

Purification and characterization of
proteins involved in metronidazole
resistance in *Giardia lamblia*

Master thesis in Pharmacy
Tiril Øyvor Pedersen



Centre for Pharmacy,
Department of Biomedicine and
Department of Clinical Science

University of Bergen
May 2018

ACKNOWLEDGEMENTS

This thesis was done in collaboration with Inari Kursula lab at the Department of Biomedicine, Bergen Giardia research group at the Department of Clinical Science and Centre for Pharmacy, University of Bergen. The laboratory work of this thesis was completed between September 2017 and April 2018 at Inari Kursula lab, laboratory D at the Department of Biomedicine.

It has been an educational and challenging year. I have gained insights into and experienced the scientific world. I have faced adversity and seen progress.

First and foremost, I would like to give a huge thank to my main supervisor, Dr. **Juha Kallio**, for all help and willingness to share his knowledge. I would also like to give a huge thank to my two co-supervisors, Dr. **Kurt Hanevik** and **Christina Skår Saghaug**. Next I would like to thank Prof. **Inari Kursula** and the people from her research group, especially the people at laboratory D, at the Department of Biomedicine for helping me at the laboratory. **Ju Xu** has been very helpful during my laboratory work. I would also like to thank **Erik Hallin** for performing SAXS and SEC-MALS experiment at the Diamond Light Source. I could not have done this kind of work, if it hadn't been for you all!

Through the toughest and most stressful times I found support and comfort in my boyfriend, family and friends. I would like to thank my boyfriend, **Per Thomas Byrkjedal**, for always supporting, comfort and believe in me. I would like to thank **Elisabeth Gåseland Egeland** for nice lunch breaks. Finally I would like to thank **Christina Lam, Erik A. Guldborg, Viola Lekve, Tonje Sevland** and the rest of the pharmacy students in my class for five memorable years.

ABBREVIATIONS

AA	Amino acid
ALA	Aminolevulinic acid
CD	Circular Dichroism
DNA	Deoxyribonucleic acid
DTT	Dithiothreitol
DLS	Dynamic light scattering
DMSO	Dimethyl sulfoxide
<i>E.coli</i>	<i>Escherichia coli</i>
EF1- γ	Elongation factor 1- γ
FAD	Flavin adenine dinucleotide
Fd	Ferredoxin
FDP	flavodiiron protein
Fe-S	Iron-sulfur
FlHb	Flavo-hemoglobin
FMN	flavin mononucleotide
FT	Flow Through
g	gram
gFlHb	Giardia flavo-hemoglobin
h	hour
IPTG	Isopropyl β -D-1-thiogalactopyranoside
ITB	Inoue transformation buffer
kDa	Kilodalton
KpnI	Enzyme from <i>Klebsiella pneumoniae</i>
LB	Luria-Bertani broth
LS	Light scattering
M	Molar
mAU	milli Absorbance Unit
mg	Milligram
mL	Milliliter
mM	Millimolar
MS	Mass Spectrometry

MTZ	Metronidazole
NAD	nicotinamide adenine dinucleotide
Ni-NTA	Nickel- nitrilotriacetic acid
ng	Nanogram
nL	Nanoliter
nm	Nanometer
NMR	Nuclear magnetic resonance
NO	Nitric oxide
NR1	Nitroreductase 1
NR2	Nitroreductase 2
PAGE	Polyacrylamid Gel Electrophoresis
PBS	Phosphate-buffered saline
PCR	Polymerase chain reactor
PFOR	Pyruvate flavodoxin/ferredoxin oxidoreductase
PMF	Peptide mass fingerprinting
RI	Refractive index
rpm	Revolutions per minute
Rg	Radius of gyration
SAXS	Small Angle X-ray Scattering
SDS	Sodium Dodecyl Sulphate
sec	Seconds
SEC	Size exclusion chromatography
SLS	Static Light Scattering
SOB	Super optimal broth
TAE	Tris-acetate-EDTA
TB	Terrific broth
TEV	Tobacco etch virus
T _m	Melting temperature
TrxR	Thioredoxin reductase
μL	microliter

ABSTRACT

Giardia lamblia is a common intestinal protozoan parasite, causing the gut infection giardiasis. In humans, the infection may vary from asymptomatic carrier stage to a more severe malabsorption syndrome or chronic sequelae. The first line treatment in Norway is metronidazole (MTZ). Recently, MTZ-refractory cases of giardiasis have increased. The mechanism of MTZ-resistance in *Giardia* is not well understood.

This study examined the flavohemoglobin (FIHb) from the two human-infective *Giardia* assemblages (A and B). FIHb is an enzyme thought to play a major role in nitric oxide (NO) detoxification in *Giardia*. In addition to FIHb, the enzymes nitroreductase 1 (NR1) and nitroreductase 2 (NR2) from assemblage B, which are thought to activate and inactivate MTZ respectively, and also be involved in MTZ resistance, were examined.

FIHb from assemblage A was expressed using *E.coli* as host cells, and purified using the Ni-NTA affinity purification system, followed by size exclusion chromatography (SEC). Structural characterization and stability tests including circular dichroism (CD), thermofluor, light scattering (LS), small angle X-ray scattering (SAXS) were performed for FIHb from assemblage A. In addition, screening for crystallization conditions was done to get crystals for crystallographic studies. NR1 and FIHb from assemblage B were cloned and expressed using *E.coli* as host cells, and initial purification trials were done, but time did not allow further research of these. NR2 was not straightforward to clone, and the time limit put an end for further attempts.

From our results, it was observed that FIHb from assemblage A separated into two different variants during SEC. One of the variants containing heme, and one variant without heme. Comparison from structural characterization suggested that the heme-less variant has a more open structural shape and behaved like a dimer. The heme-containing variant, which is the active variant of interest, is monomeric and has a closed, more globular shape.

In conclusion, further research needs to be performed and crystals need to be obtained for NR1, NR2 and FIHb in order to obtain more information about the protein functions and to better understand their role in MTZ resistance in *Giardia*.

TABLE OF CONTENTS

ACKNOWLEDGEMENTS	III
ABBREVIATIONS	IV
ABSTRACT	VI
TABLE OF CONTENTS	VII
1. INTRODUCTION	1
1.1 The parasite <i>Giardia lamblia</i>	1
1.1.1 Epidemiology.....	1
1.1.2 Biology of <i>Giardia</i>	1
1.1.3 Assemblages of <i>Giardia</i>	3
1.1.4 The life cycle of <i>Giardia</i>	3
1.2 The disease giardiasis	4
1.2.1 Treatment of giardiasis	5
1.2.2 Metronidazole – the main drug	5
1.3 Drug resistance	7
1.4 Oxidative and nitrosative stress	8
1.4.1 Oxygen detoxification.....	9
1.4.2 Nitric oxide detoxification.....	9
1.5 Proteins	9
1.5.1 Enzymes.....	10
1.5.2 Flavohemoglobin	10
1.5.3 Nitroreductase 1 and 2 are responsible for metronidazole activation/inactivation.....	12
1.6 Production of <i>Giardia</i> proteins	12
1.6.1 DNA cloning	12
1.6.2 Recombinant protein expression in <i>Escherichia coli</i>	13
1.7 Characterization of proteins	14
1.7.1 Circular dichroism.....	14
1.7.2 Static light scattering	15
1.7.3 Thermofluor.....	15
1.7.4 Small angle X-ray scattering.....	16
1.7.5 Protein crystallization.....	16
1.8 Aims of the master project	17
2. MATERIALS	18
2.1 Strains and vectors	18
2.1.1 Strains.....	18
2.1.2 Vectors	18
2.2 Sequences of the proteins of interest	21
2.2.1 Flavohemoprotein (assemblage A).....	21
2.2.2 Flavohemoprotein (assemblage B)	22
2.2.3 Nitroreductase 1 (assemblage B).....	24
2.2.4 Nitroreductase 2 (assemblage B).....	24
2.3 Equipment	25
2.3.1 Tubes.....	25
2.3.2 Pipettes	26
2.3.3 Kits	26
2.3.4 Columns.....	27
2.3.5 Other consumables	27

2.3.6 Buffers.....	28
2.3.7 Protocols.....	28
2.3.8 Instruments.....	29
2.3.9 Software.....	30
3. METHODS.....	31
3.1 Preparing the competent cells (TOP10).....	31
3.1.1 Transformation into competent cells.....	31
3.1.2 Plasmid isolation.....	32
3.2 Cloning.....	32
3.2.1 Primer design.....	32
3.2.2 Polymerase chain reaction.....	33
3.2.3 Linearization of the vector.....	34
3.2.4 T4 DNA polymerase exonuclease treatment.....	35
3.2.5 Sequencing.....	37
3.3 Protein expression.....	37
3.3.1 Expression of flavohemoglobin from assemblage A.....	37
3.3.2 Expression of flavohemoprotein138.....	38
3.3.3 Expression of nitroreductase 1.....	38
3.4 SDS-PAGE.....	39
3.5 Protein purification.....	39
3.5.1 Ni-NTA affinity chromatography.....	39
3.5.2 Size exclusion chromatography.....	40
3.6 Concentrating the proteins.....	41
3.7 His-tag cleavage.....	42
3.7.1 His-tag cleavage using TEV.....	42
3.7.2 His-tag cleavage using thrombin.....	43
3.7.3 His-tag cleavage using 3C protease.....	43
3.8 Characterization of protein.....	43
3.8.1 Peptide mass fingerprinting.....	43
3.8.2 Multi angle light scattering.....	44
3.8.3 Circular dichroism.....	44
3.8.4 ThermoFluor.....	45
3.8.5 Small angle x-ray scattering.....	45
3.9 Protein crystallization.....	46
4. RESULTS.....	48
4.1 Cloning.....	48
4.1.1 Primer design.....	48
4.1.2 Cloning and transformation into competent cells.....	48
4.1.3 Sequencing.....	51
4.2 Protein expression.....	51
4.2.1 Expression of flavohemoglobin assemblage A with pET14b-gFlHb.....	51
4.2.2 Expression of flavohemoglobin assemblage A with pJ401-gFlHb.....	52
4.2.3 Expression of flavohemoglobin138.....	53
4.2.4 Expression of nitroreductase 1.....	54
4.3 Purification.....	55
4.3.1 Affinity purification of flavohemoglobin from assemblage A.....	55
4.3.2 Affinity purification of flavohemoglobin138.....	56
4.3.3 Affinity purification of nitroreductase 1.....	57
4.3.4 Size exclusion chromatography.....	58
4.4 His-tag cleavage.....	59
4.5 Characterization of flavohemoglobin assemblage A.....	59
4.5.1 Light scattering.....	59

4.5.2 Circular dichroism.....	61
4.5.3 Thermoflour.....	63
4.5.4 Small angle X-ray scattering.....	64
4.6 Protein crystallization	67
5. DISCUSSION.....	68
5.1 Genes used in this study.....	70
5.1.1 Flavohemoglobin	70
5.1.2 Nitroreductases	71
5.1.3 Other metronidazole activating and O ₂ detoxifying enzymes have been characterized	71
5.2 Cloning.....	72
5.3 Expression.....	72
5.3.1 Time, temperature and obtained protein concentration	72
5.3.2 Expression with addition of IPTG	73
5.3.3 Cell lysis	73
5.4 Purification	74
5.4.1 Affinity purification – binding to the resin	74
5.4.2 Affinity purification – chemical environment.....	75
5.4.3. Freeze the purified protein	75
5.5 His-tag cleavage	76
5.6 Colorful and colorless variants of flavohemoglobin from assemblage A.....	77
5.6.1 Addition of heme	78
5.7 Exposure to oxygen	78
5.8 Crystallization.....	79
5.8.1 Optimization of crystal conditions.....	79
5.8.2 Homogeneity	79
5.8.3 Obtained flavohemoglobin crystals in literature	80
6. CONCLUSION.....	80
6.1 Flavohemoglobin	80
6.2 Nitroreductases.....	81
7. FURTHER RESEARCH.....	81
REFERENCES	83
APPENDIX 1.....	VII

1. INTRODUCTION

1.1 The parasite *Giardia lamblia*

1.1.1 Epidemiology

Giardia lamblia (also known as *Giardia duodenalis* and *Giardia intestinalis*, from now on *Giardia*) belongs to a group of binucleate organisms named diplomonads (1, 2). It is the most common intestinal protozoan parasite infecting humans in both developing (3) and developed countries (4). Due to poor hygiene and sanitation, infection with *Giardia*, giardiasis, is more prevalent in developing countries, where children often are affected. It has been reported that approximately 15 % of children aged 0-24 months in the developing world are infected (5). Giardiasis is further a contribution to the global problem of enteric infections, which is one of the most important causes of morbidity, mortality and poor linear growth in children under 5 years old (6, 7). Annually, *Giardia* is responsible for up to 280 million infections worldwide (8).

In Norway and other developed countries, giardiasis is mainly considered to be an imported disease (9). In the autumn of 2004, a large outbreak of giardiasis occurred in Bergen, Norway. The outbreak was caused by *Giardia* assemblage B (10), and roughly 1500 people had laboratory confirmed giardiasis (11, 12). Some patients did not respond to metronidazole (MTZ) treatment but were cured with second and third line alternatives. (13). For some people, the infection gave long-term consequences even though they received treatment. Many of the persons exposed to giardiasis during the outbreak, later reported chronic sequelae like irritable bowel syndrome (IBS), chronic fatigue (CF) and food intolerance (14, 15).

