure 35) was observed by immunofluorescence staining. We predicted that, given the previously established localization of SIL1 in the ER, the SIL1-TAT fusion protein would colocalize with the ER marker HSP47 (red, figure 35). The result, however, was that no entry of SIL1-TAT was observed in HF-MSS cells. Instead, the fluorescent signal localized mostly in the extracellular space, indicating that the fusion protein was external to the cells. Concurrently, increased signal intensity was detected with increasing SIL1-TAT concentration. These data suggest that, under the conditions employed, the presence of CPP alone was not sufficient to mediate cellular uptake. Further research is needed to identify the reason for the inhibition of cellular uptake and to establish the conditions for efficient delivery of SIL1-TAT into HF-MSS cells.

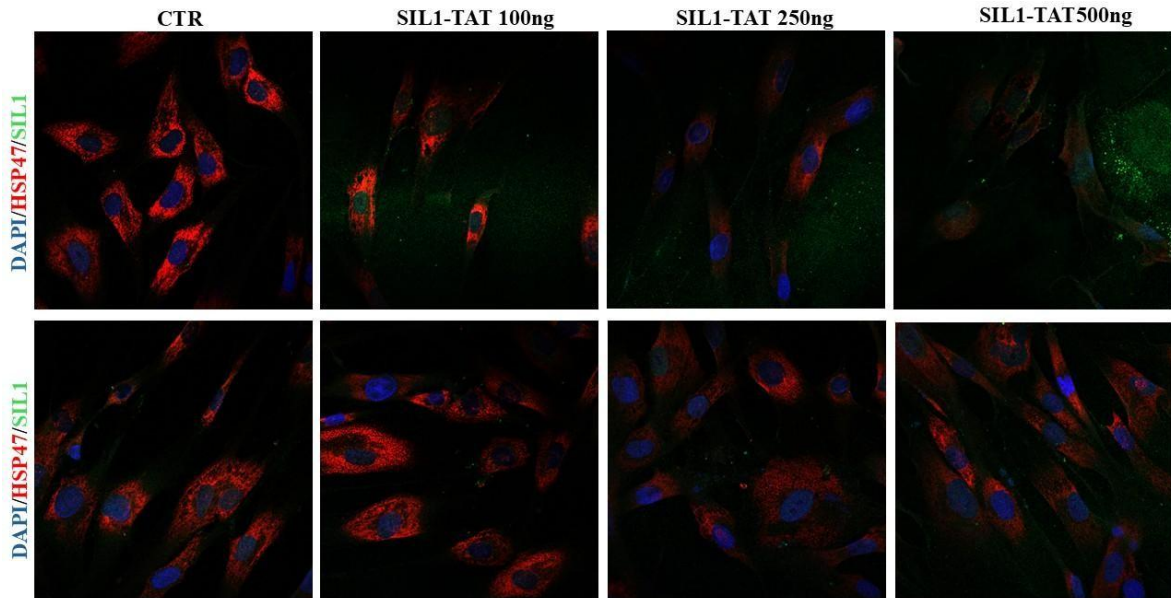


Figure 35. Evaluation of intracellular internalization of SIL1-TAT fusion protein. Fluorescence microscopy of HF-MSS (MSS) plated on glass coverslips, treated with increasing concentrations of SIL1-TAT in 50 mM Tris-HCl, 400 mM NaCl, pH 7.5 for 4 hours, then fixed in cold methanol and processed for immunofluorescence. Representative images from two independent experimental replicates (top and bottom panels) are shown. Confocal images of cells stained for HSP47 (red) and SIL1 (green) are shown.

4.6 Molecular and functional characterization of Marinesco-Sjögren syndrome

Given the potential of PRT as a therapeutic strategy for MSS, this study employed a comprehensive multifactorial approach integrating in vitro and in vivo models. Patient-derived cellular models were also used to elucidate the molecular mechanisms underlying the disease and subsequently validated in vivo. Concurrently, a wozy model was thoroughly characterized to validate its representativeness of the human disease and to define key parameters for assessing disease progression and treatment response, providing a robust platform for preclinical validation of PRT. The results presented demonstrate the synergistic value of combining these in vitro and in vivo approaches, ultimately accelerating the development of effective therapies for MSS.

4.6.1 Unravelling disease mechanisms through transcriptomic analysis

To investigate the pathogenic mechanisms underlying MSS, we performed a transcriptomic analysis of skin fibroblasts derived from a young patient diagnosed with the condition¹⁰⁶. Total RNA was extracted from patient fibroblasts cultured under standard conditions, and gene expression profiling was conducted using outsourced next-generation sequencing.

Differential gene expression analysis identified 664 differentially expressed (DE) transcripts in patient fibroblasts compared to controls. The top 50 most variably expressed genes, selected based on unbiased hierarchical clustering of log₂ fold change (logFC) values, are presented as a heatmap in Figure 36, panel A.

Gene ontology (GO) and pathway enrichment analysis of the DE genes revealed significant disruptions in several key cellular processes. Notably, this analysis confirmed a significant dysregulation of genes involved in membrane trafficking within patient fibroblasts. Furthermore, a substantial number of DE transcripts, approximately 300, were associated with the extracellular space/extracellular matrix (ECM) and cell adhesion machinery, suggesting a profound impact on these processes (Figure 36, panel B).

To validate the transcriptomic data, we performed quantitative real-time PCR (qPCR) on a subset of 13 DE genes, including *SIL1* itself. These genes were selected based on a combination of criteria: high logFC, low false discovery rate (FDR), and representation of key functional categories implicated in the disease pathogenesis, including ECM organization, cytoskeleton dynamics, vesicle trafficking, and endosome-lysosome-autophagy pathways. Specifically, seven genes were confirmed to be upregulated in patient fibroblasts, while five genes were downregulated (Figure 36, panel C).

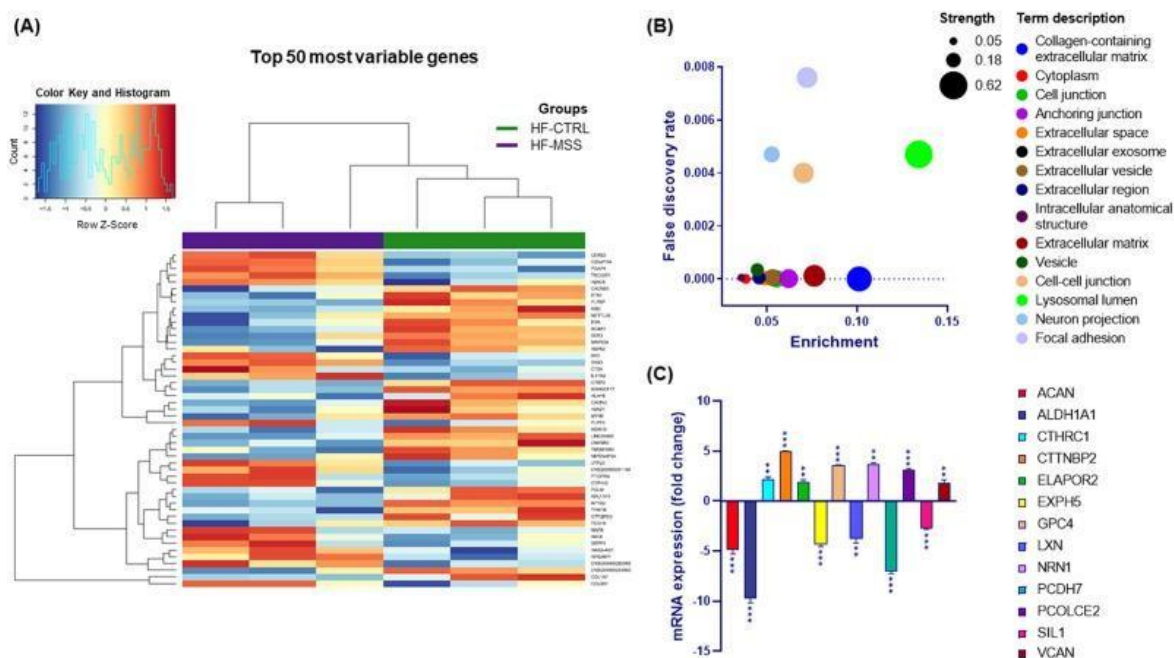


Figure 36. Differential gene expression in fibroblasts from MSS patients. (A) Heatmap of the top 50 most variably expressed genes (based on log2 fold change) in MSS patient fibroblasts compared to healthy controls. Upregulated genes are shown in red, downregulated in blue. (B) Bubble plot of enriched cellular component Gene Ontology terms. Enrichment is represented by the ratio of observed to total genes associated with each term, while strength reflects the log10 of the observed/expected gene ratio. (C) qPCR validation of 12 differentially expressed genes. Data are presented as fold change (mean $\pm$ SEM, n=3 independent experiments in triplicate). Statistical significance was determined by one-way ANOVA: $p > 0.01$, $**p > 0.001$. From “*Sil1-deficient fibroblasts generate an aberrant extracellular matrix leading to tendon disorganisation in Marinesco-Sjögren syndrome*” Amodei L. et al.

