CZECH UNIVERSITY OF LIFE SCIENCES PRAGUE

FACULTY OF ENVIRONMENTAL SCIENCES

BACHELOR THESIS

2024 Daniel Wuthrich

CZECH UNIVERSITY OF LIFE SCIENCES PRAGUE

FACULTY OF ENVIRONMENTAL SCIENCES

DEPARTMENT OF ECOLOGY

Cloning and characterization of selected proteases from the Schistosoma mansoni egg transcriptome

BACHELOR THESIS

Supervisor: RNDr. Jan Dvořák, Ph.D.
Author: Daniel Wuthrich

2024

CZECH UNIVERSITY OF LIFE SCIENCES PRAGUE

Faculty of Environmental Sciences

BACHELOR THESIS ASSIGNMENT

Daniel Wuthrich

Environmental Engineering

Thesis title

Cloning and characterization of selected proteases from the Schistosoma mansoni egg transcriptome.

Objectives of thesis

Secreted proteins from blood flukes – schistosomes specifically modulate their hosts, and their detailed study will contribute to elucidate the processes during infection course and to understand host-parasite interactions at the molecular level. Schistosomiasis (bilharziosis) is a chronic infectious disease caused by blood flukes of the genus Schistosoma (Platyhelminthes, Trematoda). The disease is a global problem and around 230 million people are infected while another 600 million people are at risk. Bioactive molecules secreted by parasites are critical for their survival in mammalian hosts. In the case of schistosome eggs, their migration in tissues depends on the secretion of their E/S products stimulating host physiological processes. Several mechanisms are known to allow mature eggs to pass from the blood vessels through the intestinal wall, such as angiogenesis, endothelial activation, interaction with coagulation processes, and a variety of immunomodulatory processes promoting the formation of granulomatous tissue around the eggs. These host-pathological reactions caused by mature eggs are responsible for the typical manifestations of schistosomiasis. Several proteolytic enzymes identified from the tissue migrating eggs will be cloned and expressed in the bacterial expression system. Leishmanolysin-like protease, Endothelin-converting enzyme, Aminopeptidase-N will be preferred genes of interest.

Methodology

Selected genes will be amplified from parasite RNA and converted to DNA. Genes will be cloned to expression bacterial plasmids, and sequenced for their integrity. Subsequently, T7 polymerase-modified bacterial strains will be used for recombinant expression. Successfully produced proteases will be purified and available for the following biological and biochemical studies.

The proposed extent of the thesis

30 pages

Keywords

Fluke; Host-parasite interactions; Schistosoma; Proteases.

Recommended information sources

Fajtová P, Stefanic S, Hradílek M, Dvořák J, Vondrásek J, Jilková A, Ulrychová L, McKerrow JH, Caffrey CR, Mareš M, Horn M. Prolyl Oligopeptidase from the Blood Fluke Schistosoma mansoni: from Functional Analysis to Anti-schistosomal Inhibitors. PLoS Neglected Tropical Diseases. 2015; 9(6): e0003827. DOI: 10.1371/journal.pntd.0003827

Jedličková L, Dvořák J , Hrachovinová I, Ulrychová L, Kašný M, Mikeš M, A novel Kunitz protein with proposed dual function from Eudiplozoon nipponicum (Monogenea) impairs haemostasis and action of complement in vitro. International Journal for Parasitology, 2019 Feb 20. pii: S0020-7519(19)30026-8. DOI: 10.1016/j.ijpara.2018.11.010

Jedličková L, Dvořáková H, Dvořák J, Kašný M, Ulrychová L, Vorel J, Žárský V, Mikeš L. Cysteine peptidases of Eudiplozoon nipponicum: a broad repertoire of structurally assorted cathepsins L in contrast to the scarcity of cathepsins B in an invasive species of haematophagous monogenean of common carp. Parasites & Vectors. 2018 Mar 6;11(1):142. DOI: 10.1186/s13071-018-2666-2

Leontovyč A., Ulrychová, L., O'Donoghue, A., Mareš, M., Vondrášek J., Marešová L., Hubálek M., Fajtová P., Chanová M., Jiang Z., Craik CS., Caffrey CR., Dvořák, J.*, Horn*, M. SmSP2: a serine protease secreted by the blood fluke pathogen Schistosoma mansoni with anti-hemostatic properties. PLoS Neglected Tropical Diseases. 2018 Apr 20;12(4):e0006446. (*shared senior authorship) DOI: 10.1371/journal.pntd.0006446

Peterkova K., Vorel J., Ilgova J., Ostasov P., Fajtova P., Kasny M., Konecny L., Chanova M., Kasny M., Horn M., Dvorak J. Proteases and their inhibitors involved in Schistosoma mansoni egg-host interaction revealed by comparative transcriptomics with Fasciola hepatica eggs. International Journal for Parasitology. 2023 May;53(5-6):253-263. doi: 10.1016/j.ijpara.2022.12.007. Epub 2023 Feb 6.

Expected date of thesis defence

2023/24 SS – FES

The Bachelor Thesis Supervisor

RNDr. Jan Dvořák, Ph.D.

Supervising department

Department of Ecology

Electronic approval: 4. 3. 2024

prof. Mgr. Bohumil Mandák, Ph.D.

Head of department

Electronic approval: 25. 3. 2024

prof. RNDr. Michael Komárek, Ph.D.

Dean

Prague on 27. 03. 2024

AUTHOR STATEMENT

I hereby declare that I have independently elaborated the bachelor/final thesis with the topic of: .....
and that I have cited all of the information sources that I used in the thesis as listed at the end of the thesis in the list of used information sources.

I am aware that my bachelor/final thesis is subject to Act No. 121/2000 Coll., on copyright, on rights related to copyright and on amendments of certain acts, as amended by later regulations, particularly the provisions of Section 35(3) of the act on the use of the thesis.

I am aware that by submitting the bachelor/final thesis I agree with its publication under Act No. 111/1998 Coll., on universities and on the change and amendments of certain acts, as amended, regardless of the result of its defence.

With my own signature, I also declare that the electronic version is identical to the printed version and the data stated in the thesis has been processed in relation to the GDPR.

In ..... on .....

.....

ACKNOWLEDGMENT

I want to express my sincere gratitude to RNDr. Jan Dvořák, Ph.D., for giving me hands-on experience in molecular biology by offering this thesis work.

I also want to express my special appreciation and gratitude to Mgr. Vojtech Vacek for his immaculate guidance and for sharing his expertise throughout the entire thesis completion process. I genuinely enjoyed working in the lab.

Lastly, I would like to thank Selam Muguleta for sharing his tips and tricks on protein-related topics.

ABSTRAKT

Schistosomiasis je rozšířená parazitní infekce, která postihuje více než 210 milionů jedinců po celém světě, převážně v oblastech s nedostatečnou sanitací a omezeným přístupem k čisté vodě, zejména v tropických a subtropických oblastech. Nemoc je způsobena trematody rodu Schistosoma, přičemž vajíčka jsou hlavními viníky. I přes jejich významnost, jejich molekulární vlastnosti proteinů kódovaných v těchto vajíčkách zůstávají nedostatečně pochopeny. Rozjasnění interakce vajíčka-hostitel vyžaduje podrobnější porozumění těmto proteinům, což může pomoci při vývoji vakcín proti Schistosomóze. Tento výzkum se zaměřuje na vyjádření a čištění následujících proteáz: aminopeptidáza-N 1 (APN 1), endotelin-konvertující enzym 1 (ECE 1), leishmanolysin-podobná peptidáza (SmLeish), které jsou nejvíce exprimovanými proteázami identifikovanými ve vajíčkách *Schistosoma mansoni*. Cílem je příprava proteáz pro další biochemická studia. K dosažení tohoto cíle jsme extrahovali RNA z vajíček a převedli ji na cDNA, která byla poté amplifikována pomocí PCR a ligována do plazmidových vektorů pro transfekci do kompetentních buněk. Během procesu jsme sledovali přítomnost zájmových transkriptů pomocí sekvencování, zarovnání s BLAST a gelové elektroforézy. Po indukci exprese proteinů jsme analyzovali proteiny pomocí SDS-PAGE. Nicméně, podařilo se nám dostat pouze SmLeish do fáze exprese, přičemž jsme získali neuspokojivé výsledky z SDS-PAGE, což nám zabránilo čištění SmLeish. Pokud jde o APN 1 a ECE 1, narazili jsme na problémy během ligace do vektoru pET-SUMO a rozhodli jsme se neopakovat proces z důvodu časového omezení. Tato zjištění však odhalila obtíže při čištění a pracování s takovými proteázami, což může poskytnout cenné poznatky pro vývoj léčebných postupů proti schistosomóze.

KLÍČOVÁ SLOVA

Ploštěnci; Hostitel/Parazit interakce; Aminopeptidáza-N 1; Endotelin konvertující enzym 1; Leishmanolysin-podobná peptidáza

ABSTRACT

Schistosomiasis is a prevalent parasitic infection that affects over 210 million individuals worldwide, primarily in regions characterized by inadequate sanitation and limited access to clean water, particularly in tropical and subtropical areas. The disease is caused by the trematodes of the Schistosoma genus, with the eggs being the primary culprits. However, despite their significance, the molecular properties of the proteins encoded in these eggs remain inadequately understood. Shedding light on the egg-host interaction requires a more comprehensive understanding of these proteins, which may aid in developing vaccines against schistosomiasis. This paper focuses on expressing and purifying the following proteases: aminopeptidase-N 1 (APN 1), endothelin-converting enzyme 1 (ECE 1), leishmanolysin-like peptidase (SmLeish), which are the most highly expressed proteases identified in *Schistosoma mansoni* eggs. We intend to make them accessible for further biochemical studies. To achieve this, we extracted RNA from the eggs and converted it to cDNA, which was then amplified using PCR and ligated into plasmid vectors for transfection into competent cells. Throughout the process, we monitored the presence of the transcripts of interest via sequencing, alignment with BLAST, and gel electrophoresis. After inducing protein expression, we analyzed the proteins using SDS-PAGE. However, we managed to get only SmLeish into the expression stage, but we were getting unsatisfactory results from the SDS-PAGE, preventing us from purifying SmLeish. As for APN 1 and ECE 1, we encountered issues during the ligation into pET-SUMO vector and decided not to repeat the process due to time constraints. Nevertheless, these findings unveiled the difficulties of purifying and working with such proteases, which can provide valuable insights into developing treatments for schistosomiasis.

KEYWORDS

Fluke; Host-parasite interaction; Aminopeptidase-N 1; Endothelin converting enzyme 1; Leishmanolysin-like peptidase

Table of Contents

1. Introduction

Schistosoma is a genus of helminths that causes schistosomiasis, a parasitic infection affecting over 210 million people globally in tropical and subtropical regions, particularly in impoverished areas with a lack of access to safe drinking water and proper sanitation facilities (Grevelding, 2004; Nelwan, 2019; *Schistosomiasis*, n.d.). There are a few schistosomiasis-causing trematode species around the globe, but this paper will focus solely on *Schistosoma mansoni*. *S. mansoni* is the most common cause of schistosomiasis in human hosts due to successfully adapting and evolving alongside humans since the 16th century (Crellen et al., 2016). The trematode itself is not harmful to humans; the eggs are primarily responsible for the detrimental effects of schistosomiasis (Grevelding, 2004; Peterkova et al., 2023).