1.1.2 Biology of *Giardia*

Antony van Leeuwenhoek was the first to observe and describe the *Giardia* parasite in 1681. It was later named after the French zoologist Alfred Mathieu Giard (16). *Giardia* is a microaerophilic eukaryotic protozoan parasite (17). It is thought to be one of the most primitive extant eukaryotic organism and it is missing mitochondria, nucleoli and peroxisomes (1).

Giardia has two major stages in its life cycle; the trophozoite and the cyst. The trophozoite has a pear-shape, containing two nuclei, cytoskeleton made of microtubules and microtubule-associated proteins that assemble in structures such as the adhesive disc, the median body, the funis and four pair of flagella. By light microscopy, *Giardia* looks like a smiling face, where the eyes are the two nuclei and the mouth is the median body line (8). The structure of the trophozoite and the cyst and their components are presented in Figure 1.1.

The size of the trophozoite measures 12-15 μm in length and 5-9 μm in width. The cyst is about the same size (18). The cyst is the robust form, where the cyst wall and an inner layer consist of two membranes protecting the organism. The adhesive disc and flagella are here disassembled and stored inside. The cyst has four tetraploid nuclei (18).

Giardia is lacking mitochondria, but a different mitochondrial remnant organelle, the mitosome, has some of the same functions. The mitosome is suggested to be involved in the synthesis of iron-sulfur (Fe-S) clusters (19). Fe-S clusters are thought to play important roles in the reduction and oxidation reactions of mitochondrial electron transport, energy metabolism, regulation of gene expression and a number of metabolic pathways (20).

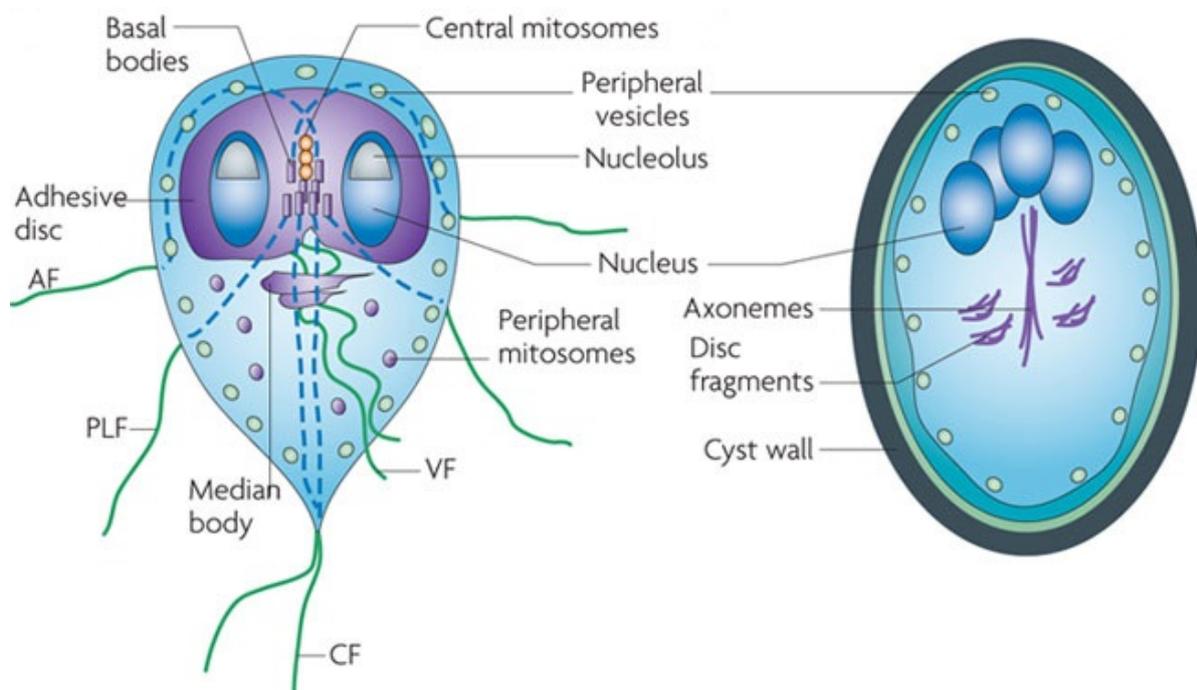


Figure 1.1: Structure of the *Giardia* trophozoite and cyst. The trophozoite is on the left hand side and the cyst is on the right hand side. The flagella are organized in four pairs. AF (anterior flagella), PLF (posterior/lateral flagella), CF (caudal flagella) and VF (ventral flagella). Figure taken from Ankarklev et. al (18).

1.1.3 Assemblages of *Giardia*

Initially, *Giardia* was divided into species based on the host from which the organism was obtained. Later, it was observed that the *Giardia spp.* from different hosts were morphological identical (1). According to this observation, three species were named (*Giardia agilis*, *Giardia lamblia* and *Giardia muris*) on the basis of morphological differences detected by light microscopy (1). Further advances in technology have provided new tools, new techniques such as molecular and electron micrographic have been used to classify *Giardia spp.* In recent times, reclassification and new species can be done by sequence comparison of the small-subunit ribosomal RNA (2). Eight assemblages (A-H) of *G. lamblia* have been identified whereas assemblage A and B infects humans (21, 22). Assemblages are further divided into smaller subdivisions, based on several factors, including comparison of electrophoretic mobility of enzymes and chromosomes, and sequencing of genes (22).

1.1.4 The life cycle of *Giardia*

The life cycle consist of two main processes: excystation (cyst to trophozoite) and encystation (trophozoite to cyst). Outside the host, the parasite is encapsulated in a cyst wall that protects it from hypotonic lysis in the environment (18). In the cyst, the parasite metabolism is downregulated (23). After entering the stomach of the host, stomach acid triggers excystation, resulting in the trophozoite. The trophozoite is the disease-causing stage of the life cycle of *Giardia*. In this stage, the parasite has high mobility because of the flagella and is able to attach strongly to the enterocytes found in the upper small intestine of the host, with its adhesive disc (2). The life cycle is shown in Figure 1.2.

Encystation is the opposite process. This process transforms the trophozoites into the infective cyst. Encystation is triggered by host-specific factors such as high levels of bile, low levels of cholesterol and a basic pH (24). The trophozoite is gradually changing into a cyst by loosing its flagella, building the cyst wall and downregulate its metabolism (18, 23). The cysts will then pass through the last part of the colon and are shed in the feces and may infect other hosts (18). The infective cysts may survive for months in soil or water, waiting for a new host to drink or eat contaminated water or food (25). *Giardia* cysts may be spread through consumption of untreated drinking water or food (25), zoonotic transmission (10), person-to-person contact, and the infective dose is low (10 cysts or less) (26).

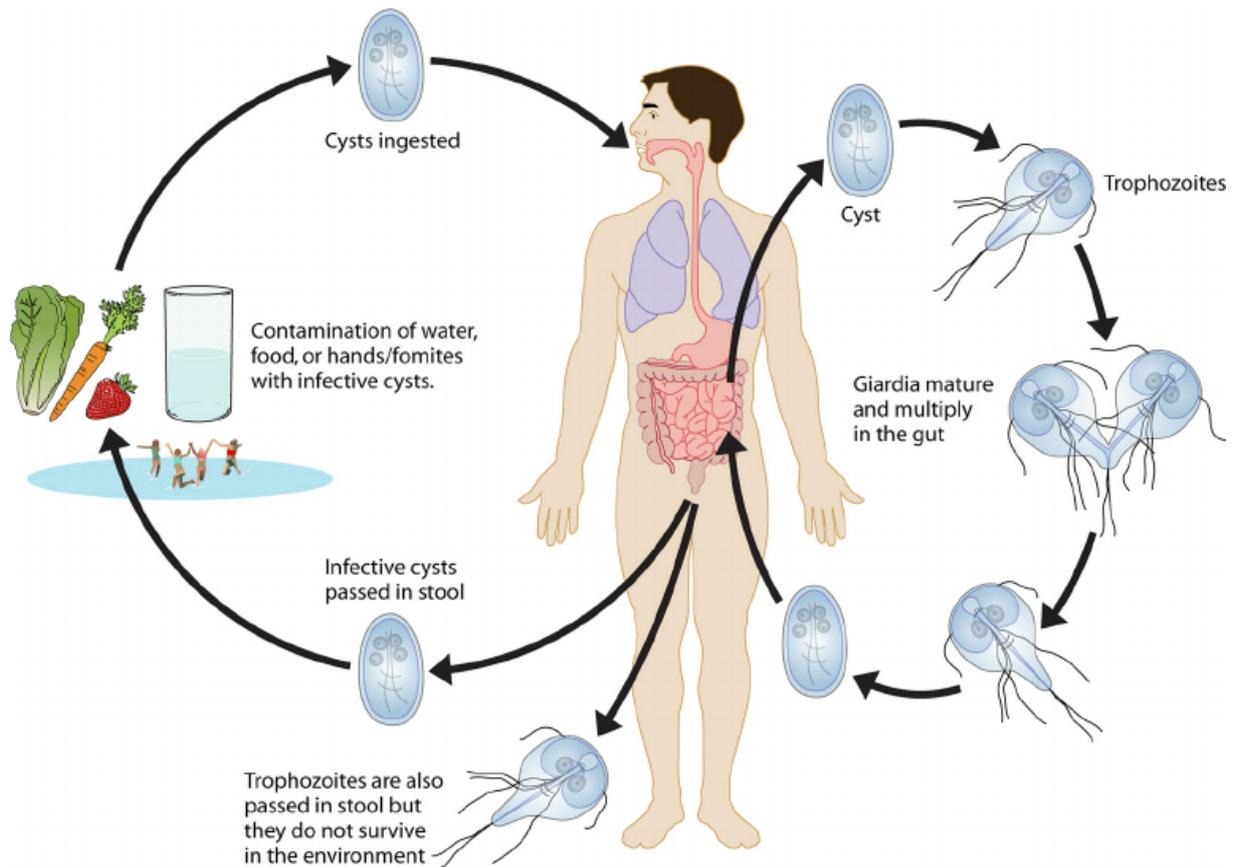


Figure 1.2: The life cycle of *Giardia lamblia*. Humans get infected by cysts from contaminated food/water or other humans. The cysts undergo excystation, and become trophozoites. The trophozoites are causing the disease giardiasis. They can undergo encystation and become cysts again. Cysts and trophozoites are passed out in the stool, but only the cysts can survive outside the host and infect new hosts. Figure taken from Esch et al. (25)

1.2 The disease giardiasis

The infection caused by *Giardia* is named giardiasis. The infection may have two clinical outcomes, symptomatic or asymptomatic (27). The symptomatic parasitic infection is characterized by watery diarrhea, epigastric pain, nausea, vomiting and weight loss, and the clinical effects of the infection may vary from the asymptomatic carrier stage to a more severe malabsorption syndrome or chronic sequelae (28, 29). Symptoms usually appear 6-15 days after infection. Often the infection is asymptomatic (27, 30).

Giardiasis can be divided into two disease phases: acute and chronic. The acute phase is usually short-term, characterized by flatulence and abdominal distension with cramps and frequent and watery diarrhea. Later the feces becomes bulky, greasy and and the patient may

experience offensive odors related to toilet visits (31). The chronic phase is associated with malaise, weight loss and stools are usually yellow/pale, frequent and of small volume.

The disease is diagnosed by antigen detection assays, nucleic acid detection assays and stool microscopy (32). Today, a sensitive and specific method for the detection is often used to diagnose intestinal parasitic infections. The method is real-time PCR, which uses fluorescent detection probes (33).

1.2.1 Treatment of giardiasis

Giardiasis is usually treated with MTZ or other nitroimidazoles (18). MTZ is the first line treatment in Norway and many other countries. MTZ has an estimated effect of 60 - 90 % (27). In cases where MTZ do not show any effect, or is not recommended (such as during pregnancy), other antiparasitic drugs have shown good effect. The drugs that may be used and their treatment efficacy are the following: tinidazole (74-100%), quinacrine (92-100%), albendazole (24-100%) and furazolidone (80-100%). In cases of pregnancy, promomycin (55-90%) is recommended (13). When treatment failure occurs, treatment with a combination of drugs is recommended. The combination of albendazole (or quinacrine) and MTZ has been effective and well tolerated as treatment of MTZ-refractory giardiasis (13).

1.2.2 Metronidazole – the main drug

MTZ is listed as an “essential medicine” by WHO (34), and was developed in the late 1950s against the microaerophilic parasite, *Trichomonas vaginalis*. The drug showed to be effective and was soon tested against other anaerobic and microaerophilic pathogens, including *Giardia* (35).