To investigate the functional consequences of the observed gene expression changes in patient fibroblasts, we performed a series of *in vitro* assays. These assays demonstrated a reduced capacity for ECM remodelling, decreased cell motility, and slower cell spreading during adhesion assays in patient fibroblasts compared to control cells (*data not shown*).

To investigate ECM remodelling in patient fibroblasts, metalloprotease activity was assessed by measuring DQ-collagen digestion in conditioned medium. Significantly lower fluorescence, indicative of reduced collagenolytic activity, was observed in patient-derived medium, confirming impaired metalloprotease function. While this likely contributes to observed defects in cell adhesion and motility, other intrinsic cellular factors, such as cellular stress responses, may also be involved (*data not shown*).

To explore the potential consequences of altered ECM processing, we performed negative-stained transmission electron microscopy (TEM) analysis in collaboration with Matteo Ardini. Conditioned medium from control fibroblasts revealed a network of thin protein filaments (11 ± 1.8 nm) (Figure 37, panel A). Medium from patient fibroblasts also showed

a network, but with thicker, ribbon-like filaments (48.8 ± 16.2 nm) that appeared to be aggregated, forming primitive fibrils reminiscent of collagen organization (Figure 37, panel B). These larger fibrillar aggregates were absent in control medium. These results suggest that while control fibroblasts deposit immature ECM at 24 h, patient fibroblasts show premature and potentially disorganized fibril formation, likely due to altered ECM remodelling.

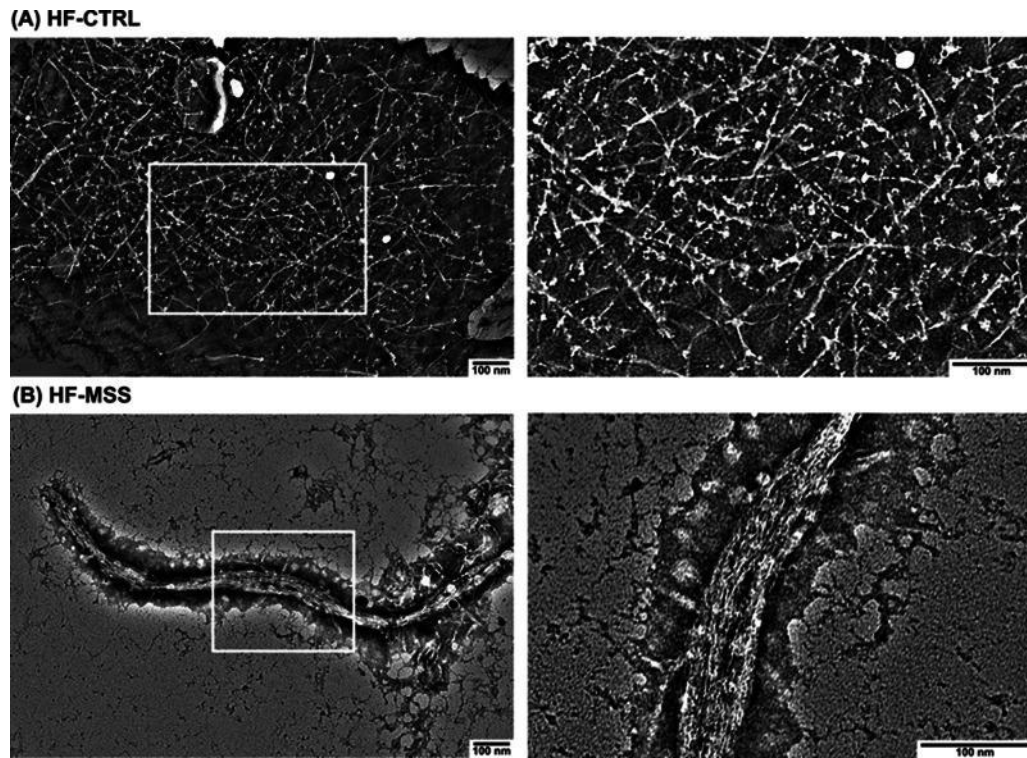


Figure 37. TEM analysis of extracellular matrix. Negative-stained ECM from control (A) and patient (B) fibroblasts, cultured in serum-free HBSS for 24 hours, was imaged by TEM (low and high magnification). From “*Sil1*-deficient fibroblasts generate an aberrant extracellular matrix leading to tendon disorganisation in Marinesco-Sjögren syndrome “Amodei L. et al. TEM experiments were conducted in collaboration with Matteo Ardini.

TEM, in collaboration with Laura Pietrangelo, also was used to examine the ultrastructure of the myotendinous junction (MTJ) in the *woozy* mouse model of MSS and compare it to control mice. Analysis of semi-thin sections stained with toluidine blue (Figure 38, panel A and C) revealed a clear difference in the organization of the connective tissue at the Achilles tendon-soleus muscle boundary. Control animals exhibited well-defined, darkly stained regions indicative of organized collagenous fibres with a distinct striation pattern (Figure 38, panels A and B). In contrast, the tendon matrix in *woozy* mice showed less organized fibrous

connective tissue, with collagen fibres appearing less compact, as evidenced by the irregular staining (Figure 38, panel C).