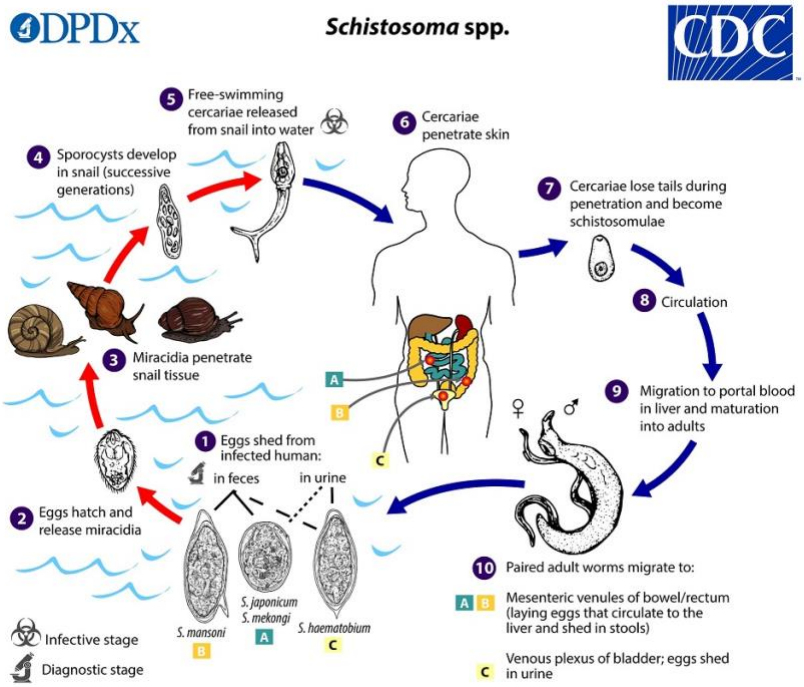
There are two main types of schistosomiasis: intestinal and urogenital. In cases of intestinal schistosomiasis, individuals may experience symptoms such as abdominal pain, diarrhea, and blood in the stool (Nelwan, 2019). As the condition worsens, liver enlargement often occurs, accompanied by fluid buildup in the abdominal cavity and hypertension in the blood vessels of the abdomen (*Schistosomiasis*, n.d.). The primary indication of urogenital schistosomiasis is the presence of blood in the urine; advanced cases may lead to kidney damage, bladder fibrosis (thickening and stiffening of the bladder), and bladder cancer may arise in later stages (*Schistosomiasis*, n.d.; Schwartz & Fallon, 2018). In females, symptoms may include vaginal bleeding, painful intercourse, and genital lesions (*Schistosomiasis*, n.d.). In males, urogenital schistosomiasis can affect the prostate and other organs, which could potentially lead to infertility (*Schistosomiasis*, n.d.). It is essential to mention that schistosomiasis can be deadly, causing over 11,000 deaths per year globally (*Schistosomiasis*, n.d.).

1.1 Life cycle and embryotic development of *S. mansoni*

As depicted by the CDC in Figure 1, the life cycle of *S. mansoni* begins with the eggs being excreted from infected mammals. Under good conditions in freshwater, the eggs hatch into miracidia, which infects intermediate host - water snails. In snails, miracidia change into sporocysts and reproduce asexually to generate numerous

cercariae (Grevelding, 2004; Wright, 1967). Once released into the water, the cercariae infects mammals, its definitive host, by penetrating through the skin, which causes the cercaria to lose its tail and become schistosomula (*CDC - DPDx - Schistosomiasis Infection*, 2019). In the human body, the schistosomula spends several days in the skin before traveling through the bloodstream into the lungs (Grevelding, 2004). Over two weeks later, they travel to the hepatoportal system (blood vessels that link the gastrointestinal tract to the liver), where they mature into adult male or female trematodes upon encountering a mate of the opposite sex (Schwartz & Fallon, 2018). Once paired, they move to the mesenteric veins (blood vessels draining blood from the intestines, pancreas, and spleen, directing it towards the liver), where they mate and begin producing eggs ~28 days later (Schwartz & Fallon, 2018). The female trematodes lay eggs on the inner cellular lining of capillaries, from where the eggs can travel through the bloodstream to other organs or pass through the intestinal epithelium into the intestinal lumen (Schwartz & Fallon, 2018). Once mature, the eggs form granulomas necessary to pass through the intestinal tissues (Schwartz & Fallon, 2018; Takaki et al., 2021). Eventually, the rest of the eggs are excreted in feces, which completes the *S. mansoni* life cycle.

Figure 1 Schistosoma mansoni life cycle



Junger et al. (2009) proposed eight stages for the embryonic development of *S. mansoni* eggs, from which it may be generally accepted that stages 1 to 6 are associated with immature eggs and stages 7 to 8 with mature eggs (Peterkova et al., n.d., 2023; Sarvel et al., 2006). Immature eggs are morphologically smaller, involving blastomeres' formation and miracidia's total development (Jurberg et al., 2009; Peterkova et al., n.d.). On the other hand, mature eggs are morphologically bigger, entailing the transformation of miracidia into cercaria, which contributes to the life cycle by infecting its definitive host (Jurberg et al., 2009; Peterkova et al., n.d.). About 22% of those eggs continue the life cycle, and the other 78% are lodged inside the definitive host, triggering the known symptoms of schistosomiasis (Mooee & Sandgeound, 1957). However, molecularly, the proteins accountable for the previously mentioned processes are not fully understood (Qokoyi et al., 2021).

2. Objective of the thesis

Understanding the roles of proteins released by *S. mansoni* eggs can aid in identifying potential drug targets and advancing vaccine development efforts against schistosomiasis, as they are the primary pathogenic agent (Qokoyi et al., 2021). The infiltration of eggs into the human body entails a range of protein interactions across species. They produce proteins not only to aid in overcoming the skin’s barriers, both physical and biochemical, but also to circumvent the immune system, enabling their survival within the host for extended periods (Knudsen et al., 2005). However, only a limited set of these protein interactions has been thoroughly studied. Therefore, understanding the molecular mechanisms behind the egg-host interaction is crucial for advancing novel treatment strategies.

Therefore, we decided to express three proteolytic enzymes, the most expressed proteases from *S. mansoni* eggs, using a bacterial expression system. The proteases include aminopeptidase-N 1 (APN 1), endothelin converting enzyme 1 (ECE 1), and leishmanolysin-like protease (SmLeish). The successfully produced proteases will be purified and made accessible for subsequent biological and biochemical investigations.

3. Proteases of interest

According to Peterkova et al. (2023), amongst the most expressed proteases in *S. mansoni* eggs are aminopeptidase-N 1, endothelin converting enzyme 1, and leishmanolysin-like protease, which are the proteases of interest in this paper.

3.1 Aminopeptidase-N 1

In studies conducted on humans, mice, and other model organisms, APN 1 has been associated with angiogenesis (the development of new blood vessels from existing ones) (Karamysheva, 2008; Sato, 2004). Despite this association, its precise function within *S. mansoni* is poorly understood. Additional research on APN 1 is needed to shed light on its influences on parasite growth and host-parasite interactions.

3.2 Endothelin converting enzyme 1

Originally isolated from human cells, ECE 1 is one of the most potent vasoconstrictive agents (Johnson et al., 2002; Kawanabe & Nauli, 2011). Based on those findings, ECE 1, present in *S. mansoni*, may be expected to execute parallel functions regarding vasoconstriction. However, more research needs to be conducted on the role and mechanism of ECE 1 across the life cycle of *S. mansoni* to fully understand its impact on vasoconstriction and other implications for parasite biology and pathogenicity.

3.3 Leishmanolysin-like protease

SmLeish is a metalloprotease that shares amino acids and exhibits structural resemblances to leishmanolysin, a metalloprotease enzyme originally found in Leishmania parasites (Hambrook et al., 2018; Kulkarni et al., 2008; Olivier et al., 2012; Schlagenhauf et al., 1998). SmLeish has been analyzed by Hambrook et al. (2018) using proteomic and transcriptomic methods on schistosomes across their various life stages and has revealed that this metalloprotease is abundantly produced throughout the life cycle of *S. mansoni*, with the highest concentrations occurring during the miracidia and cercaria stages. An example of this is the release of SmLeish through secretion generated by sporocysts during snail infection. Its presence enhances survival within intermediate hosts by assisting sporocyst encapsulation (Vondráček et al., 2022). According to studies done on *S. mansoni* and other parasites, SmLeish may facilitate host cell invasion,

modulate the host immune response, promote parasite survival, as well as assisting the degradation of extracellular matrix, which promotes migration (Knudsen et al., 2005; Kulkarni et al., 2008; Olivier et al., 2012; Schlagenhauf et al., 1998). Despite being more analyzed than APN 1 and ECE 1, SmLeish still needs more research to characterize the mentioned functions fully.

4. Methodology

This section will describe the methodological process of amplifying and purifying APN 1, ECE 1, and SmLeish transcripts from *S. mansoni* eggs using bacterial expression systems. The following steps were identically conducted for all three transcripts. The only things that will differ for each transcript are the primers and template DNA used, which are specific to each transcript.

4.1 Primer design

Primers are short DNA pieces that commence DNA replication by DNA polymerase. For PCR amplification of transcripts of interest, the designed forward (FWD) and reverse (REV) primers must be specific to the transcript. By tailoring the primers to the target sequence, we can produce an accurate transcript suitable for downstream applications. Table 1 illustrates the composition of the primers for the transcripts of interest. Table 2 displays the primers of both the transcripts and the employed plasmid vectors. Primers were designed especially for cloning into pET-SUMO vector, which allows TA cloning. The coding sequences of APN 1, ECE 1, and SmLeish are shown in Figure 2, Figure 3, and Figure 4, respectively.

Table 1 Composition of primers

Primers	Base pairs	GC content	Melting temperature
APN 1 FWD	22	40.9%	53.2 ℃
APN 1 REV	26	34.6%	51.8 ℃
ECE 1 FWD	33	21.2 %	52.3 ℃
ECE 1 REV	27	33.3 %	53.3 ℃
SmLeish FWD	21	38.1 %	53.2 ℃
SmLeish REV	24	33.3 %	53.3 ℃

Table 2 Sequence of primers

Primers	Sequence
APN 1 FWD	5’ – ATG TTG GCT GTG GAT AAT TCT G – 3’
APN 1 REV	5’ – CTA TAA TAA ATG GGT AAC TAC TGG GA – 3’
ECE 1 FWD	5’ – ATG ACT GAT TGT CAT CAA TAT ATT AAT AAT GAT – 3’
ECE 1 REV	5’ – TTA CCA TAC ATG ACA TTT CTT TAC AGG – 3’
SmLeish FWD	5’ – ATG CAA CAT GAA ACA GCT CAT – 3’
SmLeish REV	5’ – TCA TTT TGG ACA TTT ATC CTT GCA – 3’
pJET1.2 FWD	5’ – CGA CTC ACT ATA GGG AGA GCG GC – 3’
pJET1.2 REV	5’ – AAG AAC ATC GA TTT TCC ATG GCA G – 3’
pET-SUMO FWD	5’ – AGA TTC TTG TAC GAC GGT ATT AG – 3’
pET-SUMO REV	5’ – TAG TTA TTG CTC AGC GGT GG – 3’

Figure 2Figure 3Figure 4Figure 2 APN 1 transcript (2958 bp) with underlined FWD (top) and REV (bottom) primer sequence