MTZ is a 5-nitroimidazole prodrug. To become activated, MTZ, needs to be reduced at its nitro group (see Figure 1.3). One reduction mechanism is thought to be caused by electrons donated from the enzyme pyruvate flavodoxin/ferredoxin oxidoreductase (PFOR). This enzyme transfers electrons from pyruvate to ferredoxin (Fd). The reduced Fd is then reoxidized by ferredoxin:NAD oxidoreductase transferring its electrons to NAD(P). Resulting in NAD(P)H, which may then transfer its electrons to oxygen (O₂). The reaction is catalyzed

by NAD(P)H oxidase (36).

Thioredoxin reductase (TrxR) is another enzyme that may activate MTZ using NADPH as an electron donor. A third enzyme, has also been showed to be able to activate MTZ, and is known as nitroreductase 1 (NR1) (35). NR1 has a Fd domain and it could potentially have a similar activating mechanism of MTZ as the PFOR and Fd coupled activation mechanism (37). Figure 1.3 illustrates how the enzymes here may activate MTZ.

Nitroimidazoles may capture electrons directly from the reduced Fds or from the NAD(P)H oxidase, thus yielding toxic radicals causing irreversible damage in the parasite (38). If O₂ is present, the active nitro radical form is converted back to the non-toxic parent compound (21). Because of this, the drug works only under anaerobic or microaerophilic conditions. The cell damaging effects of MTZ include damage of DNA in microaerophiles and anaerobic microorganisms by introducing double strand breaks and degradation of elongation factor 1- γ (EF1- γ) (important in protein translation) (35). In addition, MTZ may cause oxidative stress by forming adducts with cysteine and inhibits the redox enzyme TrxR (35, 39). Cysteine has antioxidant properties and is thought to protect *Giardia* against oxidative stress (40). In addition to activate MTZ, TrxR acts as an antioxidant to protect the parasite against oxidative stress (41). MTZ may also be inactivated and detoxified by an enzyme assumed to be responsible for inactivation, and is known as nitroreductase 2 (NR2) (35). Figure 1.3 illustrates how NR2 potentially detoxifies MTZ.

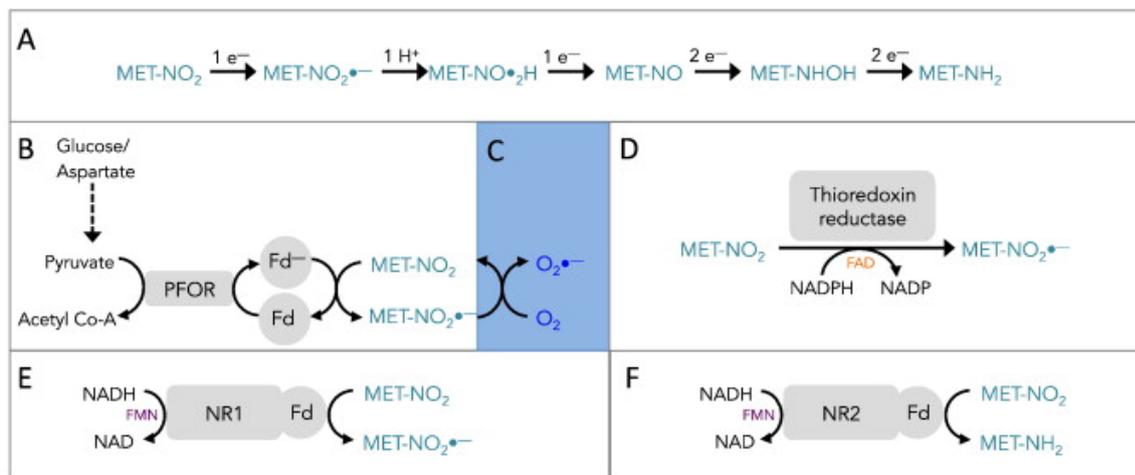


Figure 1.3: Activation and inactivation of metronidazole in *Giardia*. A: Reduction and oxidation of the prodrug (MET-NO₂) resulting in toxic intermediates and the final inert amine (MET-NH₂). B: PFORs role in activation of MTZ C: MTZ may react with oxygen and create free radicals. D: Thioredoxin reductase activates MTZ using NADH as electron donor. E: NR1 activates MTZ using NADH as electron donor. F: NR2 inactivates MTZ by completely reducing the prodrug to the inert amine, using NADH as electron donor. Figure taken from Ansell et. al (37).

1.3 Drug resistance

Drug resistance in *Giardia* is complex and far from fully understood. Recently, treatment-refractory cases of giardiasis have increased and appeared with a high prevalence. In a study from the hospital of tropical diseases in London, England, treatment failure increased from 15.1% in 2008 to 40.2% in 2013 (42). The best described mechanism of resistance is the loss of the parasite's ability to activate 5-nitroimidazole prodrugs to toxic radicals by reduction, thus effectively allowing the parasite to avoid suicidal drug activation (17).

MTZ-resistance has been associated with decrease in activity of the enzymes PFOR, Fd and NR1, by comparing MTZ-susceptible and laboratory induced MTZ-resistant *Giardia* lines (43-46). Some of the MTZ-resistant *Giardia* isolates have exhibited lower PFOR expression levels, meaning that PFOR may be one of the main targets for 5-nitroimidazole in *Giardia* (47). Evidence has however shown that PFOR may not be the only enzyme capable of activating nitro drugs in microaerophilic or anaerobic pathogens (Figure 1.3). Recent studies on *Giardia* clones and MTZ or nitroimidazole resistant strains have revealed that resistance may occur without downregulation of PFOR (48, 49). Although some 5-nitroimidazole drugs seem to interact directly with PFOR, it is unlikely that Fd directly performs the reduction of

the nitro group, as it is more likely that this reaction is catalyzed by NRs (47). It has been suggested that loss of either the PFOR/ferredoxin couple or NR1 can lead to MTZ resistance (44, 47, 50, 51). Upregulation of NR2 may also result in MTZ resistance (37). Two genes, *fdp* and *nadhox* which, respectively, encodes for the enzymes flavodiiron protein (FDP) and NADH oxidase have recently been found to be upregulated in some albendazole resistant clones of the parasite, pointing to a possible linkage between drug resistance and O₂ metabolism in *Giardia* (21). FDP and NADH oxidase may play a major role in O₂ detoxification (52). In essence, resistance in *Giardia* probably constitutes several mechanisms of both downregulation of MTZ activating enzymes, and upregulation of MTZ inactivating enzymes (50).

1.4 Oxidative and nitrosative stress

Giardia trophozoites are attached to the intestinal mucosa of the host, and will thus be exposed to O₂ and nitric oxide (NO). NO is produced by the NO-synthases (NOSs) in the intestinal epithelial cells and derived from reduction of NO₂⁻ to NO (21). It is also produced by the immune system in humans to fight microbial pathogens (51). To survive, the parasite needs to handle both O₂ and NO. *Giardia* has an antioxidant defense system to protect itself (51) from the harmful effects of these two substances (Figure 1.4).

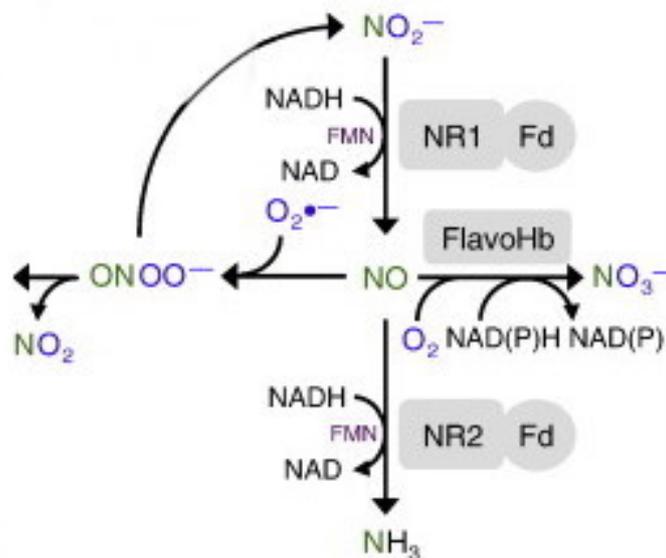


Figure 1.4: O₂ and NO are detoxified in *Giardia*. The cofactor FMN and the enzyme FlHb, NR1 and NR2 with their Fds are all involved in detoxification. Figure cropped and taken from Ansell et. al (37).

1.4.1 Oxygen detoxification

NADH oxidases are flavoenzymes, which catalyze the reduction of O_2 , resulting in the production of $O_2^{\cdot-}$, H_2O_2 or H_2O . NADH oxidase from *Giardia* trophozoites have shown to be able to convert O_2 to H_2O without producing partially reduced oxygen; $O_2^{\cdot-}$. This enzyme uses NADPH or NADH as electron donor (21, 51). Another group of enzymes involved in detoxifying O_2 are FDPs. The cores of these enzymes contain two redox centers: a flavin mononucleotide (FMN) and a non-heme Fe-Fe active site (diiron site). FMN is the electron entry site into the enzyme. Electrons derived from NADH are transported through FMN into the diiron site and catalyzes the full reduction of O_2 to H_2O (51).

1.4.2 Nitric oxide detoxification

One mechanism of NO detoxification is that electrons derived from NADH are transported through FMN into the diiron site and catalyze NO to N_2O . *Giardia* codes for a flavohemoglobin (FIHb), which is a NO-detoxifying bacterial enzyme acquired through lateral transfer from prokaryotic organisms (53). This enzyme is thought to play a major role in the NO detoxification in *Giardia* (51) (see Figure 1.4). FIHb is able to catalyze the degradation of NO to NO_3^- using O_2 as a co-substrate (54). Figure 1.4 illustrates this process. It also illustrates how FMN with help from NADH converts NO to non-toxic compounds.

1.5 Proteins

Proteins are biological macromolecules built from amino acid (AA) residues (55). Proteins are essential in biological processes. They can function as catalysts, transporters, generate movement and control growth and cell differentiation to mention some of the main functions (55). The structure of the protein and its functional groups will decide the function of the protein. The structure of proteins can be divided into levels of structure; primary, secondary, tertiary and quaternary (56). The primary structure is defined as the AA sequence (plus intra- and interchain covalent cross-links if any) in the polypeptide chain (56). The polypeptide chain is flexible and the primary structure can turn into multiple conformations. The main chain conformations are known as the secondary structure. The two main types of secondary structures are α -helices and β -sheets (56). The different combinations of α -helices and β -sheets create the tertiary structure, a three dimensional (3D) structure. In several cases, the

proteins contain more than one subunit, and the assembly of the subunits is called the quaternary structure (56).

The genetic code of a protein is translated from nucleic acid sequences of genes to the AA sequence of the protein. The unit of coding for one AA is a triplet of nucleotides, named a codon (56). The unique AA sequence of a protein determines its 3D structure. The function of a protein is dependent on its structure, which depends on physical and chemical parameters (57). In addition to AA, ions, small organic ligands and water molecules are integral parts of several protein structures (56). Studying proteins, their structure, folding and compositions are important to understand their function and role in biological processes (57).

1.5.1 Enzymes

Proteins are classified based on their functions. Enzymes are in most cases proteins (some are RNA molecules, ribozymes) (58, 59) that work as biological catalysts (59). Enzymes speed up reactions by stabilizing transition states, the highest energy point on a reaction pathway (59). Enzymes binds the transitions state more tightly than they bind substrates to altering the energy barrier between the substrate and the product (56). They are specific to the substrate due to complementary between the enzyme and the substrate, known as the “lock and key” model. Another model is “induced fit”, where the enzyme undergoes conformational changes upon binding to the substrate (60).

1.5.2 Flavohemoglobin

FlHb (also named flavohemoprotein) genes are found in both prokaryotes and eukaryotes and they are recognized by a heme pocket inside (61). The FlHb family belongs to the globin superfamily. All globins are structurally related to hemoglobins (Hbs) and myoglobins (Mbs) with conservation of the heme pocket (62). Even though FlHbs appear in various species, sequence alignments show that the FlHb family is a very homogeneous group of proteins (62). FlHb consists of a heart shape structure with three different domains, the C-terminal NAD-binding, the FAD-binding, and the N-terminal globin domain. The globin domain structure has six helices (A-H) with a long H-helix. The D-helix is substituted by a large loop region. The loop between the C and E helices (CE loop) in the globin domain forms a distinct hydrophilic cavity with a strong electron density (anion binding site). The location of the loop

on the heme edge and the strong electron density indicates that it may interact with the FAD cofactor (63).

The FAD-binding domain consists of a six-stranded antiparallel β -barrel, a small α -helix capping the β -barrel on the bottom, and a long loop connecting β 2, β 3 β -sheets on the top. The loop between β 5 and the α -helix is responsible for the different orientations of the adenosine molecule in space (61). Figure 1.5 is an example of a FIHb, and illustrates where the heme is bound in its pocket, and where FAD is bound.