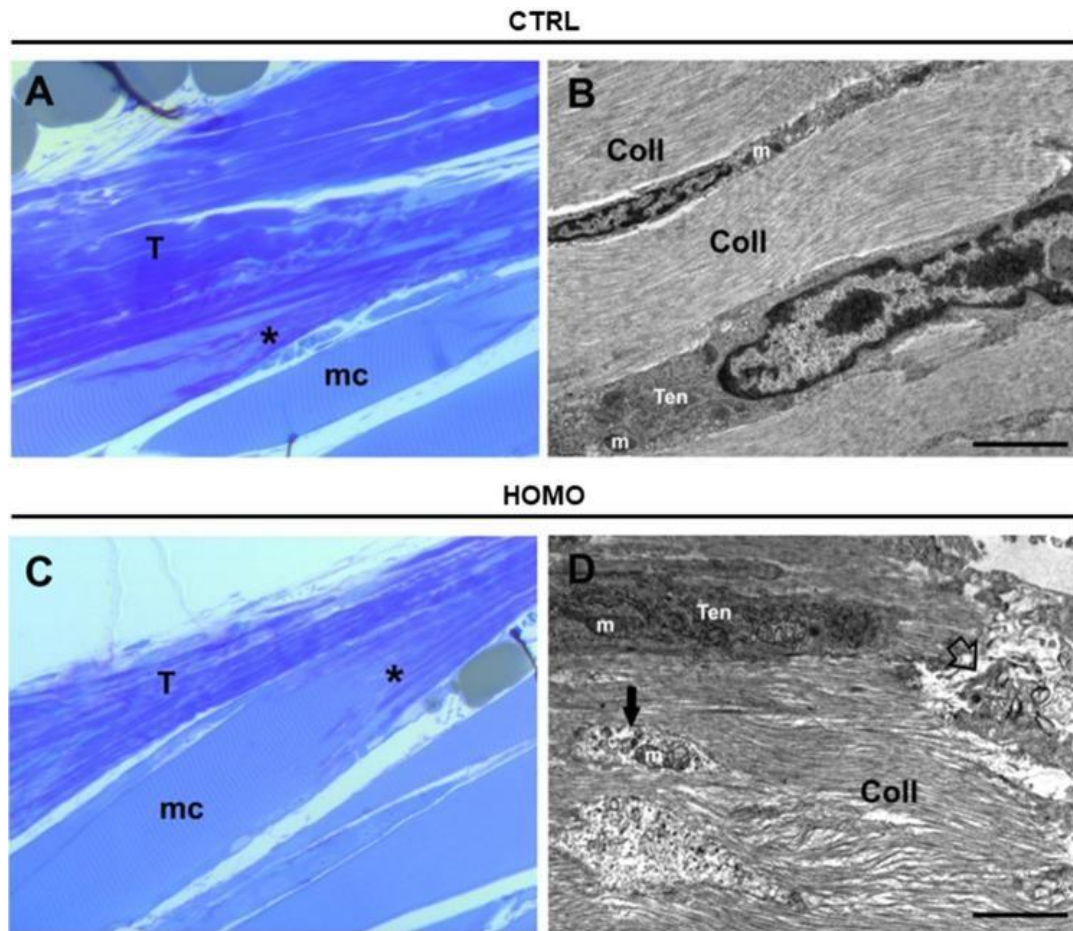


Figure 38. Ultrastructural analysis of the achilles tendon in control and woozy mice. (A, C) Toluidine blue-stained semi-thin sections of the Achilles tendon (T) and soleus muscle (mc) boundary from control (A) and woozy (C) mice. Asterisks indicate the myotendinous junction (MTJ). (B, D) TEM images of the connective tissue in control (B) and woozy (D) mice, showing tenocytes (Ten) among collagen fibrils (Coll). Mitochondria (m), cell debris (black arrow), and areas of degeneration (open arrow) are indicated. *From “Sill-deficient fibroblasts generate an aberrant extracellular matrix leading to tendon disorganisation in Marinesco-Sjögren syndrome” Amodei L. et al. TEM experiments were conducted in collaboration with Laura Pietrangelo.*

Ultrastructural analysis at higher magnification using TEM confirmed these observations. At the MTJ, control samples showed well-aligned collagen fibres (Figure 38, panel B), while woozy samples exhibited a more irregular collagen fibre pattern (Figure 38, panel D), indicative of disrupted ECM organization.

4.6.2 Baseline motor function in the Woozy Mouse

To establish a baseline for evaluating potential therapeutic interventions for MSS, we performed a comprehensive motor characterization of the *Sil1*^{wz} mouse model. Motor coordination and balance were assessed using accelerating rotarod and beam walking tests. The accelerating rotarod test revealed a significant and progressive decline in motor coordination and balance in *Sil1*^{wz} mice compared to control mice (Figure 39, panel A). Beginning at 9 weeks of age, *Sil1*^{wz} mice exhibited a statistically significant decrease in latency to fall ($p = 0.0112$). This decline in motor performance worsened progressively through week 14, at which point the average latency to fall for the *Sil1*^{wz} group was approximately one-third that of the control group. These results demonstrate a clear and progressive impairment of motor coordination and balance in the *Sil1*^{wz} mice.

The beam walking test corroborated the findings from the accelerating rotarod. *Sil1*^{wz} mice displayed a significantly increased time to traverse the beam, starting as early as 7 weeks of age compared to controls ($p = 0.0135$). This difference in beam traversal time continued to increase throughout the 14-week testing period (Figure 39, panel B, left). Analysis of contralateral foot slips during beam traversal also revealed a significant increase in the *Sil1*^{wz} group, starting from the first week of testing ($p = 0.0089$) (Figure 39, panel B, right). The increased beam traversal time and the elevated number of foot slips provide further evidence of impaired motor control and balance in the *Sil1*^{wz} mice.

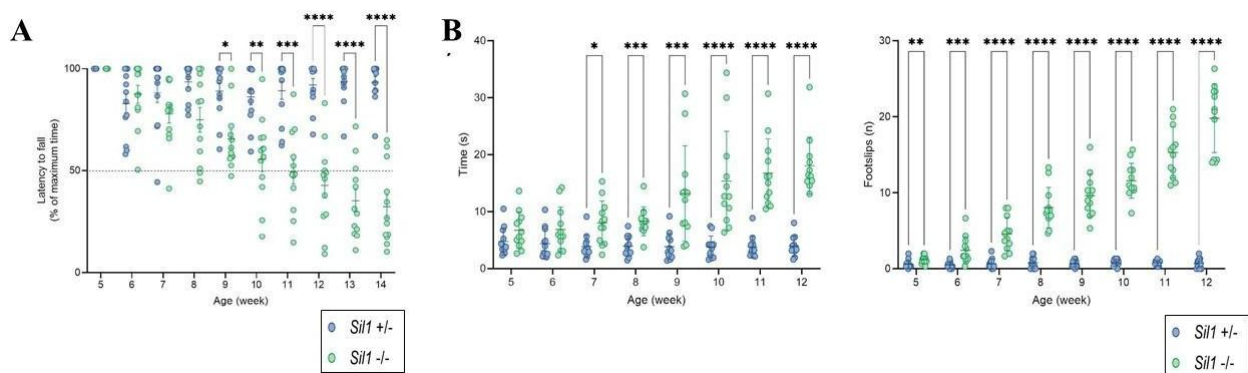


Figure 39. Progressive motor deficits in the *Sil1*^{wz} mouse model. (A) Accelerating rotarod performance (weeks 5-14), presented as a percentage of each individual's maximum time on the apparatus. (B) Beam walking test (weeks 5-12), showing mean beam traversal time and mean number of missteps ($n=3$ trials per mouse).

In summary, both the accelerating rotarod and beam walking tests consistently demonstrate progressive motor deficits in the *Sil1*^{wz} mice. These data provide a robust baseline for future studies aimed at evaluating the efficacy of potential therapeutic strategies for this neurodegenerative disease.

5. DISCUSSION

Marinesco-Sjögren syndrome (MSS), a recessive autosomal disease, results mainly from loss-of-function mutations in the SIL1 gene that lead to a deficiency of functional SIL1 protein⁵. The defect impairs its nucleotide exchange factor activity for BiP, a key ER chaperone, resulting in ER stress, UPR activation, and degeneration of cells, particularly cerebellar Purkinje cells and muscle tissue^{12,13}.

While current treatments are aimed at supportive care and symptom relief, they do not address the underlying cause of the disease¹. Protein replacement therapy (PRT) is a promising strategy that would directly compensate for the SIL1 protein deficiency, restoring normal BiP function, and avoiding the pathological cascade.

Clinical applications of PRT, and more specifically enzyme replacement therapy, have been highly successful in the treatment of genetic and metabolic diseases³¹. Furthermore, the evidence of the ability to produce active homologs of yeast Sil1 in bacterial expression systems has created a valuable precedent in considering *E. coli* as an expression system for the production of recombinant human SIL1 protein^{111,112}. As highlighted, the PRT is not without its weaknesses. Immunogenicity due to immune responses against the therapeutic protein and low tissue penetration are highly demanding^{36,39}.

5.1 Development of a robust purification protocol for SIL1: from structural characterization to therapeutic interventions for MSS

This research aimed to address the challenge of intracellular delivery of SIL1 by employing cell-penetrating peptides (CPPs), short fragments of a peptide capable of crossing a cell membrane^{39,40}. The HIV-1 transcriptional activator-derived peptide TAT was evaluated for its high cellular internalization capabilities^{46,48}. To counteract the sensitivity of TAT to enzymatic cleavage by proteases, its synthetic analogue TATk was examined, featuring targeted amino acid substitutions to abolish furin cleavage sites⁴⁹. In addition to CPPs, the potential to exploit endogenous cellular trafficking mechanisms, namely the KDEL receptor-mediated retrograde transport pathway, was examined in this work^{50,51}. This physiological cellular process suggests the potential to utilize exogenous KDEL-tagged proteins as a means of intracellular delivery. The high expression of the recombinant protein and the

extensive evidence for high internalization led us to choose to proceed with the CPP-containing fusion protein TAT.

A critical aspect of this study involved the optimization of a robust and scalable purification method for recombinant SIL1. Maximum protein expression levels were achieved through a comprehensive optimization process using several important parameters: the selection of an appropriate expression vector that offered tight transcriptional control, tight control of induction parameters such as concentration and induction time, application of measures to inhibit inclusion body formation, and elimination of bacterial impurities.

The initial purification of His-tagged SIL1-CPP presented significant challenges, prompting a strategic re-evaluation of the expression and purification methodology. To overcome these difficulties, improve protein yield and purification efficiency, the SIL1-CPP gene was subcloned into the pGEX6P-1 vector, resulting in a GST-tagged fusion protein. This transition was motivated by the well-established advantages associated with larger affinity tags such as GST. These advantages include increased binding to the affinity matrix, improved solubility and stability of the protein, and easier folding^{113–115}. Furthermore, the removal of the native signal peptide was based on fundamental differences in protein trafficking mechanisms between eukaryotic and prokaryotic cells. Human signal peptides are integral to directing nascent polypeptides to the endoplasmic reticulum and Golgi apparatus, organelles that are absent in *E. coli*¹¹⁶. Consequently, the presence of a human signal peptide in a protein expressed in *E. coli* would not only render it nonfunctional but could also lead to detrimental outcomes such as protein mis-localization, potentially leading to cytoplasmic aggregation or rapid proteolytic degradation.