5'**ATGTTGGCTGTGGATAATTCTG**ATGGAAGATACTTTGATTCCACAGATCACTGTCAATGTACCCACTCAATGCATGCA
AAAAATTTAACAAACAAAACTTTAATACAGTCAAAGAAATCATTTATCATTTTGATTGTGTTGACTCTCGAATTAATATTT
GTTGACCATGCTTCAGCAGTGAGCAGTTCTAGTTATTTCGAAAATGGTCCAGTGTTCAAAAATATATTAACTACTAGCGC
GACTGAAAATATCGATGATGAAAATAACATGTCGAATTATTCATCAGAAGATTTTAGATTACCACACACATTGTTTCCACA
TTTCTACGATTTGTCAATTCAAGTCCATTTGCATGATAACAAGTCACAAGCATTTTTCTTTAATGGATCTGTGACAATAAA
AGTTTTCTGTAATGTATCGACAACCGAAATCTTCGTCCATGCATACAAAACTCTGGATGTGAAATTGGATAAAATCCAAA
TGGTTCCTCTTCATGAACGAAATCAAACAAGTTCTATTATTGAAATAGAAGACATAAATTATTTTCAAGATGCTGAGTGTT
ATCGAATAAAACTCAAAATGCCTTTGCGACCAAACACTTATTACAACCTGACATTTGGACAGTTTTGTTCTGATCTAAAC
TATGTTTCTGAAGGACTATATCTTAGCAGATATTTGGAAAATGGAGTTTACGAATATTTGGCAAGTACTCAATTGGAAGCA
AACTATGCTAGGCGTGTATTTCCATGTTGGGATGAGCCGGAATTTAAAGCGAAATTCAAGGTCAATATTATACGTCACAA
AAGTTTCCATTCGCTTTCGAACATGAATTTGGAAAGTACGAAAGTGTTGTACGATAATTGGTGTCTCGACACATATAACA
CTAGTGTGAAAATGTCCACATATTTGTTGGCAATTGTAGTTTCCCGTTTTTCAAATATTCGGAGGACAGACAATAGAGGA
AGAAACTTTACTGTTTGGGCAAGACCAGAAAAAATTCGATCAGCCAAATACGCACTTGATGTTGGCATAAAATTATTGG
AACACTTCGAAGATTATTTTGGTATCCCCTATTCACTTCATAAAATGGATATGATCGCGATTCCTAACTCTTCAATAACGG
CAATGGAAAACTGGGGTTTGATAACTTTTAGAGAAAATTTAATGTTATGGAATCCAGAAAACGGTTCAATAGCCTCCCC
AATTGATGTGGCTTCCACTGTTTCACATGAACTTTCTCATCAGTGGTTCGGAAACATAGTGACAATGAAATGGTGGGATA
ATTTGTGGTTGAACGAAGGCTTCGCAACGTTTATGGAATACATTGGAGTACAGTCATTACATCCTGAGTGGAAAGTAGAT
GAACTATTCATATTGGACGAATTAATTCAAGTACTTCGTAGTGATACTTCGAAGAAATCAAGACCAGTTATTTTTCCAGCC
AATACTACGACACAAATTAAAAGAATGTTTGATTTAATCACCTACAGTAAGGGTGCCTCCTTGATTTGGATGATGGAGAA
ATTCATGGGCCGTAGTGCATTTCAAAACGGCTTAAAAAAATATTTAATTCAAAATCAATATAAGAACACAAATGAAAAGG
ATTTATGGAAAGCATTATCTAACGAATGGAATACTCAAGGAAATCATTTAGACATTGAATTCATTATGGATTCTTGGACTA
AACAAATGAATTATCCACTAGTAATTGTAAGAAGACATGGGTCAAACATGTTCTGTTTTGAGCAGACACGTTACCTTCAA
AACTGTGATAACAATTCACTTCCAGAGAATCATGAATACTCCTGGATAATACCAATCACTTATGGGTCAGCGAAAACGGT
GAACTGGACAGATGCTGATATTTTGTGGATGGTGAACAAGACAATGAATCAAACAATGAACATAAGTTCAGACGACTGG
TATTTGTTAAATGTAAGACAAGAAGGATTATATCGAGTTCATTATGCCGACAATAATTGGAAATTATTAGTCAATCAATTG
CAACGGAACTTCATGGCTATTCCAGTTTACTCACGCACACAAATCTTGAATGATTTATTCAGCCTTGCAAATCATCATATC
GTACCGTATACTCTTTTCCTGAACGCGACAAAGTATTTGGACCGTGAAAATCAGTTTATTGTATGGATAACAAGCAGCCG
AGCGTTTCGATACATTAACAGCATGTTTGTTTTGAATGAGAATTATGACGTTTATCAAGCATATTTGCGAACCCTCCTTGA
TAATCAAATTCGATCAGTGAATTGGTCATTTGTTAGTGAAGTTCAACATCCTCTAAAGATGCGTTCTTACTATGCCGTGAT
CAAGCTCGCTTGTATAGCTGAACATCATCTTTGTGTGAACAAAACTACGGAACTCTTCAAACAATGGATGTCGGAGCCC
AAAACAAACCCAATACGACCTGATATGAGACTCATTACTTATTGTACAGCTATTCGTTTTGGTGGTCAAAAAGAATGGAA
ATTTTTAAGAAGTCAGCTGATGTTAAATGATTCTGTGAACGAGGAAGAGAATGAAAAGAAAGTGCTGGCGTTGGCGTG
TTCTCGTGATGTAGGGATTATGAAAAAGCACTTAAATTGGGTAAAAGAGAACAAGGAGTTTTGGGACACGCTTTACTAT
TTTGCGCAATCTCCAGTCGGTAACCGACTTTTATGGGACTACCTGAGTGATATAGATAATTTTGTTGAACAGAAAGACTT
CACGAAAGCCATACTAAACACACTAATTGAATATCGTTACTTAGTGTATGATGGAACGAATCATGAAAAGATCCTTGTAA
CAGAAGATAGCGATCAAATAGACCCTGAACTTCAAAAGAAGATTGAAAAATTACTTCAGAAAACCAGAAAAAAAATGA
AATGGACTGAAGACTTTTCTGACGTCATTATTCAGTGGTTGAATGAAAATG**TCCCAGTAGTTACCCATTTATTATAG** 3'

Figure 3 ECE 1 transcript (2403 bp) with underlined FWD (top) and REV (bottom) primer sequence

5'**ATGACTGATTGTCATCAATATATTAATAATGAT**AAACATTCAGTTATTCTAGATGGTAATTGTAT
AAAAGATGATTGTATAGCATTTAGTAATATAAAAGATAATCGTACATCTAATACAAGTATATTTATTG
ATAATGAACATAAAAATGGTCAAAATTCAATACATTTATGGAATTCACATAAATATAATCATGTAAA
TATATGGAGACCACGTAAAAAATGGGAATTTTGTTTAATATTTTCATTAATAATTATATTAATTTTATT
GATTATATTTATTACTTTATGGGGTGTTCAATTGGTATTTAATCAAAATAAACCCTGCTTAGAACCTA
CATGTATTAAAGTTGCTTCAGAAATTCTTAGCAATATGGATACAAGTGTATCACCTTGTGATGATTT
CTTTGCATATTCCTGCAATGGATGGATAAAACAAAATTATATACCACAAGGTCATAATTCTTGGTCT
GTAATGCGTAAACTATCCAAATATGATGAGTATTTCACCAAAGAATTACTTGAAAATCGTTCCCAT
ACAGACACTAGTCGCGGTTTCACTTTGGCACAAATTTATTACAATTCATGTATGAATGAATCTGTT
ATTGATAGCCGAAAATTAACACCTCTTTATAATTATATTAAAAAAATATTTAATGGTTGGTTATTATT
ACCAACTGGGAATCCAGCAACTGGGCAACAAAATGCTGATCTTTTTCGTCCTGACAAATTTGATT
TAACAGATATGTATTTACCTCTGTTAAAGTATTCAGGTGCAACACCATTATCCCGTATTTTAGTTGG
ACAAGATCAGATGAATTCTTCTCGATTTGTAATAGATCTTGTAGAAGGTTATCTTAGTTTACAAAG
AGAATATTATGTAAACGATTCGTATCCAGCTTACTCAAAAAAAACTGGAGCATTTCGTCGATTTAT
GCGGAATTATTCCCTCTTGTTAGGAGTACCAAACTCATCGTTACATGAAGTAGATCAAATTTATGA
ATTTGAAAAACAAATTGCATTGAGAACTGAAGATCGTAGTGAAAGAGACCCGGAGAAAAACTAC
GAACTTGTTACTCTACAAAATTTATCTACTATTTGTCCTGTGCTTAATTGGACGAAATTATTCGATT
ATTTATATGCACCACTAAACTTCAAACTTCCTCGAGATCAAGTGATTGCATTACACGATCGTACATT
TTTCAGAGAACGATGCGCTTTATTTAAAGAATATTTGAAAACAGAGACTGGTATTCGAACTTTACA
CAATGCTGCTGTATGGAGTTTTATGTGGAAAACAGTCTCACGCATGCCAAAAGATGTTAGCGAAA
TGTTGGAAGAATACAGAGAAGCAGAACTAGGATTGAAAGTTGATCCAGATCGCTGGCAGATATG
CGTTAATGAAGTACAATTTCCATTCGGAATGGTGATTGGTAGACATTTTGTTCACGAAAGATTTAA
TCAAAAAAGTAAAGAAGCGGCTACAGAAATGATCACTGAAATCAAGACTGCTTTTAAAGAGAAT
TTCGCTAGTGTCAAATGGATGCAAGAAGCTGATAAACTTAAAGCAATTGAAAAGGTGGATTCAAT
GAAAGCTTCAGTTGGTTATCCACAAAATATTAATAATATAACAAATGAAAATAAGGAGTTTTCATAT
TTTGTTGATTTGAATGAATCAACTTATTTTGAAAATGCATTACATTGTTCTGAAGCATTATTTCTTG
AAATTCTACGAGAACTAATAATCCAGGATCCAGATAGTTGGTCTTTACCAGTTCATATTGTAAATG
CATTTTATAAAGAAAATTCAAATCATATATTTTTCCCTGCTGGTATACTTCAAAGTCCATTATATAAT
CCTGGACAACCATTATCATTGAATTTTGGTGGAATTGGTATGGTTGTTGGACATGAAATAACTCAT
GCTTTTGATCAACATGGTGCTAAATTCGATGCTAAAGGAAATATGCGAGACTGGTGGAGTGCAGA
AACATTGGCAGCTTTCGAAAAAAATTCACAATGTATGATTGATCAGTATAGTAATTATTCAATTCTT
AATACAAGTCTTAATGGAAAAATGACATTAGGTGAAAACATCGCTGATAATGGAGGTCTTAAAGC
AGCCTATAAAGCGTTTAAAAAGTTGGAGGCCAAATATTCCGACAAACCTATATTACCTGGTTTAAA
TTTCACTCCAGATCAGTTATTTTTCATTGGATTTGCCCAGTTATGGTGCATTAAAAGCCTCCCTCAG
TCAGTTCTTAATACTGTGCTTTTTGATGTTCATACAGTGGAACCATACAGAGTAATTGGTACCATAA
CAAATTCAGAAGAATTTGCTAAAGTTTTCAATTGTCCTCCAGGATCTTCAATGAAT**CCTGTAAAG**
AAATGTCATGTATGGTAA 3'

Figure 4 SmLeish transcript (1008 bp) with underlined FWD (top) and REV (bottom) primer sequence