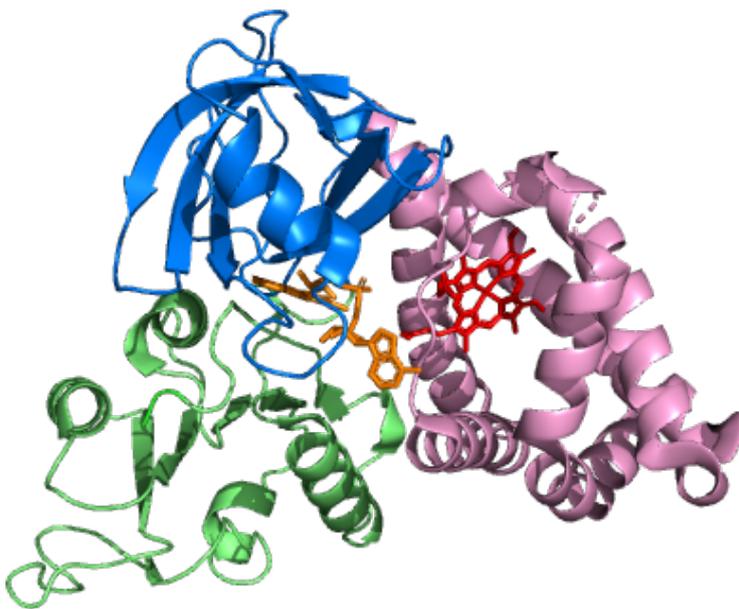


Figure 1.5: Crystal structure of flavohemoprotein from Alcaligenes eutrophus. The globin domain is colored pink, the FAD binding domain is colored blue and the NAD binding domain is colored green. Heme is showed in red inside the heme-pocket, and NAD is showed in orange. The figure is created with PyMOL, using pdb-file 1CQX from RCSB PDB and the article from Erlmer et. Al (64).

The crystal structure of FIHb in *Giardia* has not yet been solved. By solving the crystal structure of the enzyme, it will give more information of the function of the enzyme and possible affect of mutations found in *Giardia* genes. Interestingly FIHb gene from human pathogenic strain, assemblage B, codes for a FIHb that has long extension (about 140 AA) in its N-terminus with unknown function.

1.5.3 Nitroreductase 1 and 2 are responsible for metronidazole activation/inactivation

The polypeptide sequence of NR1 is rather similar to that one of NR2. Both proteins contain a Fd domain with two 4Fe-4S-clusters at their N-terminus and a FMN-reductase domain at their C-terminus (38). The Fd domain found in NR1 is thought to act as a second electron donor for the reduction of nitrodrugs, i.e. MTZ (43). The activation may partially reduce nitro compounds to toxic intermediates (46).

NR2 on the other hand, decreases the susceptibility to MTZ and detoxify nitro compounds, and could thus acts as an inactivator (38, 47). A study (38) observed that under both aerobic and semi-aerobic conditions, bacteria expressing giardial NR1 alone had a significantly higher susceptibility to nitrodrugs than control bacteria. Bacteria expressing NR2 alone, or both nitroreductases, were completely resistant to MTZ (38).

NRs have been identified in *Giardia* (GL50803_22677 from WB clone C6, and GL50803-6175). NRs from other anaerobic or microaerophilic pathogens (e. g *Helicobacter pylori*) nitroreductases are well documented as resistance factors (65-67). The biological role of NR1 and NR2 may be involved in the reduction of vitamin K analogues and FAD (38). To be able to understand the MTZ resistant mechanism of *Giardia*, further investigations are needed. The crystal structures of NRs from *Giardia* have not yet been solved. By solving the crystal structure, more information about the enzymes structure function relations can be obtained.

1.6 Production of *Giardia* proteins

1.6.1 DNA cloning

To be able to characterize proteins, they first need to be expressed using for example bacteria as host cells, and further solubilized and purified. Before the protein can be expressed, the gene of the protein of interest (GOI) needs to be inserted into a vector by cloning. The basic principle of DNA cloning is that the DNA of interest is inserted to a cloning vector (also known as plasmid), which is further incorporated into cultured host cells. The term “gene cloning” includes both *in vivo* and *in vitro* cloning. Gene cloning *in vivo* involves the use of restriction enzymes and ligases using vectors and then clone the recombinant DNA into host cells. For *in vitro* gene cloning, the polymerase chain reaction (PCR) method is used to create copies (amplify) of DNA (68).

For cloning *in vivo*, a restriction enzyme is required. There are two types of restriction enzymes: “blunt end cutters” and “sticky end cutter”. The “blunt end cutters”, cut both the strand of the target DNA at the same spot creating blunt ends. Sticky end cutters cut both strand of the target DNA at different spots creating 3'- or 5'-overhangs of one to four nucleotides. This are called “sticky ends” (69).

Plasmids are circular, double-stranded DNA molecules that are separate from a cell's chromosomal DNA. The circular DNA occurs naturally in bacteria, yeast and also in some higher eukaryotic cells. They may exist in a parasitic or symbiotic relationship with their host cells. Plasmids are, like chromosomal DNA, duplicated before every cell division. They have its independent replication system, which controls the number of copies in a cell. High copy numbers gives a higher chance of the daughter cells to contain the plasmid. Low copy numbers causes less metabolic burden on their host, but they have a higher probability of loss during cell division (70). The plasmids encode for an antibiotic resistance marker, which allows the cells containing the plasmid to grow, and kill the plasmid-free cells in a growth medium containing the selective antibiotic (71). A cloning vector is a genome that can carry and accept DNA. It can increase the number of copies of the DNA of interest through its own replication. Vectors are plasmids, and are selected based on the insert size, copy number, cloning site, antibiotic resistance and other desired properties important for the use (72).

1.6.2 Recombinant protein expression in *Escherichia coli*

The choice of host cell for protein expression depends on different factors. The host systems that are available include bacteria, yeast, filamentous fungi, unicellular algae and eukaryotic cells (insect and mammalian) (73). They all have strengths and weaknesses that make them suitable for different proteins (74).

Escherichia coli (*E.coli*) is a common host organism for recombinant protein expression. The advantage of using *E. coli* is that it grows fast and it is cost effective. In glucose-salt media and optimal environmental conditions, its replication time is about 20 minutes and plasmid transformation of *E.coli* can be performed in as little as 5 min. Plasmids can be put together by combinations of replicons, promoters, selection markers, multiple cloning sites, and fusion protein/fusion protein removal strategies. The replicon consists of one origin of replication together with its associated *cis*-acting control elements. Promoter sequences are DNA

sequences that define where transcription of a gene by RNA begins. The T5 promoter is recognized by *E.coli* RNA polymerase (75). Next to the T5 promoter, a *lac* operator (*lacO*) is inserted to control the expression. The expression can be turned on by the adding the inducer isopropyl β -D-1-thiogalactopyranoside (IPTG), which binds to *lacO* and turns on the transcription (74). The T7 promoter is recognized by the phage T7 RNA polymerase (T7 RNAP) (74). The system can be induced by IPTG under the control of *lacO*, or the expression can be leaky (74). The T7 promoter can be plain as well, meaning that there are no *lacO* site controlling the expression (76). In this case, the expression is leaky, meaning that there is some basal level of protein expression during cell growth (74).

Normal *E. coli* cells cannot take up plasmid DNA from the environment (77) and the use of chemical treated *E. coli* cells, termed competent cells, are therefore required. Competent cells are more permeable for DNA uptake (78).

1.7 Characterization of proteins

Protein sequencing has earlier been the only method used for elucidating the primary structure of proteins. Today, DNA sequencing is used (79). The secondary and tertiary structure of the protein is more complicated and often involves multiple techniques. Techniques for finding the molecular weight (Mw) of the protein include sodium dodecyl sulphate polyacrylamide gel electrophoresis (SDS-PAGE) and other electrophoresis techniques, mass spectrometry (MS) and static light scattering (SLS). In order to gain information about the structure of the protein, techniques for measuring the ratio of secondary structure including α -helices and β -sheets with circular dichroism (CD) and techniques for finding the 3D structure such as crystallography, can be used (56).

1.7.1 Circular dichroism

Circular dichroism (CD) is a good method for rapidly evaluating the secondary structure, folding and binding properties of proteins. It is a form of light absorption spectroscopy that measures the difference in absorbance of right- and left-circularly polarized light by a substance (80). A CD signal will occur when a chromophore is chiral (optically active) (81). Different structural elements have characteristic CD spectra, which can be recognised from

the measured protein solution of interest. α -helical proteins have negative bands at 222 nm and 208 nm, and a positive band at 193 nm. β -sheets can be recognized by negative bands at 218 nm and positive bands at 195 nm (80). By comparing the CD spectra of the protein of interest with the known CD spectra of α - β - and random coiled proteins, the secondary structure of the protein of interest can be estimated (80).

1.7.2 Static light scattering

Static light scattering (SLS) is a technique used for measuring the molecular weight of a protein in solution. It can be used as a quality control to estimate that it is the right protein in the solution, in addition to determine if the protein is monomeric or not (82). The principle is based on the relationship between the intensity of light scattered by a molecule and its molecular weight and size (83). The intensity of the light is measured as a function of the scattering angle, θ . The measured intensity is divided by the intensity of the incident laser beam, which is measured by a reference detector (84). If the system uses multiple detector in different angles we are talking about multi-angle light scattering (MALS). For a pure protein solution, it is expected that all of the protein molecules within the sample should have the same molecular weight. SLS is often performed together with size exclusion chromatography (SEC) to combine SLS with a purification step (82, 85, 86). The light scattering and concentration are measured for each elution fraction, resulting in an independent determination of mass and size at the elution position (82). Concentration of the protein sample during the experiment is measured using protein absorption at 280 nm or refracting index (RI).

1.7.3 Thermofluor

Thermofluor (Differential scanning fluorimetry) can be used to estimate the apparent melting temperature (T_m) of the protein of interest. When the temperature is low, the protein is completely and correctly folded and no hydrophobic areas are exposed. When the temperature starts to rise, the protein starts to unfold and hydrophobic areas become exposed. A fluorescent chemical compound (fluorophore) can now bind to these areas and fluorescence occurs. The fluorescence can be plotted as a function of temperature and form a curve that can be used to estimate the melting temperature of the protein by finding the midpoint of the transition curve (87).

1.7.4 Small angle X-ray scattering

Small angle X-ray scattering (SAXS) is a high-resolution characterization technique (88). This method can be used to study the overall shape and structural transitions of biological macromolecules in solution (89). The principle of SAXS is that a X-ray beam is used to illuminate the protein in solution, and the scattered radiation is registered by a detector (90). The scattering signal from the solvent (buffer) must be subtracted from the protein solution, to only get the scattered signal only from the protein. From the scattered data, the molecular weight, shape and model for the overall structure of the protein can be obtained (91).

1.7.5 Protein crystallization

The first protein crystal was observed by Hünefeld in 1840 (92). The protein was hemoglobin from earthworm. This observation made it clear that protein crystals could be obtained by controlled evaporation of concentrated protein solutions. In other words, protein crystals could be produced by slow dehydration (92).

To study the 3D structure of proteins, there are two main techniques: X-ray crystallography and nuclear magnetic resonance (NMR) spectroscopy. X-ray crystallography requires protein in crystal form, where all protein molecules are lined up in the same orientation (59).

Crystallization is a complicated process, where the parameters have to be individualized. First the factors that will make the crystals have to be found, then the individual variables need to be optimized to achieve the best possible crystals. When screening for crystallization conditions, systematic variation of different important variables such as temperature, pH and concentration need to be tested (93).

Crystals can only be grown from a supersaturated solution. Supersaturation is a state where a solution contains more of the dissolved material than could be dissolved by the solvent under normal circumstances. Supersaturation will occur under specific chemical and physical conditions, and the protein solution needs to be pure and homogeneous (93). Different methods can be used during the crystallization process, including *the sitting-drop vapor-diffusion method*, *the hanging-drop vapor-diffusion* and microbatch crystallization (93, 94).

1.8 Aims of the master project

The aim of this study is to clone, express and purify the two enzymes NR1 and NR2, which are involved in MTZ resistance, and the oxidative stress management enzyme FIHb from the pathogenic parasite *Giardia lamblia*. In addition, structural studies of FIHb should be performed in order to characterize the protein structure of FIHb. The structural studies include prediction of the secondary structure, stability tests and measure the size of the protein. Crystallization trials will be set up in order to obtain crystals, which further can be used to obtain the 3D structure of FIHb using macromolecular crystallography.

2. MATERIALS

2.1 Strains and vectors

2.1.1 Strains

E. coli Rosetta (DE3) (competent cells prepared by Ju Xu)

E. coli TOP10 (competent cells prepared by Tiril Ø. Pedersen)

The strains are obtained from the technician, Ju Xu, at laboratory D at the Department of Biomedicine, University of Bergen.

2.1.2 Vectors

pET14b-gFIHb is a low copy vector (<5 copies per cell) that was used as an expression vector for wild type *Giardia* FIHb (gene ID: GL50803_15009 from *Giardia* assemblage A1 (WB)). It has an N-terminal histidine tag (His₆-tag) that can be cleaved by thrombin. *E. coli* Rosetta (DE3) is used as an expression host for this gene. The plasmid codes for ampicillin resistance, and uses a plain T7 promoter (no lac operator site) so it does not require IPTG induction. When a plain T7 promoter is used, the transcription is leaky, meaning that good expression of heme-bound protein is obtained without the need of IPTG induction (e.g. use of a *lacUV5*). See Figure 2.1 for the vector map of pET14b.