The importance of a well-defined purification protocol was further underscored by the need to characterize the structural integrity and stability of SIL1-TAT. Through extensive biophysical characterization with techniques such as DLS and CD, we evaluated the structural stability and integrity of SIL1-TAT. The proper folding of SIL1-TAT was confirmed by CD spectroscopy, where it showed an α -helical secondary structure in agreement with our *in silico* predictions and with the deposited crystal structure of the yeast homologue of SIL1 (PDB code: 3QML)¹⁵. DLS confirmed that SIL1-TAT exists predominantly as a tetramer under solution conditions and has a Rh of 59.07 nm. Thermal denaturation measurements demonstrated significant thermal stability and reversibility of unfolding, indicating a well-folded and stable protein. Determination of the correct

secondary structure provides a background for determination of the protein's functional behavior and potential interactions¹¹⁷.

In addition, the lower production yield of SIL1-WT compared to SIL1-TAT confirmed that the TAT CPP is accountable for the overall stability and solubility of the recombinant SIL1 protein. The TAT CPP may promote the protein solubility and prevent its aggregation since it has a cationic nature^{118,119}.

The elucidation of the three-dimensional structure of SIL1 was recognized as a pivotal step in defining its functional mechanisms and assessing its therapeutic potential. However, the pursuit of high-resolution structural data encountered significant challenges related to protein crystallization. Initial crystallization trials, while yielding valuable preliminary insights, consistently produced crystals of suboptimal quality. These crystals exhibited characteristics such as limited size and poor morphology, resulting in diffraction data of insufficient resolution for comprehensive structural analysis. Under these conditions, the vapor diffusion method appears to be an appropriate technique for optimizing crystal growth. This method, which is optimal for managing supersaturation, allows for a controlled balance of crystallization conditions and promotes the development of larger crystals with a lower level of structural defects¹²⁰. Furthermore, optimization of crystallization conditions by adjusting some key factors such as precipitant concentration, solution pH and addition of additives is essential to manage nucleation and crystal growth rates¹²¹.

Despite initial challenges, a robust reproducible purification strategy has been developed for human SIL1 protein using a GST-tagged fusion protein and optimized expression parameters. As crystallization efforts continue, the development of a purification method that can be consistently applied to both fusion proteins and wild-type SIL1 recombinant protein, will facilitate the development of effective therapeutic interventions for MSS.

5.2 Preliminary evaluation of the safety and efficacy of SIL1: insights from *in vitro* and *in vivo* studies

The therapeutic potential of recombinant SIL1 protein required a comprehensive evaluation of its safety and efficacy. In this study, endotoxin level assessments were performed to ensure the purity of the protein and minimize the risk of adverse inflammatory responses. In addition, MTT assays were used to assess cell viability, providing crucial data on the limited

cytotoxicity of the protein. The preliminary *in vivo* treatment revealed promising results in terms of protein distribution in peripheral tissues, particularly spleen and muscle. The presence of CPP TAT probably facilitated this distribution by increasing the ability of the protein to cross cell membranes in this tissue^{42,119}. However, the limited penetration into the cerebellum highlights a persistent challenge in the delivery of therapeutic proteins to the central nervous system (CNS). This observation, coupled with the unexpected finding of predominantly extracellular localization of SIL1-TAT fusion protein in HF-MSS cells by immunofluorescence microscopy, prompts critical questions about the efficacy of our current delivery strategy.

The discrepancy between *in vivo* tissue distribution and *in vitro* cellular uptake underscores the fundamental differences between these experimental paradigms. *In vivo*, protein uptake is influenced by complex physiological processes, including blood circulation, lymphatic drainage, and tissue-specific barriers, whereas *in vitro* assays simplify these processes, often relying on direct cell exposure to the protein^{122,123}. In particular, the mechanism of TAT uptake *in vivo* is complex and influenced by the physiological environment, with possible differences in the administration route and tissue-specific factors³⁹. TAT-mediated delivery the Brain-blood barrier (BBB) appears to have yielded positive results in models of ischemia and seizures has shown promising results^{124,125}. A recent study have shown that using the TAT peptide, frataxin can be delivered to the brain, where it can combat the oxidative stress that contributes to Parkinson's disease, ultimately protecting neurons affected by the disease¹²⁶.

Furthermore, previous *in vitro* transcellular delivery studies have shown limitations in the crossing of cell layers by CPP due to entrapment in endosomes¹²⁷. In one study they investigated the mechanism of cellular uptake of a TAT-derived cell-penetrating peptide (CPP) into fibroblast cells, revealing a complex interplay of internalization pathways and the critical role of heparan sulphate (HS), a glycosaminoglycan found in the ECM. In essence, CPP utilizes multiple uptake pathways, with HS playing a central role in mediating both direct and aggregation-dependent internalization¹¹⁹.

The limitations of our study are obvious. Preliminary *in vivo* evaluation provides a snapshot of tissue distribution, but did not provide us with specific information on the tissue localization of the recombinant protein. This limitation highlights the need for more specific methods, such as immunohistochemistry and confocal microscopy, to detect precise tissue localization. In contrast, *in vitro* uptake suggests that the current SIL1-TAT construct may

require optimization. To improve intracellular delivery, several strategies could be explored. First, alternative CPPs with improved CNS penetration capabilities could be investigated¹²⁵. Second, liposomal encapsulation of SIL1 could be considered to protect proteins from degradation and facilitate cellular uptake via endocytosis¹²⁸. In particular, PEGylated liposomes have been shown to improve circulation time and enhance drug delivery to target tissues¹²⁹. These findings highlight the complexity of therapeutic protein delivery and underscore the need for a comprehensive approach integrating *in vitro* and *in vivo* studies to optimize treatment strategies for MSS.

5.3 Towards a complete characterization of Marinesco-Sjogren syndrome

In this study, we also aimed to investigate the molecular mechanisms underlying MSS, using both patient-derived cellular models and an animal model. Transcriptomic analysis of patient fibroblasts revealed significant perturbations in gene expression, particularly in membrane trafficking, ECM organization and cell adhesion. These data were further validated by TEM analyses, which revealed disorganized ECM structures in both patient fibroblasts and the mouse model. In muscle skeletal tissue, differential collagen gene expression is in agreement with the established function of the ECM in muscle homeostasis and regeneration^{130,131}. Similarly, disruption of the ECM has been implicated in cataract¹³² and skeletal malformation formation^{133,134}. We propose that aberrant ECM remodelling can be a critical pathological mechanism underlying the heterogeneity of MSS's clinical manifestations.

The observed discrepancy between SIL1-TAT uptake *in vitro* and *in vivo* may be related to the ECM abnormalities identified in patient fibroblasts and in *woozy* mouse. In particular, the impaired organization of the ECM in HF-MSS cells may hinder TAT-mediated internalization¹¹⁹. The ECM acts as a dynamic reservoir for growth factors and signalling molecules, and its integrity is critical for proper signalling and cellular uptake mechanisms¹³⁵.

Finally, to establish a functional baseline for therapeutic evaluation, motor function was assessed in the *woozy* mouse model, with significant and increasing motor impairment, thus confirming the value of this model for the evaluation of future therapies, such as SIL1-TAT administration. Future studies should incorporate additional tests, such as the pole test and the inverted grid test, to better evaluate motor parameters, including balance, coordination, and muscle strength^{136,137}. These tests have demonstrated high reliability and sensitivity in

assessing motor deficits in various neurological disorders. In addition to motor function, cognitive deficits are increasingly recognized as a significant component of neurodegenerative diseases, including MSS. Therefore, future studies should incorporate cognitive tests, such as the nesting test and the novel object recognition (NOR) test, to assess cognitive function in the woozy mouse model ^{138,139}. The nesting test evaluates nest-building behaviour, which is sensitive to cognitive decline, while the NOR test assesses recognition memory. These tests are widely used in preclinical studies of neurodegenerative diseases and can provide valuable insights into the cognitive impact of MSS.

6. CONCLUSION

This project aimed to develop a protein replacement therapy (PRT) strategy for Marinesco-Sjögren syndrome (MSS), a debilitating autosomal recessive disorder mainly caused by mutations in the SIL1 gene.