5'**ATGCAACATGAAACAGCTCAT**ATACTGGGATTACATCCTTCAATTTACGATTCTCAAAAGT
TTAGATCAGCGAAAATTCCATCTGTTCAAAACATAACACTATCTTGGTTATCAAGTAAAGGAA
ACTACGAAGTGCAAAAGACAATTCTCAGTCTACCAAAGATGCTGAAGGAAGCAAGAGAAC
ATTTTGACTGCCAAGAACTACAGGGAATCGAGCTGGATGGAATCCATTTCAGTCATAGAATA
ATGGGAAATGATTTGATGGCAACTTATCTATTGGAATCAACTAGTGTTTCGAGAATAACTCTG
GCTTATTTCGAAGATATAAACATGTATGAAGTCGACTATTCCATGGCCGATGATTTCAAATGG
GGAAAAGGTTTGGGTTGTGATTTCGTGTTGAAGAGCTGTTATGAATACATCAAGAAGAGAA
AAAGCAGAGGACAAGATATTCAGCCCTACTGTGATGTTCCCTTAGAACAAAAATGTGCCAGT
TATGGAAACGGGATCGGTACTTGCGTACTGTTTAAACACAAGAATCAATTAAATGAAGTCAA
TCAGTATATGGATGATTCGTTACCTTTTACTGACACTGAAAAGGAAAAGTATGGAGGCTTTCC
ATTTTTCGATTACTGTCCAGTTCTCCTGGTTCACCCATATGAAGAAGGTGACACAGCCCTATG
CGAGACGAAAATTGATTTAAAACCAGATTCACCACTAGACGCATTTCTGGATTATAGGGGTC
CTGATTCAGCTTGTTTTATGGACGAAACAATTAAATACGTCAATGGGTCAAGAACACACATT
GTTGAAAAGAAACCATCCTGTCATAAATTCCGATGTTCAAAAAAATTTGGTGTTGAAGTGAT
ATATCACGGACATGCGTTTCAGTGTCCTGTTGAAGGTGGGATAATAAATATAGGGGATCAATT
GAAAGATGGATATTTTTTTATCGATGTACAATGTCCAAAGTGCACATCATTG**TGCAAGGATA**
AATGTCCAAAATGA 3'

4.2 RNA extraction

The reason for first extracting and using RNA instead of DNA is that proteins are synthesized from mRNA molecules during translation. So, when isolating the mRNA linked to a specific protein and converting it into cDNA, it is then feasible to produce that specific protein that regulates a certain gene expression (Alberts et al., 2002). Proteins have various functions, such as enzymes, which are the focus of this paper (*Protein*, n.d.; *What Are Proteins and What Do They Do?*, n.d.). Another reason to use cDNA from mRNA is that the introns are cut out during translation, leaving out only the exons (*Biochain Institute Inc.*, n.d.). TRIzol™ is used to extract RNA from cells, which is then separated through chloroform extraction; the protein will be in the organic phase, the DNA in the interface, and the RNA in the aqueous phase (Rio et al., 2010). The RNA is then precipitated with isopropanol, disrupting its solubility, making it clump together, and creating a visible solid pellet after centrifugation. Turbo DNase is added to break down leftover DNA, and a DNase inhibitor is added afterward to avoid RNA degradation over time. For this step, we used the Total RNA Extraction Protocol from Bio-protocol.

In an RNase-free Eppendorf tube, 50 µl of packed eggs were disrupted using an electric handheld homogenizer with a 1.5 ml tube grinding pestle until no whole eggs were left visible under a light microscope. Then, 1 ml of TRIzol™ (Life Technologies, Gibco®, catalog number: 15596-026) was added to 50 µl of the homogenized eggs and

let to stand at room temperature (RT) for 5 minutes. Then, working in the hood, we proceeded with chloroform extraction, where 200 µl of chloroform was added to the sample; the tubes were inverted for 15 seconds and let sit for 3 minutes at RT for phase separation. Next, the tubes were centrifuged at 12,000 x g at 4 ℃ for 15 minutes. The RNA is now in the aqueous supernatant. Back in the hood, we proceeded with isopropanol precipitation; the top aqueous phase was moved to a new RNase-free Eppendorf tube, 350 µl isopropanol was added in the tube, then gently inverted several times to mix and left at RT for 10 min. The tube was centrifuged at 12,000 x g for 10 minutes at 4 ℃. Small white RNA pellets were visible at the bottom of the tube. The supernatant was carefully removed. The pellets were washed with ice-cold 75% EtOH and centrifuged at 12,000 x g at 4 ℃ for 5 minutes. The EtOH was removed. Once almost all the ethanol evaporated, the pellets were resuspended in 25 µl DEPC-H_2O (Life Technologies, Ambion®) and incubated at 60 ℃ for 10 minutes. Then, to degrade the DNA, we set up the Turbo DNase reaction, in which the 10x Turbo buffer was diluted in the RNA sample to 1x. 1 µl Turbo DNase (Life Technologies, Ambion®, catalog number: AM2238) was added to the sample and then incubated at 37 ℃ for 30 minutes. Finally, 2 µl of DNase inhibitor were added to the sample and centrifuged at 12,000 x g at 4 ℃ for 5 minutes. RNA being in the supernatant was carefully taken out. To measure the concentration of RNA, the sample was measured on a microspectrophotometer. The A260/A280 ratio was within range, meaning the sample is acceptable, showing little contamination. The RNA sample was then stored at -80 ℃ until further use.

4.3 cDNA synthesis

cDNA synthesis involves the creation of DNA from an RNA template using reverse transcription (*Five Steps to Optimal cDNA Synthesis - CZ*, n.d.; *Reverse Transcription (cDNA Synthesis) | NEB*, n.d.). Primers direct the reverse transcriptase enzymes to the desired location, where they start adding dNTPs, which builds the single-stranded RNA into cDNA (*Reverse Transcriptase - an Overview | ScienceDirect Topics*, n.d.). Reverse transcription is performed by adding an RNA mix and a reverse transcriptase (RT) mix. The RNA mix includes the extracted RNA, primers, dNTPs, and DEPC-treated water to

reduce RNA degradation. It is then heated to denature any secondary structures in the RNA, such as hairpin loops. The RT mix consists of SSIV buffer to provide suitable conditions for reverse transcription, DTT to support RT enzymes, RNase inhibitors to prevent the degradation of RNA, and RT enzymes. Heating the mixture at 55 ℃ enhances cDNA synthesis, then heating it to 80 ℃ deactivates RT enzymes to ensure that the cDNA product will not produce any non-specific PCR products later on. For cDNA synthesis, we followed the SuperScript® IV Reverse Transcriptase protocol by Invitrogen.

We annealed the primers to the extracted RNA by combining the components shown in Table 3 in a reaction tube. The tube containing the solution was briefly vortexed and centrifuged, followed by heating it at 65 ℃ for 5 minutes and incubating on ice for 1 minute. Then, the RT reaction mix was prepared by adding the components shown in Table 4 in a different reaction tube. The tube containing the solution was briefly vortexed and centrifuged. Next, the annealed RNA and RT reaction mix were combined. The reaction mixture was incubated at 55 ℃ for 10 minutes, followed by inactivating the reaction mixture by incubating it at 80 ℃ for 10 minutes. Finally, the new cDNA was used for PCR amplification.

Table 3 Solution to anneal primers to RNA

Product	**Volume**
50 μM Oligo $d(T)_{20}$ primers (catalog number: 18418-020)	1 µl
10 mM dNTP mix (catalog number: 18427-013)	1 µl
RNA	10 µl
DEPC-treated water (catalog number: 10813-012)	1 µl

Table 4 RT reaction mix

Product	**Volume**
5x SSIV buffer	4 µl
100 mM DTT	1 µl
RNaseOUT™ Recombinant RNase inhibitor (catalog number: 10777-019)	1 µl
SuperScript® IV Reverse Transcriptase	1 µl

4.4 PCR and gel electrophoresis

Polymerase Chain Reaction, or PCR, is a method that enables the amplification of a precise portion of DNA. The amplification happens in a thermocycler where the DNA strand goes through various temperatures, enabling different enzymes to perform their tasks. The first step, called denaturation, happens when the double-stranded DNA separates into two single strands, which allows space for the second step, called annealing, where primers anneal the DNA. Finally, thanks to the primers, the DNA polymerase knows where to start adding nucleotides, eventually establishing a new DNA strand from a 5' to 3' direction; this last step is elongation (*Polymerase Chain Reaction (PCR)*, n.d.). After these steps, the quantity of replicated DNA molecules increases twofold. A standard method to examine the PCR product is to run it through agarose gel electrophoresis. The gel works as a molecular sieve, which separates molecules based on size and charge. A ladder is added into the first well within the gel to act as a template to show the different sizes along the gel. Gel electrophoresis provides essential information regarding the size, quantity, and purity of the PCR product (Garibyan & Avashia, 2013). The solution used for PCR contains synthesized cDNA, a primeSTAR® mix which is a DNA polymerase of high performance with dNTPs and buffer components, nuclease-free water to provide a safe solvent from degrading the DNA template or disrupting the PCR reaction, and both FWD and REV primers.

The components shown in Table 5 were added to a PCR reaction tube for the PCR solution. The PCR tube was transferred into the thermocycler, which was programmed

according to Figure 5. Once amplified, 5 µl GeneRuler 1 kb Plus ladder was added to the first well of the agarose gel, and 10 µl of the PCR product was added to adjacent wells. The gel was run at 100 V for 30 minutes. Under UV light, the gel showed that the correct transcript was amplified, as seen in Figure 6.

Table 5 PCR solution

Product	**Volume**
PrimeSTAR® mix	25 µl
cDNA	1 µl
FWD gene specific primer	1 µl
REV gene specific primer	1 µl
Nuclease-free water	22 µl

Figure 5 Thermocycler program for PCR

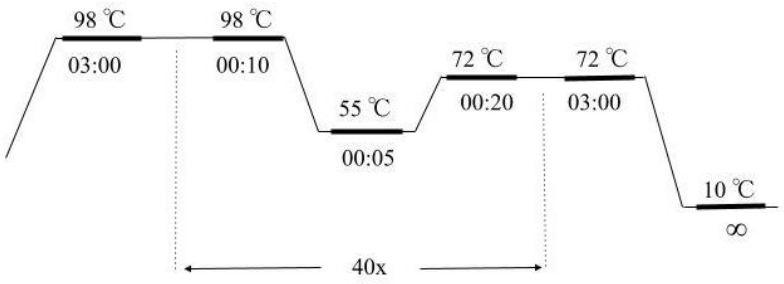
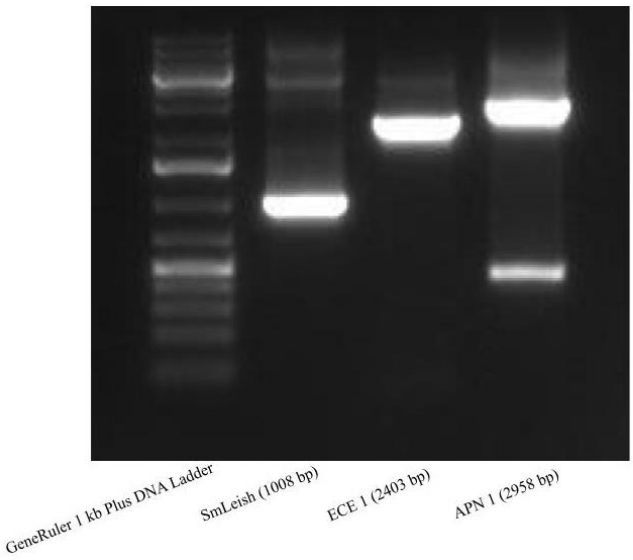


Figure 6 Presence of transcripts of interest in gel electrophoresis reading



4.5 Cleaning of PCR product

Other components besides the transcript of interest are present in the PCR product, such as primers, nucleotides, and polymerase. To ensure the purity and suitability of the PCR product for downstream applications, it is necessary to filter out those other components. DNA clean-up (purification) is conducted to efficiently retrieve DNA fragments from agarose gel, PCR, or other enzymatic reactions. PCR DNA cleanup has three steps: binding, washing, and elution. During binding, the PCR product is bound to a DF column with a DF buffer. For the washing phase, the undesired components will be flushed out from the binding column using a wash buffer, leaving the transcript in the column. Lastly, for elution, an elution buffer will be added to the column so the transcript can fall from the binding column, leaving a refined transcript product suitable for downstream applications. For DNA elution, we used the Gel/PCR DNA fragments Extraction Kit by Geneaid.