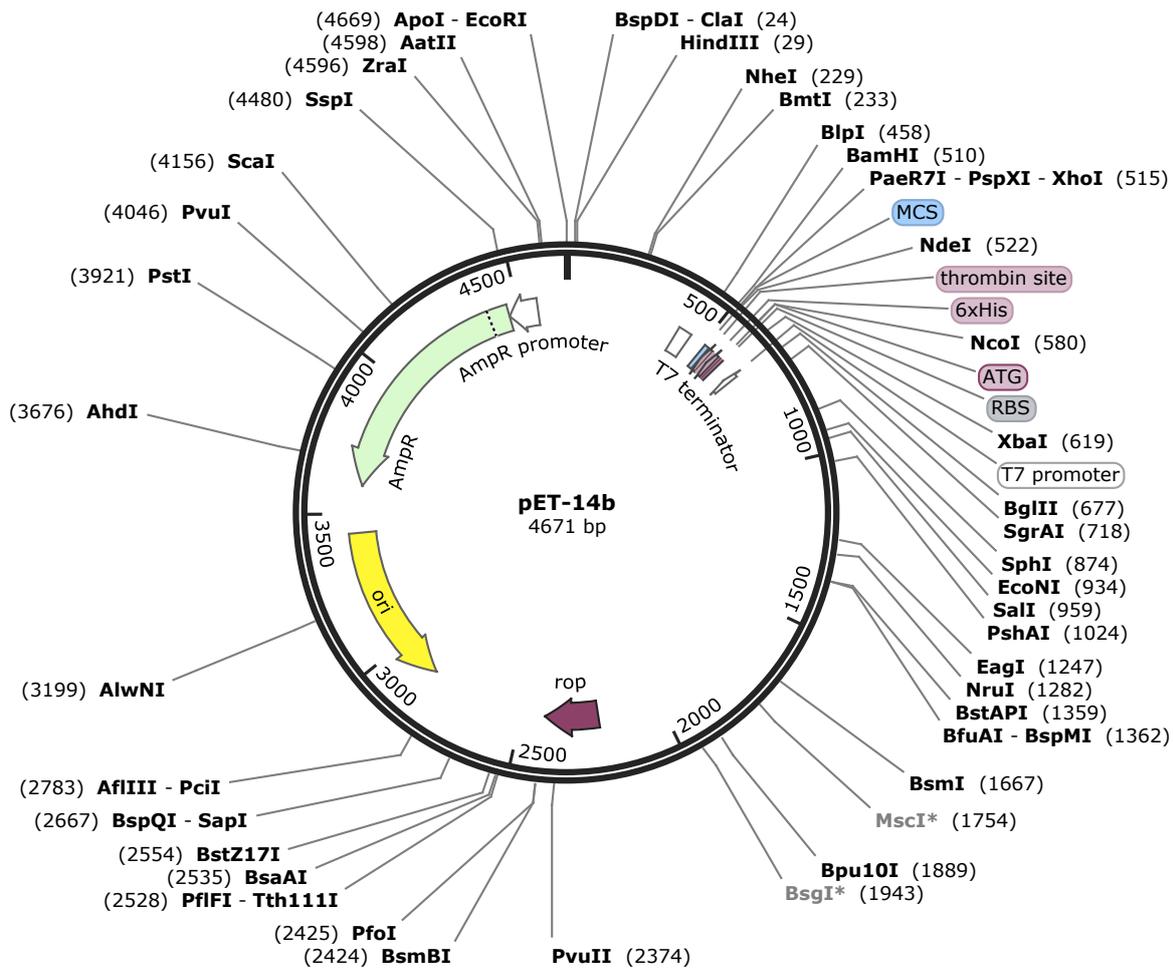


Figure 2.1: Map of the vector pET14b. Created with SnapGene® Viewer.

pJ401-gFIHb is a high copy vector (>100 copies per cell) that was another vector used for expression of wild type *Giardia* FIHb (gene ID: GL50803_15009 from *Giardia* assemblage A1 (WB)). It has an N-terminal His₆-tag that can be cleaved by tobacco etch virus (TEV). In this project Rosetta (DE3) was used as the expression host. The plasmid codes for kanamycin resistance, uses the T5 promoter and requires IPTG induction. Aminolevulinic acid (ALA) may help to obtain more heme-bound protein, as it is a heme precursor. When IPTG induction is used, the expression rate is faster than with leaky expression, and it might be that the heme synthesis is not keeping up when IPTG induction is required. ALA is therefore added when using this vector. See Figure 2.2 for the vector map of this vector.

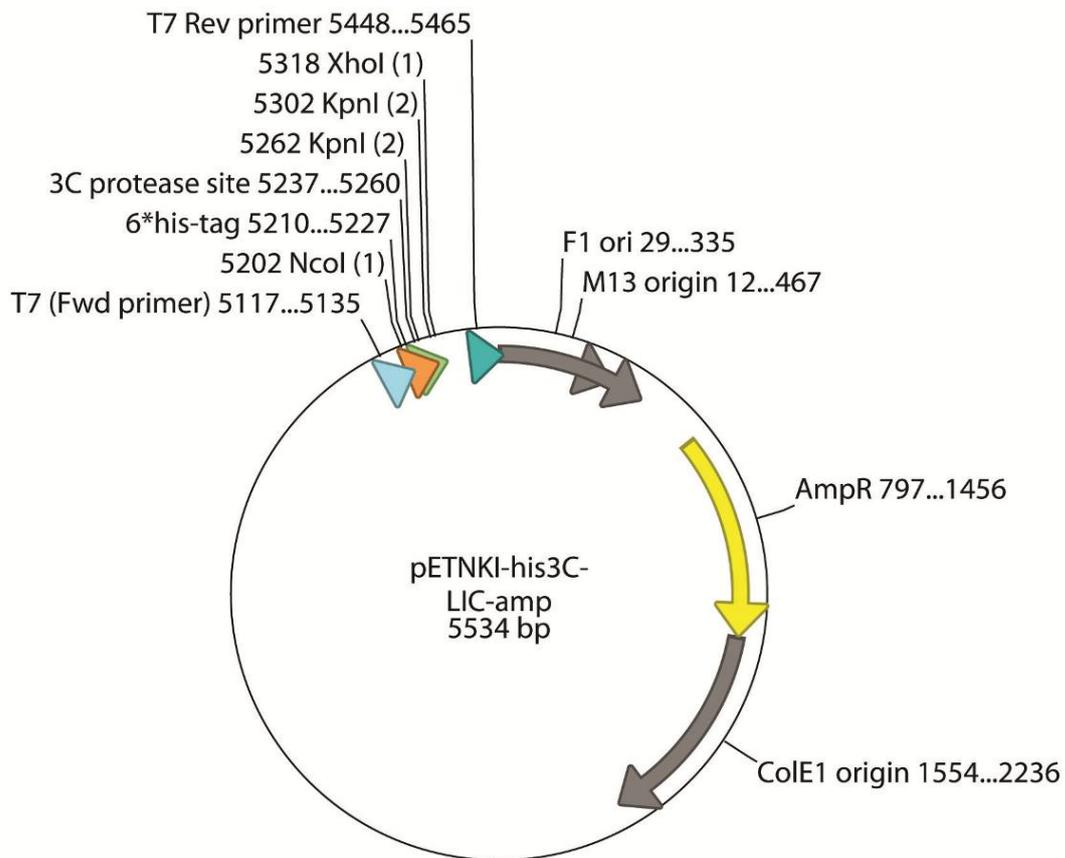


Figure 2.3: Map of the vector pETNKI-his3C-LIC-amp from NKI protein facility LIC (95).

2.2 Sequences of the proteins of interest

2.2.1 Flavohemoprotein (assemblage A)

The FIHb from assemblage A (GL50803_15009) was received from Steven Rafferty at The Chemistry Department Trent University, Peterborough, Canada. The gene of interest was in advance cloned into the vectors pET14b and pJ401. See Figure 2.4 for the protein sequence of FIHb from assemblage A.

```

MGSSHHHHHSSG|LVPRGSHMTLSEDTLRAVEATAGLIAAQGIEFTRAFYERMLT
1|          10|          20|          30|          40|          50|

KNEELKNIFNLAHQRTL RQPKALLDSLVAAYALNIRRINELYELKKGKGLPVPPEHW
60|          70|          80|          90|          100|          110|

AELQGFFSAAERVANKHTSFSGIQPAQYQIVGAHLLATIEDRITKDKDILAEWAKA
120|          130|          140|          150|          160|

YQFLADLFIKREEEIYAATEGCKGGWRQTRTRFRVEEKTRVNEIICKFRLVPAEEG
170|          180|          190|          200|          210|          220|

AGVVEHRPGQYLAI FVRSPEHFQHQQIRQYSIISAPNSAYYEIAVHRDEKGTVSR
230|          240|          250|          260|          270|

YLHDYVSTGDLLEVAPPYGDFFLRYLEADEQAPADTQASQEFQMLQSGAINFAAE
280|          290|          300|          310|          320|          330|

KTMPIVLISGGIGQTPLLSMLRFLAQKEGKETARPIFWIHAHNSRVRAFKEEVD
340|          350|          360|          370|          380|

AIRETALPSLRVVTFLSEVRATDREGEDYDFAGRINLDRISELTKLEADNANPHY
390|          400|          410|          420|          430|          440|

FFVGPTGFMTAVEEQLKTKSVPNSRIHFEMFGPFKASH
450|          460|          470|          478

```

Figure 2.4: Protein sequence of pET14b-gFIHb. The part marked in blue represents the His₆-tag (AA 1-13) and the pink part represents the cleavage site (AA 14-19). Picture was created with SnapGene® Viewer.

The sequence of pJ401-gFIHb is identical from AA 20. The first 13 AA represent the His₆-tag. The following 6 AA (AA 14-19) represent the TEV cleavage site. The cleavage site is different for the two constructs. For pJ401-gFIHb, the cleavage site is formed by the AA sequence: ENLYFQG, which is recognized by thrombin.

2.2.2 Flavohemoprotein (assemblage B)

For the FIHb from assemblage B (GSB_151570), three constructs of the protein were ordered from Genescript, Piscataway, NJ, USA. The first one is the full length of the protein (595 AA). The second one is truncated, starting from amino acid number 25 (570 AA in total). This one is truncated because the part left out of the structure was predicted to be unstructured

from the prediction (by HNN) in the secondary structure while using NKI primer design program (96). The third one is truncated as well, starting from amino acid number 138 (457 AA in total). This truncation is the corresponding construct for the one from FIHb from assemblage A. See figure 2.5 for the sequences of the three constructs.



Figure 2.5: Protein sequence of FIHb from assemblage B. The start of each construct of the protein is marked in green and the end is marked in red. Three different constructs in total. Picture was created with SnapGene® Viewer.

2.2.3 Nitroreductase 1 (assemblage B)

For the NR1 from assemblage B (GSB_153178), only the full length of the protein (287 AA) was used in this project. The construct was ordered from Genescript. See Figure 2.6 for the protein sequence of the protein.



Figure 2.6: Protein sequence of NR1. The start of the protein is marked in green and the end is marked in red. Picture was created with SnapGene® Viewer.

2.2.4 Nitroreductase 2 (assemblage B)

For the NR2 from assemblage B (GSB_22677), two constructs of the protein were ordered from Genescript used. The first one is the full length of the protein (264 AA). The second one is truncated, and starts from amino acid number 5 (259 AA in total). This one is truncated at this point because the vector has a 3C cleavage site, which will after the cleavage leave three AA (glycine, proline, glycine). Glycine is original next to tyrosine in the full length (see Figure 2.7), leaving only two extra AA in the truncated construct. See Figure 2.7 for the sequences of the constructs.

MVEGYPVFLTDTCVGCNMCVEVCPTMTLKVDPQTKVVAYANRDNCIFCGHCAAIC
 1| 10| 20| 30| 40| 50|

PSSTISMFGVTPESSEKAMKADPPAIRAIKMARTCRKYHPEPLPKDDIMKVISIA
 60| 70| 80| 90| 100| 110|

KYSASSSNVRPLHFTVLSRGTMDTLGLAIAKEVSRNPKYAKVAALMEKGIDVVFR
 120| 130| 140| 150| 160|

GAPHMLMSAPADKAAVAMADASIAGRDIQLNAESMGLGMFWCGFLLA AVASSQE
 170| 180| 190| 200| 210| 220|

LHDGCGVPEGHKILMAMGLGRP KIKFARPALRRDLEEGIDITF**K**
 230| 240| 250| 260| 264

Figure 2.7: Protein sequence of NR2. The start of each construct of the protein is marked in green and the end is marked in red. Two different constructs in total. Picture was created with SnapGene® Viewer.

2.3 Equipment

2.3.1 Tubes

Table 2.1: List of all tubes used in this project.