We developed a robust, scalable and reproducible purification protocol for the recombinant human SIL1 protein, using a GST-tagged fusion protein and optimized expression parameters. Comprehensive biophysical characterization confirmed the structural integrity and stability of SIL1-TAT, demonstrating its correct folding and tetrameric state in solution. Preliminary *in vitro* and *in vivo* studies provided valuable insights into the safety and distribution of SIL1-TAT. Furthermore, insights into the molecular pathogenesis of MSS revealed significant perturbations in gene expression in patient fibroblasts, particularly in membrane trafficking, ECM organization and cell adhesion. These findings suggest a crucial role for aberrant ECM remodelling in MSS pathogenesis.

Despite the remarkable progress made, several challenges remain. Continued optimization of crystallization conditions is essential to determine the three-dimensional structure of SIL1, which is critical to understand its function. Future studies should focus on evaluating the ability of SIL1-TAT to restore normal BiP function and alleviate ER stress in MSS cells and animal models. In this regard, it is crucial to address the limited penetration of SIL1-TAT into the cerebellum and the predominantly extracellular localization observed *in vitro*. Future efforts should explore alternative CPPs, liposomal encapsulation, and other strategies to improve delivery to the CNS.

By addressing these future goals, this research may lead to the development of effective therapeutic strategies for MSS and for other genetic diseases.

ABBREVIATIONS

AMP: ampicillin

ATF6: activating transcription factor 6

BBB: brain-blood barrier

BiP: immunoglobulin-binding protein

CD: Circular dichroism

CHL: chloramphenicol

CNS: central nervous system

CPP: cell-penetrating peptides

DE: differentially expressed

DLS: dynamic light scattering

DMSO: dimethyl sulfoxide

ECM: extracellular matrix

ER: endoplasmic reticulum

ERAD: endoplasmic reticulum-associated degradation

ERT: Enzyme replacement therapy

FDR: false discovery rate

FPLC: fast protein liquid chromatography

GO: gene ontology

GST: glutathione S-transferase

HF: human fibroblast

His: Histidine

HS: heparan sulphate

IEC: Ion exchange chromatography

IMAC: metal ion affinity chromatography

IPTG: Isopropil- β -D-1-thiogalactopyranoside

IRE: inositol-requiring enzyme 1

LB: Luria Bertani broth

MS: Mass spectrometry

MSS: Marinesco-Sjögren syndrome

MTJ: myotendinous junction

MTT: 3-(4,5-dimethylthiazol-2-yl)-2,5-diphenyltetrazolium bromide

MW: molecular weights

NEF: nucleotide exchange factor

NMR: nuclear magnetic resonance

PERK: endoplasmic reticulum kinase-like PRK

PRT: Protein replacement therapy

PTMs: post-translational modifications

qPCR: quantitative real-time PCR

Rh: hydrodynamic radius

SDS-PAGE: denaturing protein gel electrophoresis

SEC: Size-exclusion chromatography

TEM: transmission electron microscopy

TFE: trifluoroethanol

UPR: unfolded protein response

FUNDING

The PhD scholarship was co-financed with resources from the National Operational Programme for Research and Innovation 2014-2020, ESF REACT-EU resources Action IV.4 “Doctorates and research contracts on innovation topics” and Action IV.5 “Doctorates on Green topics”.

The Italian Telethon ONLUS Foundation, Rome, Italy, also financed this project with the grants N° GGP20092.

ACKNOWLEDGEMENTS

This PhD thesis was made possible thanks to the invaluable support and guidance of several people and institutions. I would like to express my sincere gratitude to Professor Luca Federici for his careful guidance and constructive feedback throughout the research process. I would also like to acknowledge the significant contribution of the biotech FYBROSIS and especially to Professor Vincenzo De Laurenzi.

I am deeply grateful to Professor Michele Sallese who diligently followed and contributed significantly to my training and development as a researcher. I would especially like to thank Professor Adriana Erica Miele, whose guidance, teachings, and encouragement were invaluable throughout my journey in Lyon.

REFERENCES

1. Anttonen, A.-K. Marinesco-Sjögren Syndrome. in *GeneReviews*® (eds. Adam, M. P. et al.) (University of Washington, Seattle, Seattle (WA), 1993).
2. MARINESCO, G. Nouvelle maladie familiale caractérisée par une cataracte congénitale et un arrêt du développement somato-neurophysique. *Encephale* **26**, 97-109 (1931).
3. Sjögren, T. Hereditary Congenital Spinocerebellar Ataxia Accompanied by Congenital Cataract and Oligophrenia: A Genetic and Clinical Investigation. *Confinia Neurologica* **10**, 293-308 (2009).
4. Anttonen, A.-K. et al. The gene disrupted in Marinesco-Sjögren syndrome encodes SIL1, an HSPA5 cochaperone. *Nat Genet* **37**, 1309-1311 (2005).
5. Senderek, J. et al. Mutations in SIL1 cause Marinesco-Sjögren syndrome, a cerebellar ataxia with cataract and myopathy. *Nat Genet* **37**, 1312-1314 (2005).
6. Krieger, M. et al. SIL1 mutations and clinical spectrum in patients with Marinesco-Sjögren syndrome. *Brain* **136**, 3634-3644 (2013).
7. Roos, A. et al. Myopathy in Marinesco-Sjögren syndrome links endoplasmic reticulum chaperone dysfunction to nuclear envelope pathology. *Acta Neuropathol* **127**, 761-777 (2014).
8. Skre, H. C Berg, K. Linkage studies on the Marinesco-Sjögren syndrome and hypergonadotropic hypogonadism. *Clinical Genetics* **11**, 57-66 (1977).
9. Brogdon, B. G., Snow, R. D. C Williams, J. P. Skeletal findings in Marinesco-Sjögren syndrome. *Skeletal Radiology* **25**, 461-465 (1996).
10. Alter, M., Talbert, O. R. C Croffead, G. Cerebellar ataxia, congenital cataracts, and retarded somatic and mental maturation. Report of cases of Marinesco-Sjogren syndrome. *Neurology* **12**, 836-847 (1962).
11. Zhao, L., Longo-Guess, C., Harris, B. S., Lee, J.-W. C Ackerman, S. L. Protein accumulation and neurodegeneration in the woozy mutant mouse is caused by disruption of SIL1, a cochaperone of BiP. *Nat Genet* **37**, 974-979 (2005).
12. Ichhaporia, V. P. et al. SIL1, the endoplasmic-reticulum-localized BiP co-chaperone, plays a crucial role in maintaining skeletal muscle proteostasis and physiology. *Dis Model Mech* **11**, dmm033043 (2018).
13. Behnke, J., Feige, M. J. C Hendershot, L. M. BiP and its Nucleotide Exchange Factors Grp170 and Sil1: Mechanisms of Action and Biological Functions. *J Mol Biol* **427**, 1589-1608 (2015).
14. Dudek, J. et al. Functions and pathologies of BiP and its interaction partners. *Cell. Mol. Life Sci.* **66**, 1556-1569 (2009).
15. Yan, M., Li, J. C Sha, B. Structural analysis of the Sil1-Bip complex reveals the mechanism for Sil1 to function as a nucleotide-exchange factor. *Biochemical Journal* **438**, 447-455 (2011).
16. Chiesa, R. C Sallese, M. Review: Protein misfolding diseases - the rare case of Marinesco-Sjögren syndrome. *Neuropathology and Applied Neurobiology* **46**, 323-343 (2020).