To prepare our sample for DNA elution, 50 µl of the PCR product was transferred into a 1.5 microcentrifuge tube with 200 µl DF buffer and mixed by vortex. Then, for DNA binding, a DF column was placed in a 2 ml collection tube, and the sample mixture was transferred to the DF column. The tube was centrifuged at 16,000 x g for 30 seconds; the flowthrough was discarded. For washing, 600 µl wash buffer with ethanol was added into the center of the DF column and let to stand at room temperature for 1

minute. The tube was centrifuged at 16,000 x g for 30 seconds; the flowthrough was discarded and followed by another centrifugation at 16,000 x g for 3 minutes to dry the DF column. For the last step, elution, the dried DF column was transferred into a new 1.5 ml microcentrifuge tube. Then, 20 µl of elution buffer was added into the center of the DF column and let to stand for 2 minutes to absorb the buffer thoroughly. Ultimately, the tube was centrifuged at 16,000 x g for 2 minutes to elute the purified DNA.

Once the DNA has eluted, ensuring the sample's integrity is essential by checking for its concentration. By knowing the concentration of the DNA sample, we can tailor the amount of DNA and other solutions needed for other experiments, ensuring optimal results. We used a microspectrophotometer to measure the DNA sample's concentration. Firstly, a blank sample was run by putting one drop of elution buffer with a pipette onto the microspectrophotometer. Once the blank sample was in the system, the elution buffer was wiped out using a cloth and replaced with one drop of eluted DNA sample; the sample was measured. The result showed a concentration within the normal range, meaning the DNA sample was ready for downstream applications.

4.6 Ligation into pJET1.2 vector

The eluted DNA product will be ligated into a plasmid vector and integrated into *E. coli* competent cells, and the cells will be left to grow. The vector containing the DNA of interest will be replicated during the cell division of the competent cells. The vector used here is the pJET1.2/blunt cloning vector by ThermoFisher, a type of vector usually utilized for molecular cloning and DNA manipulation. The vector has an ampicillin resistance gene; this ensures that only the cells with the recombinant plasmid can survive and propagate in an environment spiked with the antibiotic. The DNA is ligated into the vector by the T4 DNA ligase; it glues the fragments together by creating a phosphodiester bond between the 5’-phosphate and 3’-OH group of adjacent atoms. The reaction is done in a 2X reaction buffer, providing an efficient chemical environment. For this step, we used the Thermo Scientific^{TM} cloning protocol.

The ligation reaction was set up on ice, where the ligation reaction mixture, shown in Table 6, was added to an Eppendorf tube. The tube was vortexed and centrifuged for 5 seconds, then incubated at room temperature for 5 minutes.

Table 6 Ligation reaction mixture

Product	**Volume**
2X reaction buffer	10 µl
Nuclease-free water	7 µl
DNA sample	1 µl
pJET1.2 cloning vector	1 µl
T4 DNA ligase	1 µl

4.7 Heat-shock transfection of XL1-blue competent cells

The *E. coli* strain used for this step is the XL1-blue competent cells. The term "competent" is employed to showcase that the cells have been manipulated to enhance their capacity to incorporate foreign DNA molecules. XL1-blue competent cells were chosen due to their genetic mutations that make them proficient at cloning and storing DNA. Once the cells have taken up the vector from the heat-sock, they need to recover with a microbial growth medium that provides nutrients to the cells, known as SOC medium. The composition for the SOC medium is shown in Table 7; the pH was adjusted to 7.0, autoclaved, and cooled down below 60 ℃ before adding the rest in Table 8. Once recovered, the cells were spread out on an LB plate spiked with ampicillin to a final concentration of 100 µg/ml. Table 9 shows the LB agar composition; the pH was adjusted to 7.0 with NaOH, autoclaved, and left to cool down to 55 ℃ before pouring into Petri dishes (25 ml per 100-mm plate). Once spread, the cells are then left to grow under the optimal temperature of 37 ℃. For this step, we followed the Transformation Protocol by Hopegen Biotech.

Table 7 Composition of SOC medium

Product	**Amount**
Demi H_2O	95 ml
Tryptone	2 g
Yeast extract	0.5 g
1M NaCl	1 ml
1M KCl	0.25 ml

Table 8 Remaining composition of SOC medium

Product	**Volume**
Filter-sterilized 2M Mg^{+2} (1M $MgCl_2$ + 1M $MgSO_4$)	1 ml
Filter-sterilized 2M glucose solution	1 ml
Sterile demi H_2O	To final volume of 100 ml

Table 9 Composition of LB agar

Product	**Amount**
Agar	20 g/L
NaCl	10 g/L
Tryptone	10 g/L
Yeast extract	5 g/L
Demi H_2O	To final volume of 1 L

To incorporate the ligation mixture into the competent cells, 50 µl of competent cells were thawed on ice. Once thawed, 5 µl of the ligation mixture was added into the tube containing the competent cells and left to incubate for 30 minutes on ice. Then, the competent cells were heat shocked at 42 ℃ for 30 seconds and placed on ice for 2 minutes. We added 250 µl of SOC medium into the tube containing the competent cells and put it on an incubator plate at 37 ℃ with shaking at 227 rpm for 1 hour. Once finished, 100 µl of the transformed competent cells were transferred and spread out on an LB plate spiked with ampicillin to a final concentration of 100 µg/ml, then left to grow overnight at 37 ℃.

4.8 Colony PCR and gel electrophoresis

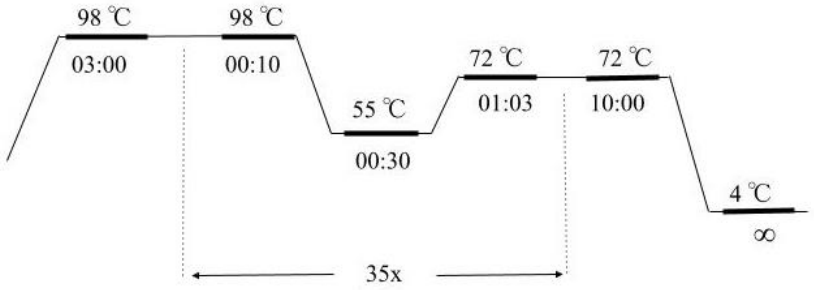
Not all grown bacterial colonies will have the desired transcript, as some might only contain an empty plasmid vector without the ligated DNA of interest. A few random colonies were selected for colony PCR to assess the desired transcript's presence by amplifying whatever is present between the plasmid FWD and REV primer binding site. The colony PCR solution includes pJET1.2 FWD and REV primers, nuclease-free water, and DreamTaq™ PCR Master Mix, which contains DNA polymerase, dNTPs, and buffer components.

We first made a PCR solution to prepare for colony PCR, as shown in Table 10. In 10 PCR tubes, 20 µl of the PCR solution was added in each tube. 10 random colonies were marked and labeled, and a small part of the colony was added to its PCR tube using a pipette. Then, the PCR tubes were put into a thermocycler, which was programmed according to Figure 7. Once amplified, the PCR product was examined through gel electrophoresis. A 5 µl GeneRuler 1 kb Plus ladder was added into the first well of the agarose gel, and 10 µl of the colony PCR product from each tube was added into adjacent wells. The gel was run at 100 V for 30 minutes. Five colonies from which the colony PCR product matched the transcript of interest were used for the miniprep.

Table 10 Colony PCR solution

Product	**Volume**
DreamTaq™ PCR Master Mix	100 µl
REV pJET1.2 primer	5.5 µl
FWD pJET1.2 primer	5.5 µl
Nuclease-free water	99 µl

Figure 7 Thermocycler program for colony PCR



4.9 Plasmid isolation

After running the gel electrophoresis to check which colonies contain our plasmid vector with the ligated transcript of interest, we need to extract the plasmid from the positive colonies. This phase, called miniprep, involves isolating and purifying plasmid DNA from bacterial cells. The miniprep procedure has an overall of 6 steps. First, the bacterial cultures that contain the transcript of interest are grown in an LB medium that supports bacterial growth. The composition of the LB medium is shown in Table 11; the pH was adjusted to 7.0 with NaOH and then autoclaved. Once the cultures have grown for 16 hours, the second step consists of harvesting the bacterial cells by centrifugation, which induces the cells to accumulate into pellets at the bottom of the tube. Third, the pellets are resuspended to stabilize the plasmid. In the fourth step, the cells are lysed, which releases the cellular content, such as the plasmid of interest. During the fifth step,

the pH of the lysate is neutralized to stabilize the DNA and prevent further degrading. The sixth step isolates the plasmid from the other cellular debris in the lysate, such as DNA, RNA, proteins, and other cellular elements. The sample goes through a series of centrifugation steps where the unwanted cellular components are washed away from the column. Once the flowthroughs containing the debris are discarded, the plasmid DNA is eluted from the column. For this step, we used the GenElute™ Plasmid Miniprep Kit.

Table 11 Composition of LB medium

Product	**Amount**
NaCl	10 g/L
Tryptone	10 g/L
Yeast extract	5 g/L
Demi H_2O	To final volume of 1 L

To grow the previously 5 chosen colonies, each colony was transferred into its own Eppendorf tube containing 5 ml of LB medium and was left to incubate overnight on an incubator plate at 37 ℃ with shaking at 235 rpm until the solution had a cloudy color, indicating a successful growth of bacteria. The following steps were done for each colony. To harvest and lyse the cells, the tubes were centrifuged for 1 minute at 12,000 x g, forming pellets at the bottom, and the supernatant was discarded. The pellets were resuspended in a 200 µl resuspension solution by pipetting up and down. 200 µl of lysis solution was added, inverted gently to mix, and left to clear for 5 minutes. To prepare the cleared lysate, 350 µl of Neutralization Solution (S3) was added and inverted 6 times to mix, then centrifuged at 16,000 x g for 10 minutes. Separately, the binding columns were prepared by putting a binding column in a collection tube for each sample, and 500 ml Column Preparation Solution was added into the column and centrifuged at 12,000 x g for 2 minutes; the flowthrough was discarded. To bind the plasmid DNA to the column, each cleared lysate was transferred into its prepared binding column and centrifuged at 12,000 x g for 2 minutes; the flowthrough was discarded. To remove the

cellular debris, 750 µl Wash Solution was added into the column and centrifuged at 12,000 x g for 2 minutes; the flowthrough was discarded. The column was centrifuged again at 12,000 x g for 2 minutes to dry. To elute the purified plasmid DNA, the columns were transferred into their new collection tube, followed by adding 100 µl Elution solution and centrifuged at 12,000 x g for 2 minutes. We now have pure plasmid DNA ready for downstream application.

The concentration of the miniprep product was measured on a microspectrophotometer. Firstly, a blank sample was run by putting one drop of elution solution with a pipette onto the microspectrophotometer. Once the blank sample was in the system, the elution solution was wiped out using a cloth and replaced with one drop from the first colony; the sample was measured. The first colony was wiped with a cloth and replaced with one drop of the miniprep product from the second colony; the sample was measured. The same steps of cleaning and replacing the drop were followed for the other three colonies. The colonies contained suitable DNA concentrations.

To check for the exact constitution of genetic information inserted in the plasmid vector, the miniprep products were sent for sanger sequencing at Seqme s.r.o. The samples were prepared according to Seqme's sequencing template. As shown in Table 12, each colony had two sequencing solutions in an Eppendorf tube: one with FWD buffer and the other with REV buffer.

Table 12 Sequencing solution

Product	**Volume**
Miniprep product	3 µl
pJET1.2 FWD/REV buffer	2.5 µl
Nuclease-free water	4.5 µl

With the DNA identified from the miniprep product, the sequences were verified by NCBI BLAST analysis and aligned with the original sequence.

4.10 PCR amplification of sequenced transcripts for cloning into pET-SUMO

With the transcript's presence confirmed in the miniprep DNA, we continued by amplifying the product via PCR and checking it with gel electrophoresis to verify the specificity of the amplification.