Tubes	Producer	Catalog number	Lot number
PCR Strips of 8 tubes 0,2 mL	BRAND GMBH & CO KG, Germany	781320	579335
Fisherbrand® PCR Strip of 8 with Cap Strip, 0,25 mL	Fisher Scientific, EU	14230215	NA
SafeSeal tube (eppendorf tube), 1,5 mL	SARSTEDT AG & Co. KG, Germany	72.706	0647/8043014
Tube 13 mL, sterile	SARSTEDT, Germany	62.515.006	6091811
Tube 15 mL (Falcon tube)	SARSTEDT, Germany	62.554.502	7042411
Tube 50 mL (Falcon tube)	SARSTEDT, Germany	62.559.001	7043212
Amicon® Ultra-15 Centrifugal Filters, Ultracel® - 30 K/10 K	Merck Millipore Ltd., Ireland	UFC903096 UFC901024	R7JA37191 R7EA90728

Amicon® Ultra-0,5 mL, Centrifugal Filters, Ultracel® - 30 K/10 K	Merck Millipore Ltd., Ireland	UFC503096 UFC501096	R5EA96214 R5DA68606
Costar® Spin-X Centrifuge Tube Filter 0,22 µm, 2 mL, non-sterile	Corning Incorporated, USA	CLS8160-96EA	11117000

2.3.2 Pipettes

Table 2.2: List of all pipettes used in this project.

Pipettes	Producer	Catalog number	Lot number
Pipette tip 10 µL	SARSTEDT, Germany	70.1130.600	NA
Pipette tip 200 µL	SARSTEDT, Germany	70.760.502	NA
Biosphere® Fil. Tip 10 neutral	SARSTEDT, Germany	70.1130.210	7052211
Serological pipette 1 mL	SARSTEDT, Germany	86.1251.001	6034E
Serological pipette 5 mL	SARSTEDT, Germany	86.1253.001	8266E
Serological pipette 10 mL	SARSTEDT, Germany	86.1254.001	8030E
Serological pipette 25 mL	SARSTEDT, Germany	86.1685.001	7257K

2.3.3 Kits

Table 2.3: List of the kits used in this project.

Kit	Producer	Catalog number
PureYield™ Plasmid Miniprep System	Promega, Norway	A1223
Wizard® SV Gel and PCR Clean-up System	Promega, Norway	A9281

Table 2.4: List of crystallization screens used in this study.

Crystallization screens	Producer	Catalog number
PACT premier™	Molecular Dimention, UK	MD1-29
JCSG plus™	Molecular Dimention, UK	MD1-40
SG1 screen™	Molecular Dimention, UK	MD1-89

2.3.4 Columns

Table 2.5: Columns used in this project.

Columns	Producer	Catalog number	Lot number
Pierce® Centrifuge Columns for affinity purification, 10 mL	Thermo Scientific, USA	89898	SF249329
PD-10 Columns for buffer exchange	GE Healthcare, UK	17-0851-01	12881046
HiLoad 16/600 Superdex 200 pg, SEC	GE Healthcare, UK	GE28-9893-35	
Superdex™ increase 200 10/300, SEC	GE Healthcare, UK	17517501	

2.3.5 Other consumables

Table 2.6: List of other consumables used in this project.

Consumable	Producer	Catalog number	Lot number
His-Pur™ Ni-NTA resin	Thermo Scientific, USA	88222	SI252668A
Instant Blue™	Expedeon, England	ISB1L	170616330
Mini-PROTEAN® TGX™ Gels	Bio-Rad Laboratories, Norway	456-1096	L006952 A
SeaKem® LE Agarose	Lonza, USA	50004	0000470886
LB-Agar (Lennox)	Carl Roth GmbH + Co. KG	X965.2	221170173
LB-Medium (Lennox)	Carl Roth GmbH + Co. KG	X964.2	034207346
T4 DNA polymerase	New England BioLabs, England	M0203L	0401706
Taq DNA polymerase	New England BioLabs, England	M0267S	0141706
FastDigest KpnI	Thermo Scientific, USA	FD0524	00367902

Isopropyl β -D-1-thiogalactopyranoside	Fischer BioReagents, USA	BP1755-10	165180
--	--------------------------	-----------	--------

2.3.6 Buffers

Table 2.7: List of buffers from companies used in this study.

Equipment	Producer	Catalog number	Lot number
FastDigest Green buffer (10X)	Thermo Fisher Scientific, USA	B72	00337394
NEBuffer™ 2.1 (10X)	New England BioLabs, England	B7202S	0231704
5X Phusion HF buffer	Thermo Scientific, USA	F-518	00583637
ThermoPol® 10X	New England BioLabs, England	M0267S	0011508

Additional buffers made in our laboratory are described in appendix 1.

2.3.7 Protocols

- Molecular methods method “The Inoue Method for Preparation and Transformation of Competent *E.coli*: “Ultracompetent” cells (97).
- NKI Protein Facility LIC Vectors (95).
- ACEMBL Expression System, User Manual, EMBL, p. 9-13 (98).
- Thermo Scientific, User guide: FastDigest KnpI (99).
- One-Step Sequence- and Ligation-Independent Cloning as a Rapid and Versatile Cloning Method for Functional Genomics Studies (100).
- Terrific Broth media (101).

2.3.8 Instruments

Table 2.8: List of all instruments used in this project.

Instrument	Use	Producer
VCX130, Ultrasonic Processor for Small Volume Applications	Sonication for cell lysis	Sonics & Materials, Inc, USA
WPA CO 8000 Biowave Cell Density Meter	Measure OD ₆₀₀	Biochrom Ltd, England
SORVALL RC-5C Plus Centrifuge	Centrifugation of higher volumes and high speed	DuPont, USA
Eppendorf Centrifuge 5810 R	Centrifugation of higher volume	Eppendorf
Eppendorf Centrifuge 5415 D	Centrifugation of smaller volume	Eppendorf
CERTOMAT® IS	Incubation of small scale bacteria culture	Sartorius, Germany
Allegra® X-15R	Centrifugation of high volume	Beckman Coulter, USA
Innova 4430 Incubator Shaker	Incubation of big scale bacteria culture	New Brunswick Scientific Co., Inc., USA
Eppendorf Thermomixer comfort, 1,5 mL	Heating, cooling and shaking of eppendorftubes (1,5 mL)	Eppendorf
MiniSpin®	Centrifuge small volumes (1,5 mL)	Eppendorf
Termaks TS 8056	Drying oven	Termaks AS, Norway
Heto OBN 8	Warm bath	Heto Lab Equipment
PowerPac® Basic Power Supply	Run SDS-PAGE	Bio-Rad Laboratories, Norway
PowerPac 200 Electrophoresis Supply	Run electrophoresis with agarose gel	Bio-Rad Laboratories, Norway
MyCycler™ thermal cycler	PCR	Bio-Rad Laboratories, Norway
ChemiDoc™ XRS+ with Image Lab™ Software	Image SDS-gels and agarose gels	Bio-Rad Laboratories, Norway
Abbatemat 500 Refractometer	Measure concentration	Anton Paar, Austria
NanoDrop® ND-1000 UV-Vis Spectrophotometer	Measure concentration	Thermo Fisher Scientific, USA
miniDAWN TREOS	SLS (MALS)	Wyatt technology, USA
RefractoMax 520	Measure the RI for concentration determination	ERC, Germany

	during SLS	
Spectropolarimeter J-810	Measure CD	JASCO, Japan
Mosquito® LCP	Preparation of crystallization plates	TTP Labtech,
ROCK IMAGER® 182	Imaging of crystallization plates	Formulatrix, USA
LightCycler® 480 II	Thermofluor assay	Roche, Switzerland
ÄKTApurifier	SEC (preparative and SLS)	GE Healthcare
Systec HX-320	Autoclave	Systec GmbH, Germany

2.3.9 Software

- Microsoft® Word for Mac 2011 (Microsoft corporation) was used for writing this thesis. Microsoft®.
- Excel for Mac 2011 (Microsoft corporation) was used to make diagrams and to transfer data from instruments.
- ATSAS online GASBOR was used to create the overall structure of the proteins with data from SAXS (102).
- Primus was used for processing and analysing of data from SAXS (103).
- SnapGene® Viewer 4.1.5 (GSL Biotech LLC) was used to compare sequences from sequencing and ordered sequences, and to create protein sequence figures.
- ASTRA 6 (Wyatt technology, USA) used for SLS.
- ND-1000 V3.5.2 (Thermo Fisher Scientific, USA) for measurements at NanoDrop®. ND-1000 UV-Vis Spectrophotometer.
- Spectra Manager for measurement of CD with Spectropolarimeter J-810.
- UNICORN 5.01 (GE Healthcare) used to create elution profile and collect fractions from SEC.
- ExPASy translate tool (104) for translate DNA sequences into protein sequences.
- ExPASy ProtParam tool (105) to compute various physical and chemical parameters for the proteins of interest.
- NKI primer design (96).
- PyMOL Molecular Graphics System, version 2.0.6 (Schrodinger LLC, USA).

3. METHODS

3.1 Preparing the competent cells (TOP10)

In order to make competent cells, the protocol described in Molecular methods “The Inoue Method for Preparation and Transformation of Competent *E.coli*: “Ultracompetent” cells (97) was used.

E.coli cells were incubated in super optimal broth (SOB) medium at 20 °C with shaking at 250 revolutions per minute (rpm) overnight. 100 mL of Inoue transformation buffer (ITB) was made. The cultures were incubated until OD₆₀₀ reached 0.55 (measured with WPA CO 8000 Biowave Cell Density Meter) and then they were transferred to an ice-bath for 10 min. The cells were harvested by centrifugation at 2500 \times g for 10 min at 4 °C (Allegra X-15R Centrifuge). The cells were resuspended gently in 80 mL of ice-cold ITB and harvest again with the same centrifugation procedure. Further the cells were resuspended in 10 mL ice-cold ITB and 750 μ L of dimethyl sulfoxide (DMSO) was added. Lastly the cells were dispensed into chilled eppendorf tubes in a cold room (4 °C) followed by a snap-freeze in a bath of liquid nitrogen.

3.1.1 Transformation into competent cells

Plasmid (0.5 μ L) was added into competent cells, TOP10 (100 μ L), and incubated on ice for 30 min. After incubation the cells were shock-heated at 42 °C (water bath) for 45 sec and immediately put back on ice for at least 2 min. Then 500 μ L Luria-Bertani broth (LB) medium was added to the cells and incubated at 37 °C for approximately 1 h with shaking at 700 rpm (using CERTOMAT® IS). 50 μ L of transformation mixture were plated on LB agar plates containing antibiotics. The plates were incubated overnight at 37 °C (Thermaks TS 8056). One culture from each of the agar plates was transferred to a flask containing LB media and antibiotics. Flasks were incubated overnight at 37 °C with shaking at 250 rpm (CERTOMAT® IS).

LB medium was prepared by Ju Xu, at laboratory D at the Department of Biomedicine, University of Bergen. LB agar plates were prepared by adding 8.75 g LB-Agar (Lennox) into

250 mL water and autoclaved (Systec HX-320). Mixture was microwaved until the hard part was dissolved, cooled down to room temperature (RT) and divided into plates.

3.1.2 Plasmid isolation

Miniprep is a rapid method for isolating smaller amount of plasmids, which have been amplified in *E.coli* bacterial cells (106). Plasmid isolation was done using Miniprep kit (PureYield™ Plasmid Miniprep System), and the manufacture protocol. After transformation and amplification of the desired plasmid into TOP10 cells, the plasmid was purified using Miniprep.

3.2 Cloning

Sequence- and ligation-independent cloning (SLIC) method was used. The gene of interest (GOI) is PCR amplified with specific primers and the vector is linearized. The PCR amplified product and the linearized vector are separately treated with T4 DNA polymerase. T4 DNA polymerase acts as an exonuclease in the absence of dNTP and produces long sticky overhangs. The T4 DNA polymerase treated GOI and vector are mixed and annealed, followed by transformation, yielding a single gene expression cassette (98). The cloning was done using ACEMBL Procedure (98).

3.2.1 Primer design

Primers for the insert contain a DNA sequence complementary for the GOI, and overhangs, the corresponding to a region of sequence in the expression vector. In this case where NKI vectors are used, it is the site where the vector is linearized using Kpn1 enzyme. These overhangs result in the long sticky end upon treatment with T4 DNA polymerase in the absence of dNTP. Primers were design from each of the regions marked with a green start and red stop in the protein sequences. See Figures 2.5-2.7. Forward and reverse primers are designed to have identical T_m s (melting temperatures) for the PCR to be successful.

NKI primer design (96) was used for designing primers for NR1, NR2 and FIHb assemblage B.

3.2.2 Polymerase chain reaction

PCR is a widely used method for amplifying specific DNA sequences. The reaction consists of three steps; denaturation, annealing and extension. In the denaturation step, the two strands of the parent DNA are separated by heat (98 °C). Further in the annealing step, the temperature is cooled down to the T_m of the primers, so the primers can bind to their complementary sequence on the single strand template DNA. In the last step, the temperature is increased (72 °C) in order for Taq polymerase to extend the primers and synthesizes new strands of DNA. The steps are repeated in 25-35 cycles (107). The result of a successful PCR experiment, depends on the primer design, the time in each step and the temperature (108).

The protocol from ACEMBL (98) was used for PCR amplification of insert and vector, with some adjustments (Table 3.1).

Table 3.1: Components added in PCR tubes before PCR run, using 2X Phusion HF Master Mix.