17. Walter, P. C Ron, D. The unfolded protein response: from stress pathway to homeostatic regulation. *Science* **334**, 1081-1086 (2011).
18. Signal integration in the endoplasmic reticulum unfolded protein response | Nature Reviews Molecular Cell Biology. <https://www.nature.com/articles/nrm2199>.
19. Hetz, C. The unfolded protein response: controlling cell fate decisions under ER stress and beyond. *Nat Rev Mol Cell Biol* **13**, 89-102 (2012).
20. Potenza, F. *et al.* Proteomic Analysis of Marinesco-Sjogren Syndrome Fibroblasts Indicates Pro-Survival Metabolic Adaptation to SIL1 Loss. *Int J Mol Sci* **22**, 12449 (2021).
21. Moreno, J. A. *et al.* Oral treatment targeting the unfolded protein response prevents neurodegeneration and clinical disease in prion-infected mice. *Sci Transl Med* **5**, 206ra138 (2013).
22. PERK inhibition delays neurodegeneration and improves motor function in a mouse model of Marinesco-Sjögren syndrome | Human Molecular Genetics | Oxford Academic. <https://academic.oup.com/hmg/article/27/14/2477/4990255?login=false>.
23. PERK inhibition attenuates the abnormalities of the secretory pathway and the increased apoptotic rate induced by SIL1 knockdown in HeLa cells - ScienceDirect. <https://www.sciencedirect.com/science/article/pii/S092544391830231X>.
24. Restelli, E., Masone, A., Sallese, M. C Chiesa, R. Neuroprotective modulation of the unfolded protein response in Marinesco-Sjögren syndrome: PERK signaling inhibition and beyond. *Neural Regeneration Research* **14**, 62 (2019).
25. Lavigna, G. *et al.* Trazodone, dibenzoylmethane and tauroursodeoxycholic acid do not prevent motor dysfunction and neurodegeneration in Marinesco-Sjögren syndrome mice. *PLoS One* **20**, e0317404 (2025).
26. Gámez, A. *et al.* Protein misfolding diseases: Prospects of pharmacological treatment. *Clinical Genetics* **G3**, 450-458 (2018).
27. Kashimada, A. *et al.* Targeting the enhanced ER stress response in Marinesco-Sjögren syndrome. *J Neurol Sci* **385**, 49-56 (2018).
28. Gulani, A. C Weiler, T. Genetics, Autosomal Recessive. in *StatPearls* (StatPearls Publishing, Treasure Island (FL), 2025).
29. Gorzelany, J. A. C de Souza, M. P. Protein Replacement Therapies for Rare Diseases: A Breeze for Regulatory Approval? *Science Translational Medicine* **5**, 178fs10-178fs10 (2013).
30. Sanchez-Garcia, L. *et al.* Recombinant pharmaceuticals from microbial cells: a 2015 update. *Microbial Cell Factories* **15**, 33 (2016).
31. Enzyme Replacement Therapy. in *LiverTox: Clinical and Research Information on Drug-Induced Liver Injury* (National Institute of Diabetes and Digestive and Kidney Diseases, Bethesda (MD), 2012).
32. Leader, B., Baca, Q. J. C Golan, D. E. Protein therapeutics: a summary and pharmacological classification. *Nat Rev Drug Discov* **7**, 21-39 (2008).

33. Johnson-Léger, C., Power, C. A., Shomade, G., Shaw, J. P. C Proudfoot, A. E. Protein therapeutics - lessons learned and a view of the future. *Expert Opinion on Biological Therapy* **6**, 1-7 (2006).
34. The biochemical basis of disease | Essays in Biochemistry | Portland Press. <https://portlandpress.com/essaysbiochem/article/62/5/619/78428/The-biochemical-basis-of-disease>.
35. Meena, N. K., Ralston, E., Raben, N. C Puertollano, R. Enzyme Replacement Therapy Can Reverse Pathogenic Cascade in Pompe Disease. *Molecular Therapy Methods & Clinical Development* **18**, 199-214 (2020).
36. Porter, S. Human Immune Response to Recombinant Human Proteins. *Journal of Pharmaceutical Sciences* **G0**, 1-11 (2001).
37. Marshall, S. A., Lazar, G. A., Chirino, A. J. C Desjarlais, J. R. Rational design and engineering of therapeutic proteins. *Drug Discovery Today* **8**, 212-221 (2003).
38. Chirino, A. J., Ary, M. L. C Marshall, S. A. Minimizing the immunogenicity of protein therapeutics. *Drug Discovery Today* **G**, 82-90 (2004).
39. Miliotou, A. N. *et al.* Protein Transduction Domain-Mediated Delivery of Recombinant Proteins and In Vitro Transcribed mRNAs for Protein Replacement Therapy of Human Severe Genetic Mitochondrial Disorders: The Case of Sco2 Deficiency. *Pharmaceutics* **15**, 286 (2023).
40. Järver, P. C Langel, Ü. Cell-penetrating peptides—A brief introduction. *Biochimica et Biophysica Acta (BBA) - Biomembranes* **1758**, 260-263 (2006).
41. Milletti, F. Cell-penetrating peptides: classes, origin, and current landscape. *Drug Discov Today* **17**, 850-860 (2012).
42. Guidotti, G., Brambilla, L. C Rossi, D. Cell-Penetrating Peptides: From Basic Research to Clinics. *Trends in Pharmacological Sciences* **38**, 406-424 (2017).
43. Madani, F., Lindberg, S., Langel, Ü., Futaki, S. C Gräslund, A. Mechanisms of Cellular Uptake of Cell-Penetrating Peptides. *J Biophys* **2011**, 414729 (2011).
44. Duchardt, F., Fotin-Mleczek, M., Schwarz, H., Fischer, R. C Brock, R. A Comprehensive Model for the Cellular Uptake of Cationic Cell-penetrating Peptides. *Traffic* **8**, 848-866 (2007).
45. Trabulo, S., Cardoso, A. L., Mano, M. C De Lima, M. C. P. Cell-Penetrating Peptides—Mechanisms of Cellular Uptake and Generation of Delivery Systems. *Pharmaceutics* **3**, 961-993 (2010).
46. Cellular uptake of the tat protein from human immunodeficiency virus: Cell. [https://www.cell.com/cell/abstract/0092-8674\(88\)90263-2?_returnURL=https%3A%2F%2Flinkinghub.elsevier.com%2Fretrieve%2Fpii%2F0092867488902632%3Fshowall%3Dtrue](https://www.cell.com/cell/abstract/0092-8674(88)90263-2?_returnURL=https%3A%2F%2Flinkinghub.elsevier.com%2Fretrieve%2Fpii%2F0092867488902632%3Fshowall%3Dtrue).
47. Gee Jun Tye, F. N. TATκ Fusion Protein of OCT-3/4 and KLF-4: Stable Mixed Population Cell Lines Capable of Delivering Fusion Proteins to Target Cells. *J Cell Sci Ther* **05**, (2014).
48. Flinterman, M. *et al.* Delivery of Therapeutic Proteins as Secretable TAT Fusion Products. *Molecular Therapy: the Journal of the American Society of Gene Therapy* **17**, 334 (2009).

49. Trazzi, S. *et al.* CDKL5 protein substitution therapy rescues neurological phenotypes of a mouse model of CDKL5 disorder. *Human Molecular Genetics* **27**, 1572-1592 (2018).
50. Jia, J. *et al.* KDEL receptor is a cell surface receptor that cycles between the plasma membrane and the Golgi via clathrin-mediated transport carriers. *Cell Mol Life Sci* **78**, 1085-1100 (2020).
51. Newstead, S. C Barr, F. Molecular basis for KDEL-mediated retrieval of escaped ER-resident proteins - SWEET talking the COPs. *Journal of Cell Science* **133**, jcs250100 (2020).
52. Bailey, L. An Overview of Enzyme Replacement Therapy for Lysosomal Storage Diseases. *Online J Issues Nurs* **13**, (2008).
53. Solomon, M. C Muro, S. Lysosomal enzyme replacement therapies: Historical development, clinical outcomes, and future perspectives. *Advanced Drug Delivery Reviews* **118**, 109-134 (2017).
54. Safary, A., Akbarzadeh Khiavi, M., Omid, Y. C Rafi, M. A. Targeted enzyme delivery systems in lysosomal disorders: an innovative form of therapy for mucopolysaccharidosis. *Cell. Mol. Life Sci.* **76**, 3363-3381 (2019).
55. Lillicrap, D. Future prospects for the treatment of hemophilia: A global perspective. *HemaSphere* **2**, 179-181 (2018).
56. Papaneophytou, C. Design of Experiments As a Tool for Optimization in Recombinant Protein Biotechnology: From Constructs to Crystals. *Mol Biotechnol* **61**, 873-891 (2019).
57. Brondyk, W. H. Chapter 11 Selecting an Appropriate Method for Expressing a Recombinant Protein. in *Methods in Enzymology* (eds. Burgess, R. R. C Deutscher, M. P.) vol. 463 131-147 (Academic Press, 2009).
58. Tripathi, N. K. C Shrivastava, A. Recent Developments in Bioprocessing of Recombinant Proteins: Expression Hosts and Process Development. *Front Bioeng Biotechnol* **7**, 420 (2019).
59. Mukhtar, H., Sania Sahreen, Sundas Sharif, C Hamna Ahmad. Comparison of Host Expression Systems used for Efficient Recombinant Proteins Production: Host Expression System for Recombinant Proteins Production. *PPASB* **60**, 5-28 (2023).
60. Kielkopf, C. L., Bauer, W. C Urbatsch, I. L. Expressing Cloned Genes for Protein Production, Purification, and Analysis. *Cold Spring Harb Protoc* **2021**, pdb.top102129 (2021).
61. Manufacture of Recombinant Therapeutic Proteins Using Chinese Hamster Ovary Cells in Large-Scale Bioreactors - Biosimilars of Monoclonal Antibodies - Wiley Online Library. <https://onlinelibrary.wiley.com/doi/10.1002/9781118940648.ch11>.
62. Better and faster: improvements and optimization for mammalian recombinant protein production - ScienceDirect. <https://www.sciencedirect.com/science/article/pii/S0959440X14000359?via%3Dihub>.
63. Pandey, A., Shin, K., Patterson, R. E., Liu, X.-Q. C Rainey, J. K. Current strategies for protein production and purification enabling membrane protein structural biology. *Biochem Cell Biol* **G4**, 507-527 (2016).