The PCR primer working solutions were prepared according to Table 13 to make a 10x dilution for both primers. For the PCR solution itself, the components in Table 14 were added to a PCR tube. The tubes were transferred into a thermocycler, which was programmed according to Figure 8.

Once the PCR product was done, 10 µl of GeneRuler 1 kb Plus ladder was added to the first well of the gel and 40 µl of the PCR product into the adjacent well. It was run at 100 V for 40 minutes. The gel ran longer to separate the bands better, facilitating the cutting of the bands for the next step. The result from the gel electrophoresis showed that we have the correct transcript from the PCR.

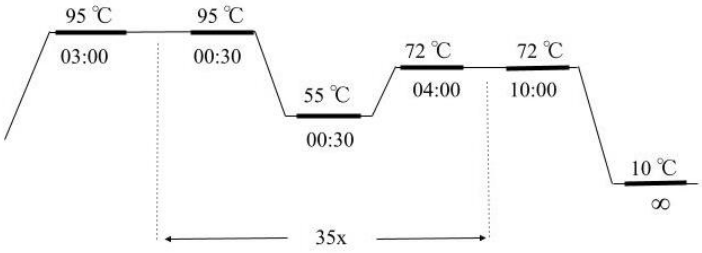
Table 13 10x dilution of gene-specific FWD/REV primer for PCR

Product	**Volume**
FWD/REV gene specific primer	2 µl
Nuclease-free water	18 µl

Table 14 PCR solution

Product	**Volume**
DreamTaq™ PCR Master Mix	25 µl
FWD pJET1.2 primer	1 µl
REV pJET1.2 primer	1 µl
DNA template	1 µl
Nuclease-free water	22 µl

Figure 8 Thermocycler program for PCR



4.11 Gel DNA extraction

To ensure that we would work with the correct transcript from the agarose gel, the band that matched the size of the desired transcript was cut out and then eluted to isolate the transcript. Gel DNA elution is the same process of eluting the DNA as PCR DNA elution; the only difference is that the gel must be cut out and dissolved before binding the DNA. To do so, we followed the Gel Extraction Protocol from Geneaid.

The gel was put into the UV machine to make the bands visible while wearing UV protective glasses, the desired band was cut out using a clean and disinfected scalpel, and any extra gel was removed. To dissolve the gel, it was transferred into a 1.5 ml microcentrifuge tube with 500 μl of DF buffer and vortexed. To ensure that the gel completely dissolved, it was incubated at 60 ℃ for 10 minutes; during that time, the gel was vortexed every 2 minutes. Once dissolved, it was let to cool down to room temperature. For DNA binding, a DF Column was placed in a 2 ml Collection Tube. Then, 800 µl of the dissolved gel was transferred to the column and centrifuged at 16,000 x g for 30 seconds; the flowthrough was discarded. For washing, 400 µl of W1 Buffer was added into the column and centrifuged at 16,000 x g for 30 seconds; the flowthrough was discarded. 600 µl of Wash Buffer was added into the column, let to stand for 1 minute at room temperature, and centrifuged at 16,000 x g for 30 seconds; the flowthrough was discarded. To dry the column, it was centrifuged again at 16,000 x g for 3 minutes. The dried column was transferred into a new 1.5 ml microcentrifuge tube. To elute the DNA, 20 µl Elution Buffer was added into the center of the dried

column, let to stand for 2 minutes to ensure that the buffer is thoroughly absorbed, and then centrifuged for 2 minutes at 16,000 x g to elute the purified DNA. We now have our desired eluted DNA.

4.12 Ligation into pET-SUMO vector

To insert the transcript of interest into competent cells, it will first need to be ligated into a plasmid vector. The vector used here is pET-SUMO, mainly utilized for protein expression. This vector is equipped with a kanamycin resistance gene to ensure that only the cells containing the plasmid will proliferate in a medium containing this particular antibiotic. We referred to the Champion™ pET-SUMO Protein Expression System by Invitrogen for this ligation.

To prepare the ligation mixture, the components in Table 15 were added to an Eppendorf tube and left to incubate at 15 ℃ overnight.

Table 15 Ligation mixture

Product	**Volume**
Eluted PCR product	2 µl
Ligation Buffer	1 µl
pET-SUMO vector	2 µl
Nuclease-free water	4 µl
T4 DNA ligase	1 µl

4.13 Heat-shock transfection of XL1-blue competent cells

With the transcript being ligated onto the plasmid vector, it will go through replication using competent cells. For this step, we followed the Transformation Protocol of Invitrogen.

2 µl of the ligation reaction was added into a thawed vial of XL1-blue competent cells and mixed gently, then left to incubate on ice for 30 minutes. The cells were heat-

shocked for 30 seconds at 42 ℃ without shaking; afterward, the cells were immediately put on ice for 2 minutes, and then 250 µl of room temperature SOC medium was added. The tubes were tightly caped and put on an incubator plate at 37 ℃ for 1 hour to shake horizontally at 200 rpm. Once finished, 100 µl of the ligated cells were spread per LB plate for two plates containing an LB medium spiked with Kanamycin to a final concentration of 50 µg/ml and left to grow overnight at 37 ℃.

4.14 Colony PCR and gel electrophoresis

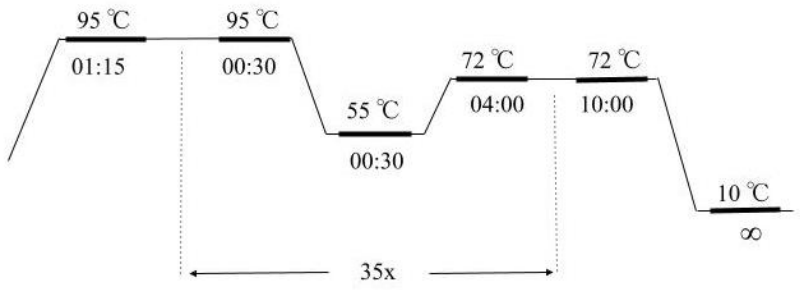
The transfected cells underwent PCR to amplify the inserted transcript to ensure we were still working with the proper transcript.

A PCR solution was made according to Table 16 to prepare for colony PCR to check the correct orientation of the insert in the vector since TA cloning is not oriented. 3 random colonies were marked and labeled, and a small part of the colony was added into 19.5 µl of PCR solution into its PCR tube. Then, the PCR tubes were put into a thermocycler, which was programmed according to Figure 9. The PCR products were loaded into an agarose gel where 5 µl of GeneRuler 1 kb Plus ladder was loaded into the first well, and 10 µl of each PCR product was loaded into adjacent wells. The gel was run at 100 V for 30 minutes. The colony from which the colony PCR product matched the transcript of interest was used for the miniprep.

Table 16 Colony PCR solution

Product	**Volume**
DreamTaq™ PCR Master Mix	90 µl
FWD pET-SUMO primer	4.5 µl
REV gene specific primer	4.5 µl
Nuclease-free water	81 µl

Figure 9 Thermocycler program for colony PCR



4.15 Plasmid isolation

To grow the colony that contains the transcript of interest, it was transferred with a pipette into its falcon tube containing 5 ml LB medium and left overnight on an incubator plate at 37 ℃ with shaking at 185 rpm. To proceed with the miniprep, we followed the GenElute™ Plasmid Miniprep Kit.

To harvest the cells, the overnight cultures were first centrifuged at 6,000 x g for 2 minutes to pellet them down, and the supernatant was discarded. The pellets were resuspended in 200 µl Resuspension Solution by pipetting up and down. Then 200 µl of Lysis Solution was added and gently inverted several times to mix, letting it clear for 5 minutes. To prepare the cleared lysate, 350 µl Neutralization Solution was added and inverted 6 times to mix, then centrifuged for 10 minutes at 22,000 x g. To prepare the binding column, 500 µl Column Preparation Solution was added into a binding column in a collection tube, then centrifuged at 12,000 x g for 1 minute, and the flowthrough was discarded. To bind the plasmid DNA onto the column, the cleared lysate was transferred into the binding column and centrifuged at 12,000 x g for 1 minute, and the flow-through was discarded. To remove the cellular debris, 750 µl Wash Solution was added to the column and centrifuged at 12,000 x g for 1 minute; the flowthrough was discarded. The column was then centrifuged again at 12,000 x g for 1 minute to dry the column. To elute the plasmid from the column, the dried column was transferred into a new collection tube. 100 µl Elution Solution was added into the column and centrifuged at 12,000 x g for 1 minute.

The concentration of the miniprep product was measured using a microspectrophotometer. Firstly, a blank sample was run by putting one drop of elution solution with a pipette onto the microspectrophotometer. Once the blank sample was in the system, the elution solution was wiped out using a cloth and replaced with one drop of the miniprep product.

To check for the exact constitution of genetic information inserted in the plasmid vector, the miniprep products were sent for sanger sequencing at Seqme s.r.o. The samples were prepared accordingly to meet Seqme's requirements, as shown in Table 17; each colony had two sequencing solutions in an Eppendorf tube, one with FWD primer and the other with REV primer.

Table 17 Sequencing solution

Product	**Volume**
Miniprep product	7 µl
FWD/REV pET-SUMO primer	2.5 µl

With the DNA identified from the miniprep product, the sequence was analyzed by alignment with the original to determine whether it corresponds to the intended transcript.

4.16 Heat shock transfection of C43(DE3) competent cells

Upon confirmation of the transcript's presence and its correct orientation in the plasmid vector, the next step involves inserting the vector into C43(DE3) competent cells, which are specialized for protein expression. These cells provide an ideal environment for the transcription and translation of the target gene, ultimately producing the desired protein.

1 µl of miniprep product was added into a vial of C43(DE3) competent cells and left to incubate on ice for 30 minutes, heat shocked at 42 ℃ for 45 seconds, and cooled down on ice for 2 minutes. Then, 450 µl of SOC medium was added to the cells to promote recovery. The cells were put on an incubator plate at 37 ℃ with shaking at 185

rpm for 1 hour. Once finished, 100 µl of the transformed cells were spread onto two LB plates spiked with kanamycin to a final concentration of 50 µg/ml and incubated overnight at 37 ℃.

4.17 Colony PCR and gel electrophoresis

To confirm the integrity of the genetic material, the cells that had undergone transfection were subjected to PCR.

A PCR solution was made according to Table 16 to prepare for colony PCR. Three random colonies were marked and labeled, and a small part of the colony was added into 19.5 µl of PCR solution into its PCR tube. Then, the PCR tubes were put into a thermocycler, which was programmed according to Figure 9. The PCR products were loaded into an agarose gel where 5 µl of GeneRuler 1 kb Plus ladder was loaded into the first well, and 10 µl of each PCR product was loaded into adjacent wells. The gel was run at 100 V for 30 minutes.

One positive colony was transferred with a pipette into 10 ml of LB medium spiked with kanamycin to a final concentration of 50 µg/ml and let to incubate overnight on an incubator plate at 37 ℃ with shaking at 235 rpm.

4.18 Protein expression and extraction

The C43(DE3) competent cells are currently inactive in terms of gene expression due to the presence of the lac repressor protein within the plasmid. This protein physically blocks the RNA polymerase from transcribing the transcript of interest. However, upon induction with an agent such as IPTG, the agent binds to the lac repressor protein, dissociating it from the lac operator region of the plasmid. This, in turn, enables RNA polymerase to initiate transcription and produce the desired recombinant protein.