Reagents		Amount added
ddH ₂ O		Add up to 25 µL
2X Phusion HF PCR Master Mix with HF buffer		12,5 µL
Forward primers	FIHb_FW_1	1 µL
	FIHb_FW_26	1 µL
	FIHb_FW_138	1 µL
	NR1_FW_1	1 µL
	NR2_FW_1	1 µL
	NR2_FW_5	1 µL
Reverse primers	FIHb_RV_595	1 µL
	NR1_RV_287	1 µL
	NR2_RV_264	1 µL
Template DNA	FIHb	1 µL ≈ 103,7 ng
	NR1	1 µL ≈ 55,6 ng
	NR2	1,5 µL ≈ 47,6 ng
Total amount added		25 µL

All components used for PCR reaction of the insert (gene of interest) is given in Table 3.1. All PCR tubes contained water (ddH₂O) and master mix, with different template DNA and one corresponding forward primer and the corresponding reverse primer. The template DNA

for FIHb was divided into three PCR reactions (1:FIHb_FW_1 and FIHb_RV_595, 2: FIHb_FW_26 and FIHb_RV_595, 3:FIHb_FW_138 and FIHb_RV_595). Similarly two PCR reactions were used for NR2 (1:NR2_FW_1 and NR2_RV_264 and 2:NR2_FW_5 and NR2_RV_264). For NR1, one PCR reaction was used (NR1_FW_1 and NR1_RV_287).

All of the prepared PCR tubes were run in the PCR instrument (MyCycler™) at this program:

1 x 98 °C
35 x [98 °C for 10 sec. → 65 °C for 30 sec. → 72 °C for 2 min]
1 x 72 °C for 10 min

The amplified PCR product was then loaded on 1 % agarose gel:

2 µL of PCR sample and 0.5 µL of dye (6x loading buffer + green dye). 1 X tris-acetate-EDTA (TAE) buffer was filled in the chamber until it covered the gel. Run at 100 V for 60 min with PowerPac 200 Electrophoresis Supply.

The rest of the PCR products were loaded on gel one more time with the same run. Each band from each of the genes was cut out of the gel. The gel bands were cleaned with PCR cleaning kit (Promega, Wizard® SV Gel and PCR Clean-up System).

1 % agarose gel was prepared by adding 0.5 g SeaKem® LE Agarose and 50 mL of 1 X TAE buffer into a flask. Mixture was microwaved until the agarose powder was dissolved, and poured into suitable forms with a comb to make the wells. The gel was solid after approximately 15 min.

3.2.3 Linearization of the vector

The vector (pETNKI-his3C-LIC-amp) was linearized with the enzyme Kpn1. The protocol from Thermo Scientific FastDigest KpnI was followed with some modification:

pETNKI-his3C-LIC-amp (20 ng/ µL)	40 µL
Kpn1 enzyme	3 µL
10 X Fast Digest Green Buffer	5 µL
ddH ₂ O	2 µL
Total	50 µ L

The components in 3.2.3 were mixed together. The mixture was incubated at 37 °C for 2 h. KpnI was inactivated at 80 °C for 5 min before the mixture was loaded on 1% agarose gel with dye (50 µL vector mixture + 8 µL (6x loading buffer + green dye)). The bands on the gel were cut out and purified with PCR cleaning kit (Promega, Wizard® SV Gel and PCR Clean-up System).

3.2.4 T4 DNA polymerase exonuclease treatment

All the inserts and the linearized vector (pETNKI-his3C-LIC-amp) were treated with T4 DNA polymerase. Two different methods were tested. The first method follows the ACEMBL protocol (98) where the inserts and the vectors are treated separately before they are mixed together. Separately they are incubated at 23 °C for 20 min, 1 µL of EDTA is added, and then they are inactivated at 75 °C for 20 min. After mixing, they are annealed at 65 °C for 10 min and then slowly cooled down to RT in a heat block. Following components were mixed in each PCR tube:

10 X NEB 2.1 buffer	1 µL
100 mM DTT	0.5 µL
2 M Urea	1 µL
Insert/ vector	7 µL
<u>T4 DNA polymerase</u>	<u>0.5 µL</u>
Total	10 µ L

The second method is a one-step treatment (100) where the insert and the vector are first mixed, and then treated with T4 in the same tube. The mixture was incubated at 50 °C for 30 sec, and then inactivated on ice. The following components were mixed in each PCR tube for this method:

Linearized vector	1 μ L
Insert	3 μ L
10 X NEB 2.1 Buffer	1 μ L
100 mM DTT	1 μ L
T4 DNA polymerase	0.2 μ L
Water	4 μ L
Total	10 μ L

All the T4 DNA treated samples were transformed into competent cells (TOP10) with standard transformation procedures. Colonies from the transformation were run with PCR to check if the cloning had worked.

Colony PCR were run with Taq DNA polymerase with following components added to each PCR tube:

10 X ThermoPol buffer	2.5 μ L
10 mM dNTPs	0.5 μ L
10 μ M FW primer	0.5 μ L
10 μ M RV primer	0.5 μ L
1 bacteria colony from agarplate	
Taq DNA polymerase	0.125 μ L
Water	20.9 μ L
Total	25 μ L

PCR program:

95 °C for 30 sec

[95 °C for 15 sec, 65 °C for 30 sec, 68 °C for 2 min] x 30 cycles

68 °C for 5 min

The finished PCR products (2 μ L of PCR sample + 0.5 μ L of dye (6x loading buffer + green dye)) were loaded on 1 % agarose gel to check if the PCR was successful. Run at 100 V for 40 min.

3.2.5 Sequencing

Positive colonies from colony PCR were sent for sequencing (at the Sequence laboratory on the 6th floor of the Laboratory building, Center for Medical Genetics and Molecular Medicine, Haukeland University Hospital). First the colonies from the fresh agar plates were transferred to LB media and bacteria was cultured at 37 °C with shaking (250 rpm) overnight. The plasmid DNA was isolated from the bacteria media before following components were mixed in each PCR tube:

Big dye 3.1	1 µL
Sequencing buffer	1 µL
T7 primer (FW or RV)	1 µL
Plasmid DNA	3 µL (200 ng)
Water	4 µL (up to 10 µL)
Total	10 µ L

PCR program:

96 °C for 5 min

[96 °C for 10 sec, 50 °C for 5 sec, 60 °C for 4 min] x 25 cycles

10 °C ∞

10 µL water was added to the each PCR tube after the PCR reaction before they were sent for sequencing.

3.3 Protein expression

3.3.1 Expression of flavohemoglobin from assemblage A

pJ401-gFIHb was transformed into Rosetta (DE3) cells with the transformation procedure described earlier (3.1.1) using antibiotics, kanamycin and chloramphenicol for the overnight culture. Bacteria culture (5 mL) was then added to two different Fernback flasks of terrific broth (TB) media (1 L). ALA 0.1 mM and Potassium hexacyanoferrate (III) ($K_3Fe(CN)_6$) 0,01 mM were added to one Fernback flask of TB media (1 L) and only $K_3Fe(CN)_6$ was added to the other Fernback flask of TB media (1 L). Antibiotics were added as well, both kanamycin and chloramphenicol. The flasks were incubated at 37 °C with shaking (200 rpm) (Innova

4430 Incubator Shaker) until the cells reached an OD₆₀₀ of 0.4-0.5 (WPA CO 8000 Biowave Cell Density Meter). 1 mM IPTG was added to both flasks and incubated further overnight at 20 °C (Innova 4430 incubator shaker). An aliquot (1 mL) was taken after 0,17 and 24 hours, centrifuged (maximum speed with MiniSpin® for 1 min) and the pellets collected. After 24 hours of incubation, the cells were collected by centrifugation at 4000 g for 20 min at 4 °C. The medium was discarded and the cells were scrapped with a spatula and put on Falcon tubes.

pET14b-gFIHb was transformed into Rosetta cells in the same way as pJ401-gFIHb, but with ampicillin and chloramphenicol as antibiotics. The bacteria culture (5 mL) was added to a Fernback flask of TB media (1 L) together with ampicillin (1 mL) and chloramphenicol (1 mL). The flask was incubated with the same procedure as the pJ401-gFIHb culture, but without the IPTG induction. Aliquots and the cells were collected in the same way as for pJ401-gFIHb.

3.3.2 Expression of flavohemoprotein138

Expression using the pETNKI vector was done after the cloning. For the truncated FIHb138 (starting from AA 138 from Figure 2.5) the same procedure as described in 3.2.1 was done, except the K₃Fe(CN)₆ was replaced with the 1000 X trace elements (109) (0.2 mL in 1L TB media), to provide a supply for the iron in ferrous state.

3.3.3 Expression of nitroreductase 1

For the NR1, the same procedure as for FIHb138 was followed (described in 3.2.2), except for the addition of ALA. 1 mM IPTG was added when the OD₆₀₀ reached 0.7.

For NR1, auto-induction was also tested. The ZYM-5052 medium (109) for auto-induction was used. After transformation into Rosetta cells, the bacteria culture was added into a Fernback flask with ZYM-5052 medium (1L). The flask was incubated at 37 °C with shaking (200 rpm) for 2 hours, then the temperature was lowered to 18 °C and the cells were incubated over the weekend (64 hours). The cells were collected in the same way as described in 3.2.1.

3.4 SDS-PAGE

SDS-PAGE is a method for separating proteins based on their molecular weight. As proteins move through the gel in response to an electric field, the gel's pore structure allows smaller proteins to travel more rapidly than larger proteins. The SDS-PAGE system is a discontinuous denaturing system. The proteins become fully denatured and dissociate from each other by heat in the presence of SDS and the reducing agent, β -mercaptoethanol (β -ME), which breaks possible disulfide bonds (110).

All the pellets from each of the sample were resuspended in 0.5 mL buffer mixture of 20 mM Hepes buffer + 150 mM NaCl with a pipette. The suspension was sonicated for 10 seconds with amplitude of 70. Further 20 μ L from each of the samples was transferred to eppendorf tubes. Then the remains were spun down at 13.4 rpm (maximum speed) for 5 min. 20 μ L of the supernatant was transferred to new eppendorf tubes. 5 μ L of SDS (+ 1 % β -ME in SDS buffer) loading buffer (5 X) was added to all of the eppendorf tubes. The tubes were heated at 95 °C for approximately 5 min (Eppendorf Thermomixer). 8 μ L of each sample loaded on the gel (Mini-PROTEAN[®] TGX[™] Precast Gels) together with 2 μ L of marker. The chamber was filled with 1 X running buffer. The gel was set to run at 200 V for 35 min with PowerPac[®]. The gel was stained with Instant blue.

3.5 Protein purification

3.5.1 Ni-NTA affinity chromatography

Affinity chromatography is based on specific binding of the compound of interest to the stationary phase. When the solution is passed through the column, the compound of interest is bound. A washing step is then done using wash buffer and the one adhering compound is eluted by changing the condition (111). This purification method is highly selective for the protein of interest. In this case, the protein of interest has a string of histidine residues (called His-tag) on protein N-terminus. The tagged proteins are then passed through a column of beads containing covalently attached, immobilized nickel (II). The His-tag binds tightly to nickel and will therefore also bind the protein of interest while other proteins flow through the column. The protein can be eluted from the column by adding imidazole, which binds to the metal ions and displaces the protein (59).

The following buffers (see appendix 1) were used for the purification of the different proteins:

- pJ401-gFIHb, pET14b-gFIHb and FIHb138: Lysis buffer 1, wash buffer 1 and elution buffer 1
- NR1: Lysis buffer 2, wash buffer 2 and elution buffer 2

The pellet collected from expression was resuspended by pipetting in 30 mL of lysis buffer with lysozyme (0.1 mg/mL). The addition of lysozyme will help to break the cell walls. The cells were then lysed by sonication 4 x 2 min, 10 sec pulses and 23 W. The lysate was centrifuged using SS-34 rotor with the speed of 17 000 rpm for 30 min (SORVALL RC-5C Plus centrifuge). After the centrifugation, the supernatant containing the soluble proteins was collected.

To prepare the column ready, 2 mL of Ni-NTA was added to Pierce® Centrifuge Columns, and washed with water to remove ethanol from the Ni-NTA subsequently followed by a lysis buffer wash. Then the supernatant was added to the column and mixed with the Ni-NTA by pipetting. The column was opened and the contents were eluted. Then the column was washed with 15 mL of washing buffer 3-4 times. Aliquots (50 µL) from each wash was collected to measure the amount of protein left in each wash, and to validate the amount of washing steps needed to remove contaminants. The concentration was measured by nanodrop (Spectrophotometer) at 280 nm. Iron (1 mM Fe(CN₆)) was mixed with washing buffer (1 mL in total) and added to the column before the last wash, to get the heme in the protein to ferric state. When the washing was done, 6-8 mL elution buffer was added. The protein was then eluted from the column and collected. Buffer exchange was done with a PD10 column to remove excess of imidazole.

If it was not possible to continue to the next purification step the next day, the protein solution of both FIHb assemblage A and FIHb138, was frozen down by snap freezing with nitrogen and stored at -80 °C.