64. *Pichia pastoris* secretes recombinant proteins less efficiently than Chinese hamster ovary cells but allows higher space-time yields for less complex proteins - Maccani - 2014 - Biotechnology Journal - Wiley Online Library. <https://analyticalsciencejournals.onlinelibrary.wiley.com/doi/10.1002/biot.201300305>.
65. Karbalaeei, M., Rezaee, S. A. C Farsiani, H. *Pichia pastoris*: A highly successful expression system for optimal synthesis of heterologous proteins. *Journal of Cellular Physiology* **235**, 5867-5881 (2020).
66. Comparison of recombinant protein expression in a baculovirus system in insect cells (Sf9) and silkworm | The Journal of Biochemistry | Oxford Academic. <https://academic.oup.com/jb/article/149/2/219/838838>.
67. Strategies for optimization of heterologous protein expression in *E. coli*: Roadblocks and reinforcements - ScienceDirect. <https://www.sciencedirect.com/science/article/pii/S0141813017323826>.
68. Rosano, G. L. C Ceccarelli, E. A. Recombinant protein expression in *Escherichia coli*: advances and challenges. *Front Microbiol* **5**, 172 (2014).
69. Factors involved in heterologous expression of proteins in *E. coli* host | Archives of Microbiology. <https://link.springer.com/article/10.1007/s00203-023-03541-9>.
70. Sun, M. *et al.* High-throughput process development from gene cloning to protein production. *Microbial Cell Factories* **22**, 182 (2023).
71. Chumappumkal Joseph, B., Pichaimuthu, S. C Srimenakshi, S. An Overview of the Parameters for Recombinant Protein Expression in *Escherichia coli*. *Journal of Cell Science & Therapy* **06**, (2015).
72. Zhang, Z. *et al.* High-level production of membrane proteins in *E. coli* BL21(DE3) by omitting the inducer IPTG. *Microb Cell Fact* **14**, 142 (2015).
73. Jhamb, K. C Sahoo, D. K. Production of soluble recombinant proteins in *Escherichia coli*: Effects of process conditions and chaperone co-expression on cell growth and production of xylanase. *Bioresource Technology* **123**, 135-143 (2012).
74. Structural Genomics Consortium *et al.* Protein production and purification. *Nat Methods* **5**, 135-146 (2008).
75. De Brabander, P., Uitterhaegen, E., Delmulle, T., De Winter, K. C Soetaert, W. Challenges and progress towards industrial recombinant protein production in yeasts: A review. *Biotechnol Adv* **64**, 108121 (2023).
76. Francis, D. M. C Page, R. Strategies to Optimize Protein Expression in. *Current Protocols in Protein Science* **61**, 5.24.1-5.24.29 (2010).
77. Georgiou, G. C Valax, P. Expression of correctly folded proteins in *Escherichia coli*. *Curr Opin Biotechnol* **7**, 190-197 (1996).
78. Secretion of recombinant proteins from *E. coli* - Kleiner-Grote - 2018 - Engineering in Life Sciences - Wiley Online Library. <https://analyticalsciencejournals.onlinelibrary.wiley.com/doi/10.1002/elsc.201700200>.

79. Galloway, C. A., Sowden, M. P. C Smith, H. C. Increasing the yield of soluble recombinant protein expressed in *E. coli* by induction during late log phase. *Biotechniques* **34**, 524-526, 528, 530 (2003).
80. Oliveira, C., Felix, W., Moreira, R. A., Teixeira, J. A. C Domingues, L. Expression of frutalin, an α -D-galactose-binding jacalin-related lectin, in the yeast *Pichia pastoris*. *Protein Expression and Purification* **60**, 188-193 (2008).
81. Gomes, T. A., Zanette, C. M. C Spier, M. R. An overview of cell disruption methods for intracellular biomolecules recovery. *Preparative Biochemistry & Biotechnology* **50**, 635-654 (2020).
82. Perović, M. N., Knežević Jugović, Z. D. C Antov, M. G. Improved recovery of protein from soy grit by enzyme-assisted alkaline extraction. *Journal of Food Engineering* **276**, 109894 (2020).
83. An Introduction to Protein Purification: Methods, Technologies and Applications. *Analysis & Separations from Technology Networks* <http://www.technologynetworks.com/analysis/articles/an-introduction-to-protein-purification-methods-technologies-and-applications-388443>.
84. Young, C. L., Britton, Z. T. C Robinson, A. S. Recombinant protein expression and purification: A comprehensive review of affinity tags and microbial applications. *Biotechnology Journal* **7**, 620-634 (2012).
85. Protein Engineering Strategies for Selective Protein Purification - Hedhammar - 2005 - Chemical Engineering C Technology - Wiley Online Library. <https://onlinelibrary.wiley.com/doi/10.1002/ceat.200500144>.
86. Babaie, M. Proteins Separation and Purification Methods with Focus on Chromatography: a Review Study. *Journal of Ardabil University of Medical Sciences* **20**, 151-175 (2020).
87. Ion-Exchange Chromatography of Proteins | Shuichi Yamamoto, Kazahiro N. <https://www.taylorfrancis.com/books/mono/10.1201/b15751/ion-exchange-chromatography-proteins-shuichi-yamamoto-kazahiro-nakanishi-ryuichi-matsuno>.
88. Oliveira, C. C Domingues, L. Guidelines to reach high-quality purified recombinant proteins. *Appl Microbiol Biotechnol* **102**, 81-92 (2018).
89. Burgess-Brown, N., Chalk, R., Strain-Damerell, C. C Mahajan, P. Analysing recombinant proteins by mass spectrometry. **2**, 19-22 (2015).
90. Protein aggregation - Mechanisms, detection, and control - ScienceDirect. <https://www.sciencedirect.com/science/article/pii/S0378517318306215?via%3Dihub>.
91. Stetefeld, J., McKenna, S. A. C Patel, T. R. Dynamic light scattering: a practical guide and applications in biomedical sciences. *Biophys Rev* **8**, 409-427 (2016).
92. Yang, J. C Mattison, K. Protein characterization using dynamic and static light scattering. *China Particuology* **1**, 219-222 (2003).
93. Siligardi, G. C Hussain, R. CD spectroscopy: an essential tool for quality control of protein folding. *Methods Mol Biol* **1261**, 255-276 (2015).