To prepare the cells for expression, 2.5 ml of the overnight cultures were transferred into 50 ml of prewarmed LB medium with kanamycin and left on an incubator plate at 37 ℃, shaking at 300 rpm, until the OD_{600} is 0.5 to 0.7. For the noninduced control, 1 ml of the sample was taken before induction, and the cells were pelleted down and resuspended in 50 µl 1x SDS-PAGE sample buffer. The noninduced control was frozen

at -20 ℃ until needed for SDS-PAGE. The expression was induced by adding IPTG to a final concentration of 1 mM. The induced cultures were grown for 4 hours on an incubator plate at 37 ℃, shaking at 300 rpm. For the induced control, 1 ml of the induced sample was taken before pelleting the cells and resuspending them in 100 µl 1x SDS-PAGE sample buffer. The induced control was frozen at -20 ℃ until needed for SDS-PAGE. The cells were then harvested by centrifugation at 4000 x g for 20 minutes.

To extract the protein, the cell pellets were resuspended in 5 ml of lysis buffer for native purification. The sample was frozen on dry ice and thawed in cold water. Once thawed, it was sonicated at 6 x 10 seconds with 10-second pauses at 200 W on ice. The sample was centrifuged at 10,000 x g at 4 ℃ for 30 minutes. The supernatant (crude extract A, soluble protein) was kept and frozen at -20 ℃ until needed for SDS-PAGE. The pellets were resuspended in 5 ml lysis buffer, making a suspension of the insoluble matter (crude extract B, insoluble protein) that was frozen at -20 ℃ until needed for SDS-PAGE.

4.19 SDS-PAGE analysis

A detergent such as SDS is used to denature proteins to ensure uniformity in their negative charge density. Samples are taken at various stages to assess the success of the protein's expression and solubility through SDS-PAGE. The stacking gel, located at the top layer of the gel, facilitates the simultaneous entry of all samples into the running gel. The gel is then stained with a dye to visualize the protein bands. Table 18 and Table 19 show the composition of the stacking gel (12%) and the running gel (4%), respectively. The recipes of the products used for both gels are also displayed in Table 20.

Table 18 Composition of stacking gel (12%)

Product	**Volume**
Buffer A	3.6 ml
Buffer B	3 ml
Demi H_2O (dH_2O)	5.4 ml
TEMED	7.5 µl
APS	75 µl

Table 19 Composition of running gel (4%)

Product	**Volume**
Buffer A	0.6 ml
Buffer C	1.5 ml
dH_2O	3.9 ml
TEMED	7.5 µl
APS	60 µl

Table 20 Recipes for products used in stacking and running gel

Product	Composition
Buffer A	- Acrylamide/Bis-acrylamide, 30% solution (Sigma-Aldrich)
Buffer B	- 3 g Tris (Trizma base), pH of 6.8 - 0.2 g SDS, pH of 8.8 - dH_2O to 100 ml
Buffer C	- 9.1 g Tris - 0.2 g SDS - dH_2O to 100 ml
TEMED	- N,N,N,N-Tetramethylethylenediamine (Sigma-Aldrich)
APS	- 10% ammonium perfusate in dH_2O

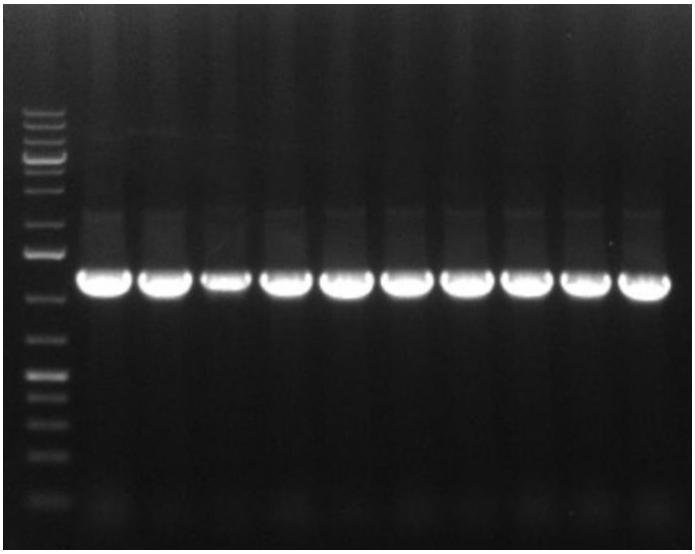
To prepare for the gel, 5 µl of 2x SDS-PAGE sample buffer was added to 5 µl of crude extracts A and B. The samples were heated with the frozen noninduced and induced samples at 95 ℃ for 5 minutes. All the samples were centrifuged at 25,000 x g for 1 minute to pellet any debris; the supernatant was used for SDS-PAGE. 10 µl PageRuler™ protein ladder was added in the first well of the 12% SDS-PAGE gel, and adjacently, 20 µl of the noninduced, induced samples, crude extract A and B, were loaded in. The gel was run at 200 V until the samples had migrated enough through the gel. Then, the gel was rinsed with distilled water and washed again on a horizontal shaking plate for 5 minutes at 100 rpm. Once cleaned, the gel was stained with EZBlue™ Gel Staining Reagent on a horizontal shaking plate for 1 hour at 100 rpm. The staining reagent was removed, and the gel was destained with distilled water on a horizontal shaking plate at 130 rpm; the water was replaced every 15 minutes for 1 hour and 30 minutes.

5. Results

We successfully isolated total RNA from S. mansoni using TRIzol reagent and then converted it into stable cDNA with reverse transcriptase. This formed the initial step

necessary for PCR amplification of specific genes. We designed primers with care to amplify transcripts for all three selected proteases, which were then expressed in the pET-SUMO vector for TA cloning. The amplified transcripts were then extracted from the gel and ligated into the pJET1.2 vector for sequencing and safekeeping. Afterward, we transfected the finished construct to XL1-blue competent cells via heat shock and screened for positive colonies with colony PCR. Following this, we grew selected positive colonies overnight, isolated plasmids with a plasmid mini-isolation kit, and sent them for sequencing to verify the presence of the chosen insert and to check for any possible errors in the nucleotide sequence. We then subcloned the inserts into a pET-SUMO vector and transfected them to XL1-blue competent cells. We repeated colony PCR to ensure the correct insert was present. This step proved to be the most challenging, as we were only able to clone SmLeish into the pET-SUMO vector successfully. Due to time constraints, we opted not to repeat the ligation step with the other two genes and instead proceeded only with SmLeish. Despite SmLeish being the only one to be successfully transfected into XL1-blue cells on the pET-SUMO vector, there were only three small colonies, one of which contained the desired transcript. After isolating the plasmids, we screened the SmLeish construct again with colony PCR and sent it for sequencing. Plasmids with the correct insert were used to transfect C43(DE3) expression cells, and we were able to begin inducing protein expression experimentally once the insert was checked with colony PCR, as seen in Figure 10. Unfortunately, we did not obtain any results from protein expression, as the SDS-PAGE gel was unreadable, probably due to the high viscosity of some samples, which may have leaked into the other wells.

Figure 10 Gel electrophoresis from colony PCR of SmLeish on pET-SUMO following transfection of C43(DE3) competent cells



The amino acid sequence of ECE 1, APN 1, and SmLeish can be found in the UniProt entry A0A3Q0KR44, A0A5K4F103, and A0A3Q0KS51, or in Figure 11, Figure 12, and Figure 13, respectively.

Figure 11 ECE 1 amino sequence with a length of 800 amino acids (aa) and a mass of 92,502 Da

MTDCHQYINNDKHSVILDGNCIKDDCIAFSNIKDNRTSNTSIFIDNEHKNGQNSIH
LWNSHKYNHVNIWRPRKKWEFCLIFSLIIILILLIIFITLWGVQLVFNQNKPCLEPT
CIKVASEILSNMDTSVSPCDDFFAYSCNGWIKQNYIPQGHNSWSVMRKLSKYDE
YFTKELLENRSHTDTSRGFTLAQIYYNSCMNESVIDSRKLTPLYNYIKKIFNGWL
LLPTGNPATGQQNADLFRPDKFDLTDMYLPLLKYSGATPLSRILVGQDQMNSSRF
VIDLVEGYLSLQREYYVNDSYPAYSKKTGAFRRFMRNYSLLLGVPNSSLHEVDQ
IYEFEKQIALRTEDRSERDPEKNYELVTLQNLSTICPVLNWTKLFDYLYAPLNFKL
PRDQVIALHDRTFFRERCALFKEYLKTETGIRTLHNAAVWSFMWKTVSRMPKD
VSEMLEEYREAELGLKVDPDRWQICVNEVQFPFGMVIGRHFVHERFNQKSKEA
ATEMITEIKTAFKENFASVKWMQEADKLKAIEKVDSMKASVGYPQNINNITNEN
KEFSYFVDLNESTYFENALHCSEALFLEILRELIIQDPDSWSLPVHIVNAFYKENS
NHIFFPAGILQSPLYNPGQPLSLNFGGIGMVVGHEITHAFDQHGAKFDAKGNMR
DWWSAETLAAFEKNSQCMIDQYSNYSILNTSLNGKMTLGENIADNGGLKAAYK
AFKKLEAKYSDKPILPGLNFTPDQLFFIGFAQLWCIKSLPQSVLNTVLFDVHTVEP
YRVIGTITNSEEFAKVFNCPPGSSMNPVKKCHVW

Figure 12 APN 1 amino acid sequence with a length of 985 aa and a mass of 116,126 Da

MLAVDNSDGRYFDSTDHCQCTHSMHAKNLTNKTLIQSKKSFIILIVLTLELIFVDH
ASAVSSSSYFENGPVFKNILTTSATENIDDENNMSNYSSEDFRLPHTLFPHFYDLSI
QVHLHDNKSQAFFFNGSVTIKVFCNVSTTEIFVHAYKTLDVKLDKIQMVPLHERN
QTSSIIEIEDINYFQDAECYRIKLKMPLRPNTYYNLTFGQFCSDLNYVSEGLYLSRY
LENGVYEYLASTQLEANYARRVFPCWDEPEFKAKFKVNIIRHKSFHSLSNMNLES
TKVLYDNWCLDTYNTSVKMSTYLLAIVVSRFSNIRRTDNRGRNFTVWARPEKIRS
AKYALDVGIKLLEHFEDYFGIPYSLHKMDMIAIPNSSITAMENWGLITFRENLML
WNPENGSIASPIDVASTVSHELSHQWFGNIVTMKWWDNLWLNEGFATFMEYIGV
QSLHPEWKVDELFILDELIQVLRSDTSKKSRPVIFPANTTTQIKRMFDLITYSKGAS
LIWMMEKFMGRSAFQNGLKKYLIQNQYKNTNEKDLWKALSNEWNTQGNHLDI
EFIMDSWTKQMNYPLVIVRRHGSNMFCFEQTRYLQNCDNNSLPENHEYSWIIPIT
YGSAKTVNWTDADILWMVNKTMNQTMNISSDDWYLLNVRQEGLYRVHYADNN
WKLLVNQLQRNFMAIPVYSRTQILNDLFSLANHHIVPYTLFLNATKYLDRENQFI
VWITSSRAFRYINSMFVLNENYDVYQAYLRTLLDNQIRSVNWSFVSEVQHPLKM
RSYYAVIKLACIAEHHLCVNKTTELFKQWMSEPKTNPIRPDMRLITYCTAIRFGGQ
KEWKFLRSQLMLNDSVNEEENEKKVLALACSRDVGIMKKHLNWVKENKEFWD
TLYYFAQSPVGNRLLWDYLSDIDNFVEQKDFTKAILNTLIEYRYLVYDGTNHEKIL
VTEDSDQIDPELQKKIEKLLQKTRKKMKWTEDFSDVIIQWLNENVPVVTHLL

Figure 13 SmLeish amino acid sequence with a length of 335 aa and a mass of 38,296 Da

MQHETAHILGLHPSIYDSQKFRSAKIPSVQNITLSWLSSKGNY
EVQKTILSLPKMLKEAREHFDCQELQGIELDGIHFSHRIMGND
LMATYLLESTSVSRITLAYFEDINMYEVDYSMADDFKWGKG
LGCDFVLKSCYEYIKKRKSRGQDIQPYCDVPLEQKCASYGN
GIGTCVLFKHKNQLNEVNQYMDDSLPFTDTEKEKYGGFPFF
DYCPVLLVHPYEEGDTALCETKIDLKPDSPLDAFLDYRGPDS
ACFMDETIKYVNGSRTHIVEKKPSCHKFRCSKKFGVEVIYHG
HAFQCPVEGGIINIGDQLKDGYFFIDVQCPKCTSLCKDKCPK

6. Discussion

It is worth noting that while the cloning of SmLeish into pET-SUMO was achieved successfully, the same cannot be said for ECE 1 and APN 1. The reason for this lack of success could be attributed to several factors. Firstly, the larger size of the two transcripts in question may have resulted in a suboptimal molar ratio between the PCR product and vector. Additionally, the gel-extracted PCR product may have been contaminated with high salt concentrations or EDTA, thus inhibiting the ligase reaction. Lastly, the proteases being investigated may have had a toxic effect on bacterial cells, further complicating the cloning process. Given sufficient time to conduct the necessary experiments, it is probable that the successful cloning of ECE 1 and APN 1 transcripts into pET-SUMO could have been achieved, as we had only attempted ligation once. Regarding SmLeish, the viscosity of the samples prepared for SDS-PAGE analysis

proved to be a challenge during preparation for gel electrophoresis. The high viscosity of the samples made it difficult to pipette them into the wells, which may have led to contamination of the surrounding wells, making the gel unreadable.

Despite the outcomes so far, more work will be conducted following the submission of this thesis to correctly express and purify ECE 1, APN 1, and SmLeish. The purified proteases could represent an advancement in parasitology regarding *S. mansoni*. Understanding the biochemical mechanism, such as the ability to infiltrate tissues and avoid the host's immune response, could help develop drugs or inhibitors that can interfere with essential biological pathways, hampering the parasite's growth. Additionally, the proteases may serve as vaccine candidates by inducing immunity against schistosomiasis and could help provide vaccines that may target the trematode at various life stages.

7. Conclusion

To conclude, this experiment successfully isolated RNA from *S. mansoni* and converted it into cDNA, which was then PCR amplified and ligated into plasmid vectors for transfection into competent cells. While the successful cloning of SmLeish into the pET-SUMO vector was achieved, ECE 1 and APN 1 proved to be more challenging due to several factors, such as the larger size of the transcripts, possible gel-extracted PCR product contamination, and potential toxicity of the proteases on bacterial cells. Furthermore, the high viscosity of the SmLeish samples during SDS-PAGE analysis made it challenging to pipette into the wells, resulting in unreadable results. However, with more time and experimentation, it could have been possible to express and purify all three proteases successfully. This experiment showcases the difficulties of expressing large genes while contributing to the ongoing efforts to understand the molecular mechanisms behind the egg-host interaction to formulate remediations for schistosomiasis.

8. Bibliography

Alberts, B., Johnson, A., Lewis, J., Raff, M., Roberts, K., & Walter, P. (2002). Analyzing Protein Structure and Function. In *Molecular Biology of the Cell. 4th edition*. Garland Science. https://www.ncbi.nlm.nih.gov/books/NBK26820/

Biochain Institute Inc. (n.d.). Biochain Institute Inc. Retrieved February 23, 2024, from https://www.biochain.com/

CDC - DPDx—Schistosomiasis Infection. (2019, August 14). https://www.cdc.gov/dpdx/schistosomiasis/index.html

Crellen, T., Allan, F., David, S., Durrant, C., Huckvale, T., Holroyd, N., Emery, A. M., Rollinson, D., Aanensen, D. M., Berriman, M., Webster, J. P., & Cotton, J. A. (2016). Whole genome resequencing of the human parasite Schistosoma mansoni reveals population history and effects of selection. *Scientific Reports*, *6*(1), 20954. https://doi.org/10.1038/srep20954

Five Steps to Optimal cDNA Synthesis—CZ. (n.d.). Retrieved February 22, 2024, from https://www.thermofisher.com/uk/en/home/life-science/pcr/reverse-transcription/5steps-cDNA.html

Garibyan, L., & Avashia, N. (2013). Research Techniques Made Simple: Polymerase Chain Reaction (PCR). *The Journal of Investigative Dermatology*, *133*(3), e6. https://doi.org/10.1038/jid.2013.1

Grevelding, C. G. (2004). Schistosoma. *Current Biology*, *14*(14), R545. https://doi.org/10.1016/j.cub.2004.07.006

Hambrook, J. R., Kaboré, A. L., Pila, E. A., & Hanington, P. C. (2018). A metalloprotease produced by larval Schistosoma mansoni facilitates infection establishment and maintenance in the snail host by interfering with immune cell function. *PLOS Pathogens*, *14*(10), e1007393. https://doi.org/10.1371/journal.ppat.1007393

Johnson, G. D., Swenson, H. R., Ramage, R., & Ahn, K. (2002). Mapping the Active Site of Endothelin-Converting Enzyme-1 through Subsite Specificity and Mutagenesis Studies: A Comparison with Neprilysin. *Archives of Biochemistry and Biophysics*, *398*(2), 240–248. https://doi.org/10.1006/abbi.2001.2708

Jurberg, A. D., Gonçalves, T., Costa, T. A., de Mattos, A. C. A., Pascarelli, B. M., de Manso, P. P. A., Ribeiro-Alves, M., Pelajo-Machado, M., Peralta, J. M., Coelho, P. M. Z., & Lenzi, H. L. (2009). The embryonic development of Schistosoma mansoni eggs: Proposal for a new staging system. *Development Genes and Evolution*, *219*(5), 219–234. https://doi.org/10.1007/s00427-009-0285-9

Karamysheva, A. F. (2008). Mechanisms of angiogenesis. *Biochemistry (Moscow)*, *73*(7), 751–762. https://doi.org/10.1134/S0006297908070031

Kawanabe, Y., & Nauli, S. M. (2011). Endothelin. *Cellular and Molecular Life Sciences*, *68*(2), 195–203. https://doi.org/10.1007/s00018-010-0518-0

Knudsen, G. M., Medzihradszky, K. F., Lim, K.-C., Hansell, E., & McKerrow, J. H. (2005). Proteomic Analysis of *Schistosoma mansoni* Cercarial Secretions. *Molecular & Cellular Proteomics*, *4*(12), 1862–1875. https://doi.org/10.1074/mcp.M500097-MCP200

Kulkarni, M. M., Jones, E. A., McMaster, W. R., & McGwire, B. S. (2008). Fibronectin Binding and Proteolytic Degradation by Leishmania and Effects on Macrophage Activation. *Infection and Immunity*, *76*(4), 1738–1747. https://doi.org/10.1128/iai.01274-07

Mooee, D. V., & Sandgeound, J. H. (1957). *The Relative Egg producing Capacity of Schistosoma mansoni and Schistosoma japonicum*. *5*, 831–840.

Nelwan, M. L. (2019). Schistosomiasis: Life Cycle, Diagnosis, and Control. *Current Therapeutic Research*, *91*, 5–9. https://doi.org/10.1016/j.curtheres.2019.06.001

Olivier, M., Atayde, V. D., Isnard, A., Hassani, K., & Shio, M. T. (2012). *Leishmania* virulence factors: Focus on the metalloprotease GP63. *Microbes and Infection*, *14*(15), 1377–1389. https://doi.org/10.1016/j.micinf.2012.05.014

Peterkova, K., Konecny, L., Macháček, T., Jedličková, L., Winkelmann, F., Sombetzki, M., & Dvořák, J. (n.d.). *Winners vs. Losers: Schistosoma mansoni intestinal and liver eggs exhibit striking differences in gene expression and immunogenicity*.

Peterkova, K., Vorel, J., Ilgova, J., Ostasov, P., Fajtova, P., Konecny, L., Chanova, M., Kasny, M., Horn, M., & Dvorak, J. (2023). Proteases and their inhibitors involved in Schistosoma mansoni egg-host interaction revealed by comparative

transcriptomics with Fasciola hepatica eggs. *International Journal for Parasitology*, *53*(5–6), 253–263. https://doi.org/10.1016/j.ijpara.2022.12.007

Polymerase Chain Reaction (PCR). (n.d.). Retrieved February 24, 2024, from https://www.ncbi.nlm.nih.gov/probe/docs/techpcr/

Protein. (n.d.). Retrieved February 23, 2024, from https://www.genome.gov/genetics-glossary/Protein

Qokoyi, N. K., Masamba, P., & Kappo, A. P. (2021). Proteins as Targets in Anti-Schistosomal Drug Discovery and Vaccine Development. *Vaccines*, *9*(7), 762. https://doi.org/10.3390/vaccines9070762

Reverse Transcriptase—An overview | ScienceDirect Topics. (n.d.). Retrieved February 24, 2024, from https://www.sciencedirect.com/topics/neuroscience/reverse-transcriptase

Reverse Transcription (cDNA Synthesis) | NEB. (n.d.). Retrieved February 22, 2024, from https://www.neb.com/en/applications/cloning-and-synthetic-biology/dna-preparation/reverse-transcription-cdna-synthesis

Rio, D. C., Ares, M., Hannon, G. J., & Nilsen, T. W. (2010). Purification of RNA using TRIzol (TRI reagent). *Cold Spring Harbor Protocols*, *2010*(6), pdb.prot5439. https://doi.org/10.1101/pdb.prot5439

Sarvel, A. K., Kusel, J. R., Araújo, N., Coelho, P. M. Z., & Katz, N. (2006). Comparison between morphological and staining characteristics of live and dead eggs of Schistosoma mansoni. *Memórias Do Instituto Oswaldo Cruz*, *101*, 289–292. https://doi.org/10.1590/S0074-02762006000900045

Sato, Y. (2004). *Role of Aminopeptidase in Angiogenesis*. *27*(6).

Schistosomiasis. (n.d.). Retrieved February 13, 2024, from https://www.who.int/news-room/fact-sheets/detail/schistosomiasis

Schlagenhauf, E., Etges, R., & Metcalf, P. (1998). The crystal structure of the *Leishmania major* surface proteinase leishmanolysin (gp63). *Structure*, *6*(8), 1035–1046. https://doi.org/10.1016/S0969-2126(98)00104-X

Schwartz, C., & Fallon, P. G. (2018). Schistosoma "Eggs-Iting" the Host: Granuloma Formation and Egg Excretion. *Frontiers in Immunology*, *9*.

https://www.frontiersin.org/journals/immunology/articles/10.3389/fimmu.2018.02492

Takaki, K. K., Rinaldi, G., Berriman, M., Pagán, A. J., & Ramakrishnan, L. (2021). Schistosoma mansoni Eggs Modulate the Timing of Granuloma Formation to Promote Transmission. *Cell Host & Microbe*, *29*(1), 58-67.e5. https://doi.org/10.1016/j.chom.2020.10.002

Vondráček, O., Mikeš, L., Talacko, P., Leontovyč, R., Bulantová, J., & Horák, P. (2022). Differential proteomic analysis of laser-microdissected penetration glands of avian schistosome cercariae with a focus on proteins involved in host invasion. *International Journal for Parasitology*, *52*(6), 343–358. https://doi.org/10.1016/j.ijpara.2021.12.003

What are proteins and what do they do?: MedlinePlus Genetics. (n.d.). Retrieved February 23, 2024, from https://medlineplus.gov/genetics/understanding/howgeneswork/protein/

Wright, C. A. (1967). The Schistosome Life-Cycle. In F. K. Mostofi (Ed.), *Bilharziasis: International Academy of Pathology · Special Monograph* (pp. 3–7). Springer. https://doi.org/10.1007/978-3-642-49777-3_2