94. Greenfield, N. J. Using circular dichroism spectra to estimate protein secondary structure. *Nat Protoc* **1**, 2876-2890 (2006).
95. Kwan, T. O. C. *et al.* Selection of Biophysical Methods for Characterisation of Membrane Proteins. *Int J Mol Sci* **20**, 2605 (2019).
96. Sreerama, N. C Woody, R. W. On the analysis of membrane protein circular dichroism spectra. *Protein Science* **13**, 100-112 (2004).
97. Gawas, U. B., Mandrekar, V. K. C Majik, M. S. Chapter 5 - Structural analysis of proteins using X-ray diffraction technique. in *Advances in Biological Science Research* (eds. Meena, S. N. C Naik, M. M.) 69-84 (Academic Press, 2019). doi:10.1016/B978-0-12-817497-5.00005-7.
98. Alberts, B. *et al.* Analyzing Protein Structure and Function. in *Molecular Biology of the Cell. 4th edition* (Garland Science, 2002).
99. Expasy - ProtParam. <https://web.expasy.org/protparam/>.
100. HisTrap HP His tag protein purification columns. Cytiva <https://www.cytivalifesciences.com/en/us/shop/chromatography/prepacked-columns/affinity-tagged-protein/hisrap-hp-histidine-tagged-protein-purification-columns-p-00250?srltid=AfmBOopEJNzxfUtvPE3mkvrOP4KgMVPXe5ZrHFMeQklrgC2PX2l66-zD>.
101. Glutathione Sepharose 4B GST-tagged protein purification resin. Cytiva <https://www.cytivalifesciences.com/en/us/shop/chromatography/resins/affinity-tagged-protein/glutathione-sepharose-4b-gst-tagged-protein-purification-resin-p-05999>.
102. Maity, R. *et al.* GST-His purification: A Two-step Affinity Purification Protocol Yielding Full-length Purified Proteins. *J Vis Exp* 50320 (2013) doi:10.3791/50320.
103. Bio-Rad Protein Assay | Bio-Rad. <https://www.bio-rad.com/it-it/product/bio-rad-protein-assay?ID=d4d4169a-12e8-4819-8b3e-ccab019c6e13>.
104. PierceTM BCA Protein Assay Kits. <https://www.thermofisher.com/order/catalog/product/23225>.
105. PierceTM Chromogenic Endotoxin Quant Kit. <https://www.thermofisher.com/order/catalog/product/A39552>.
106. Amodei, L. *et al.* Sil1-deficient fibroblasts generate an aberrant extracellular matrix leading to tendon disorganisation in Marinesco-Sjögren syndrome. *J Transl Med* **22**, 787 (2024).
107. Francis, D. M. C Page, R. Strategies to Optimize Protein Expression in *E. coli*. *CP Protein Science* **61**, (2010).
108. Wilkins, D. K. *et al.* Hydrodynamic radii of native and denatured proteins measured by pulse field gradient NMR techniques. *Biochemistry* **38**, 16424-16431 (1999).
109. Pereira, A. F., Piccoli, V. C Martínez, L. Trifluoroethanol direct interactions with protein backbones destabilize α -helices. *Journal of Molecular Liquids* **365**, 120209 (2022).
110. Malyala, P. C Singh, M. Endotoxin Limits in Formulations for Preclinical Research. *Journal of Pharmaceutical Sciences* **G7**, 2041-2044 (2008).
111. Rosam, M. *et al.* Bap (Sil1) regulates the molecular chaperone BiP by coupling release of nucleotide and substrate. *Nat Struct Mol Biol* **25**, 90-100 (2018).

112. Siegenthaler, K. D., Pareja, K. A., Wang, J. C Sevier, C. S. An unexpected role for the yeast nucleotide exchange factor Sil1 as a reductant acting on the molecular chaperone BiP. *eLife* **6**, e24141 (2017).
113. Malik, A. Protein fusion tags for efficient expression and purification of recombinant proteins in the periplasmic space of E. coli. *3 Biotech* **6**, 44 (2016).
114. Harper, S. C Speicher, D. W. Purification of Proteins Fused to Glutathione S-Transferase. in *Protein Chromatography: Methods and Protocols* (eds. Walls, D. C Loughran, S. T.) 259-280 (Humana Press, Totowa, NJ, 2011). doi:10.1007/978-1-60761-913-0_14.
115. Zhao, X., Li, G. C Liang, S. Several Affinity Tags Commonly Used in Chromatographic Purification. *Journal of Analytical Methods in Chemistry* **2013**, 581093 (2013).
116. G, von H. C L, A. Species-specific variation in signal peptide design. Implications for protein secretion in foreign hosts. *FEBS letters* **244**, (1989).
117. Linhares, L. A. C Ramos, C. H. I. Unlocking Insights into Folding, Structure, and Function of Proteins through Circular Dichroism Spectroscopy—A Short Review. *Applied Biosciences* **2**, 639-655 (2023).
118. Kramer, R. M., Shende, V. R., Motl, N., Pace, C. N. C Scholtz, J. M. Toward a Molecular Understanding of Protein Solubility: Increased Negative Surface Charge Correlates with Increased Solubility. *Biophys J* **102**, 1907-1915 (2012).
119. Ziegler, A., Nervi, P., Dürrenberger, M. C Seelig, J. The cationic cell-penetrating peptide CPP(TAT) derived from the HIV-1 protein TAT is rapidly transported into living fibroblasts: optical, biophysical, and metabolic evidence. *Biochemistry* **44**, 138-148 (2005).
120. Improving protein crystal quality by decoupling nucleation and growth in vapor diffusion | Protein Science | Cambridge Core. <https://www.cambridge.org/core/journals/protein-science/article/abs/improving-protein-crystal-quality-by-decoupling-nucleation-and-growth-in-vapor-diffusion/149A7DBEB921547CFE35B9415D4BB5CD>.
121. Optimization of crystallization conditions for biological macromolecules - PMC. <https://pmc.ncbi.nlm.nih.gov/articles/PMC4231845/>.
122. Albanese, A., Tang, P. S. C Chan, W. C. W. The effect of nanoparticle size, shape, and surface chemistry on biological systems. *Annu Rev Biomed Eng* **14**, 1-16 (2012).
123. Zhang, Y., Liu, A. T., Cornejo, Y. R., Van Haute, D. C Berlin, J. M. A Systematic comparison of in vitro cell uptake and in vivo biodistribution for three classes of gold nanoparticles with saturated PEG coatings. *PLoS One* **15**, e0234916 (2020).
124. Simon, M. J., Kang, W. H., Gao, S., Banta, S. C Morrison, B. TAT is not capable of transcellular delivery across an intact endothelial monolayer in vitro. *Ann Biomed Eng* **3G**, 394-401 (2011).
125. Zou, L.-L., Ma, J.-L., Wang, T., Yang, T.-B. C Liu, C.-B. Cell-penetrating Peptide-mediated therapeutic molecule delivery into the central nervous system. *Curr Neuropharmacol* **11**, 197-208 (2013).
126. Kim, M. J. *et al.* Tat-Frataxin protects dopaminergic neuronal cells against MPTP-induced toxicity in a mouse model of Parkinson's disease. *Biochimie* **G4**, 2448-2456 (2012).

127. El-Sayed, A., Futaki, S. C Harashima, H. Delivery of macromolecules using arginine-rich cell-penetrating peptides: ways to overcome endosomal entrapment. *AAPS J* **11**, 13-22 (2009).
128. Immordino, M. L., Dosio, F. C Cattel, L. Stealth liposomes: review of the basic science, rationale, and clinical applications, existing and potential. *Int J Nanomedicine* **1**, 297-315 (2006).
129. Allen, T. M. C Cullis, P. R. Liposomal drug delivery systems: from concept to clinical applications. *Adv Drug Deliv Rev* **65**, 36-48 (2013).
130. Barraza-Flores, P., Bates, C. R., Oliveira-Santos, A. C Burkin, D. J. Laminin and Integrin in LAMA2-Related Congenital Muscular Dystrophy: From Disease to Therapeutics. *Front. Mol. Neurosci.* **13**, (2020).
131. Kwong, A. K. *et al.* Collagen VI-related myopathies: clinical variability, phenotype-genotype correlation and exploratory transcriptome study. *Neuromuscul Disord* **33**, 371-381 (2023).
132. Wederell, E. D. C de longh, R. U. Extracellular matrix and integrin signaling in lens development and cataract. *Semin Cell Dev Biol* **17**, 759-776 (2006).
133. Gibson, B. G. C Briggs, M. D. The aggrecanopathies; an evolving phenotypic spectrum of human genetic skeletal diseases. *Orphanet Journal of Rare Diseases* **11**, 86 (2016).
134. Dubail, J. C Cormier-Daire, V. Chondrodysplasias With Multiple Dislocations Caused by Defects in Glycosaminoglycan Synthesis. *Front. Genet.* **12**, (2021).
135. Yue, B. Biology of the Extracellular Matrix: An Overview. *J Glaucoma* S20-S23 (2014) doi:10.1097/IJG.000000000000108.
136. Tillerson, J. L. C Miller, G. W. Grid performance test to measure behavioral impairment in the MPTP-treated-mouse model of parkinsonism. *J Neurosci Methods* **123**, 189-200 (2003).
137. Wada, M. *et al.* Behavioral characterization in MPTP/p mouse model of Parkinson's disease. *J Integr Neurosci* **20**, 307-320 (2021).
138. Jirkof, P. Burrowing and nest building behavior as indicators of well-being in mice. *Journal of Neuroscience Methods* **234**, 139-146 (2014).
139. Antunes, M. C Biala, G. The novel object recognition memory: neurobiology, test procedure, and its modifications. *Cogn Process* **13**, 93-110 (2012).