3.5.2 Size exclusion chromatography

Size exclusion chromatography (or gel filtration chromatography) is a method for separating molecules by their size. This is often used as a last purification step after at least one other purification step. The larger molecules will be eluted first and the smaller molecules will be

eluted later (86). Small molecules penetrate the pores in the stationary phase, and larger molecules do not. Because small molecules must pass through an effectively larger volume, the large molecules will be eluted first (111). The beads of gel filtration columns consist of cross-linked polyacrylamide, agarose, dextran, or a combination. Depending on the molecule size, it may or may not enter the beads. By using this method we can separate and collect different size species, in addition to estimate how much of the protein is a monomeric form and how much has aggregated or formed oligomers (112). The detection happens by monitoring absorbance at 280 nm.

Depending on the protein solution volume, two different columns and loops were used. For a sample volume of 1 mL, the loop size was 1 mL and the column used was HiLoad 16/600 Superdex 200 prep grade. For smaller volumes like 250 μ L, the size of the loop was 250 μ L and the column used was Superdex™ 200 increase 10/300.

For all FIHb constructs, gel filtration buffer 1 was used and for NR1 gel filtration buffer 3 was used. The buffer of use was in advance filtered and degassed to avoid introducing air in the column. Before the protein solution was injected into the instrument, it was filtered with Costar® Spin-X Centrifuge Tube Filter 0,22 μ m, 2 mL, non-sterile.

Gel filtration was done prior the crystallization screening of FIHb, to make sure that the sample is pure. In this case, FAD (1 mM) and $K_3Fe(CN)_6$ (1 mM) was added to the protein solution before filtration. The buffer of use was gel filtration buffer 2. HiLoad 16/600 Superdex 200 prep grade column was used for 2 mL of protein solution. Fractions from the elution profile were collected in a 96-well plate.

3.6 Concentrating the proteins

The purified protein solutions (up to 10 mL) were concentrated to the wanted concentration by adding the protein solution to Amicon® Ultra-15 Centrifugal Filters, Ultracel® (30 K for FIHb, 10 K for NR1), and centrifuged at 3900 rmp for approximately 10 min (depending on the concentration) at 4 °C using Eppendorf Centrifuge 5810 R.

For smaller volumes of protein solutions (up to 0.5 mL) Amicon® Ultra-0.5 mL Centrifugal Filters, Ultracel® was used, and centrifuged at maximum speed for approximately 2 min at 4 °C using Eppendorf Centrifuge 5515 D.

Measuring the concentration of the purified protein was done both with spectrophotometer (nanodrop A280) and with refractometer (Anton Paar Abbemat 500 Refractometer).

3.7 His-tag cleavage

3.7.1 His-tag cleavage using TEV

For pJ401-gFIHb, the His₆-tag can be cleaved by TEV protease. TEV protease has a chymotrypsin-like fold with an atypical catalytic triad in which cysteine replaces serine, and exhibit an absolute requirement for glutamine in the P1 position of their substrates (113).

1) After the size exclusion chromatography, FIHb fractions from the elution profile, was collected and concentrated to half the volume (400 µL). 1 mM DTT and 20 µL TEV was added to the protein solution. DTT is a reducing agent used to reduce protein disulfide bonds. Aliquots (20 µL) were taken right after TEV was added, after 2 hours and after 18 hours, and analyzed on SDS-PAGE.

2) 150 µL TEV was added to 3 mL of protein solution (2,59 mg/mL) in a dialysis tube. The tube were placed in dialysis buffer 1 (see appendix 1) and left overnight for dialysis at 4 °C.

3) 125 µL flavohemoprotein was added in 500 µL dialysis buffer 1 (appendix 1) and 20 µL TEV protease was added to the protein solution. Two samples of this mixture were prepared, one sample was incubated in room temperature and one sample was incubated at 30 °C overnight. Aliquots (20 µL) were taken after 0, 2, 4 and 24 hours and analyzed on SDS-PAGE.

Analyzes using SDS-PAGE were done to check if the cleavage had worked. If there was a shift in the band sizes comparing with and without TEV, the cleavage had worked. If the SDS-gel did not give a clear result, another method was used. This method was to run the protein solution with TEV through a Ni-NTA column. If the His₆-tag had been cleaved off,

the protein would no longer bind to the resin, but follow out through the column and be in the flow through.

3.7.2 His-tag cleavage using thrombin

For pET14b-gFIHb the His₆-tag can be cleaved off by thrombin. The first cleavage trial was done by adding 10 units of thrombin per mg protein to the protein solution. The mixture was left on ice overnight. A second trial was done in RT overnight. This time one sample contained 50 units of thrombin per mg protein and one sample contained 100 units per mg protein. To confirm if the cleavage had worked or not, the same methods as described in the section above (3.7.1), were used.

3.7.3 His-tag cleavage using 3C protease

For the constructs cloned into the vector pETKNI-his-3C-LIC-amp, the His₆-tag can be cleaved off with 3C protease. 3C protease was added to the affinity purified FIHb138 with a ratio of 1:10. The mixture was left at 4 °C overnight. To confirm if the cleavage had worked or not, the same methods as described in the section above (3.7.1), were used.

3.8 Characterization of protein

All characterization was done with FIHb from assemblage A, expressed with pET14b-gFIHb.

3.8.1 Peptide mass fingerprinting

Peptide mass fingerprinting (PMF) is a protein identification technique. The protein of interest is digested with trypsin yielding a defined number of peptides of specific length and mass. These peptides with specific masses are unique to a protein. The peptides are analyzed by using mass spectrometry (MS). The protein of interest can be identified by comparing the list of peptide masses (peptide sequences) from the protein of interest to a peptide list of a database protein (114).

PMF was used to confirm that the bands from SDS-PAGE gels were the right proteins. The bands were cut out from the gel, and put in eppendorf tubes. The samples were sent to Ulrich Bergmann at University of Oulu, who performed the PMF.

3.8.2 Multi angle light scattering

MALS was done in combination with chromatography (SEC system), SEC-MALS. By using the combination, the chromatography system will remove problems related to purity and prepare the sample for FIHb assemblage A. With SEC, the different size species were separated, followed by measurement of RI signal for the sample and light scattering (LS) data. Sample's concentration from UV or RI can then be combined with LS data. The concentration source used in this case was RI. A MALS detector collects scattered lights at many angles. It can measure molecular weight for both isotropic and anisotropic scatters.

The FIHb protein sample purified from His-tagged Ni-NTA affinity chromatography, was filtered with Costar® Spin-X Centrifuge Tube Filter 0.22 µm, 2 mL, non-sterile. The buffer used for SEC-MALS was phosphate-buffered saline (PBS).

In addition, SEC-MALS was performed by Erik Hallin at Diamond Light Source, Harwell Science & Innovation Campus, England, in conjunction with SAXS. The sample preparation is described in 3.8.5.

3.8.3 Circular dichroism

CD was done with FIHb from assemblage A. The protein solution was dialysed against dialysis buffer 2 (appendix 1) overnight in cold room, using a dialysis button. The start concentration of the protein solution was 0.71 mg/mL. The protein was run on the CD instrument, but the concentration needed to be lowered as the detector was saturated. For measuring the CD spectrum, a protein concentration of 0.175 mg/ml was used. With the given concentration, a run where the CD spectra are measured as a function of temperature was done to estimate the T_m of the protein.

Two peaks were seen close to each other in the elution profile from SEC, therefore it was desirable to compare the secondary structure of the proteins from these two peaks. This was done with CD. FIHb assemblage A was run with SEC with gel filtration buffer 2 (appendix 1). Fractions were collected from SEC respectively to the two peaks. CD spectra were compared.

3.8.4 ThermoFluor

A thermoFluor assay was performed to determine the T_m and to screen for the optimal buffer condition where FIHb assemblage A is most stable. First a thermoFluor buffer/salt screen was made in a 2-mL deep 96-well block. See Table 3.3 for the different condition in each well. Second the thermoFluor plate was prepared with 16 μ L of each condition from the screen, 1 μ L 5.28 mg/ml protein and 3 μ L of 100x SYPRO Orange in 50 % DMSO in each well of the plate (96 wells in total). The plate was placed in the instrument LightCycler® 480 II, where the thermoFluor assay was performed.

Table 3.3: Buffer and salt screening components for the ThermoFluor.

	1	2	3	4	5	6	7	8	9	10	11	12	SALT
A													0 mM NaCl
B													100 mM NaCl
C													250 mM NaCl
D													500 mM NaCl
E													0 mM NaCl
F													100 mM NaCl
G													250 mM NaCl
H													500 mM NaCl
BUFFER	20 mM Bis-tris, pH 6		20 mM Tris, pH 7,5		20 mM Tris, pH 8,5		20 mM Bis-tris, pH6		20 mM Tris, pH 7,5		20 mM Tris, pH 8,5		

No additives	10 mM MgCl ₂
50 mM (NH ₄) ₂ SO ₄	10% glycerol

3.8.5 Small angle x-ray scattering

The sample was carefully purified with Ni-NTA affinity purification and then stored on ice until it was measured. The concentration of this sample was 10.5 mg/mL. SEC coupled SAXS was performed with FIHb assemblage A.

The SAXS experiment was executed by Erik Hallin at Diamond Light Source, Harwell Science & Innovation Campus, England. Collected SEC-SAXS frames were processed using a script written by Hallin. Data was processed and analysed using Primus (103) from ATSAS package. Models were created using GASBOR (102).

3.9 Protein crystallization

Protein concentration of 12.67 mg/mL was used for initial crystallization screening for the FIHb assemblage A. 1 mM FAD was added to half of the protein solution. Sitting drop vapour-diffusion method was used. Plates were prepared using Mosquito HT crystallization robot. One droplet contained FIHb and one droplet contained FIHb + FAD mixed 1:1 with crystallization solution. MRC SD2 96 well plate was used for the droplets. Crystallization screens JCSG-plus HT 96 and PACT-premier was used for screening for optimal crystallization conditions. The plates were stored in 20 °C and each droplet were imaged by ROCK IMAGER® 182.

New plates were set up with lower concentrations of FIHb from assemblage A (9 mg/mL and 4.7 mg/mL) and FAD (1 mM) added to all of the droplets. Crystallization screen PACT-premier and JCSG-plus HT 96 were used. The plates were stored in 20 °C and each droplet were imaged by ROCK IMAGER® 182.

More screening was done to find the optimal crystallization condition. 4 new crystallization plates were prepared. Two plates containing PACT-premier and two plates containing SG1. Fe and FAD were added to the protein solution before gel filtration, and the concentration of the solution was 10.3 mg/ml. One plate with PACT and one plate with SG1 were put in 8 °C, and the other two plates were put in 20 °C.

Crystallization plates were also set up directly after affinity purification and buffer exchange. The buffer used was gel filtration buffer 1 (appendix 1). The protein concentration for this experiment was 11.1 mg/mL. Plates were set up with the crystallization screens JCSG-plus HT 96, PACT-premier and SG1. The plates were stored in 20 °C and each droplet were imaged by ROCK IMAGER® 182.

Two different fractions from SEC were collected and new trials for crystallization were set up. The buffer used for SEC was gel filtration buffer 1 (appendix 1). The concentration of the colorless fraction was 9.3 mg/mL and the concentration of the colourful fraction was 10 mg/mL. The same crystallization screens were used (JCSG, PACT and SG1). The plates were stored in 20 °C and each droplet were imaged by ROCK IMAGER® 182.

UV-positive crystals were tested for diffraction at the beam line P11 at DESY/PETRA III by Juha Kallio.

4. RESULTS

4.1 Cloning

4.1.1 Primer design

Primers in Table 4.1 were designed:

Table 4.1: Primers designed in this study

Gene	Base pair	Primer name	Sequence
FIHb	1785	FIHb_FW_1	CAGGGACCCGGTGTTCATGTTCCGCATAGCGT
	1710	FIHb_FW_25	CAGGGACCCGGTCGTAAAGCACCGGCATGG
	1371	FIHb_FW_138	CAGGGACCCGGTATGCCGCTGAGCGAAGATA
		FIHb_RV_595	CGAGGAGAAGCCCGGTTACTGCAGCGGTTTAAACGGA
NR1	861	NR1_FW_1	CAGGGACCCGGTATGCTGATCAACATTTTTGCG
		NR1_RV_287	CGAGGAGAAGCCCGGTTACAGAAAGGTAACATCTGC ACCTTC
NR2	792	NR2_FW_1	CAGGGACCCGGTATGCTGATCAACATTTTTGCG
	777	NR2_FW_5	CAGGGACCCGGTATTTTTGCGACCCTGAATATGAG
		NR2_RV_264	CGAGGAGAAGCCCGGTTAACACCTAAAACAAATGCCAGC

The part of the sequence marked in red, is the “overhang” of the primer, which is complementary to the vector backbone. The black part of the sequence is the complementary region for the GOI. T_m of the primers are estimated to 60-65 °C.

4.1.2 Cloning and transformation into competent cells

PCR was done to amplify each of the protein coding sequence from synthetic genes from assemblage B. The PCR products were run on a 1 % agarose gel, stained with GelGreen and imaged under UV-light. The results from the gel electrophoresis indicate the size of the PCR products. The picture (Figure 4.1) shows bands corresponding to the expected size of genes of

interest. FIHb is approximately 1400-1800 base pair (bp) in size (see Table 4.1) corresponding to 1.4-1.8 Kb in the Figure 4.1. NR1 and NR2 are corresponding to their gene length of 0.8-0.9 Kb (Figure 4.1).

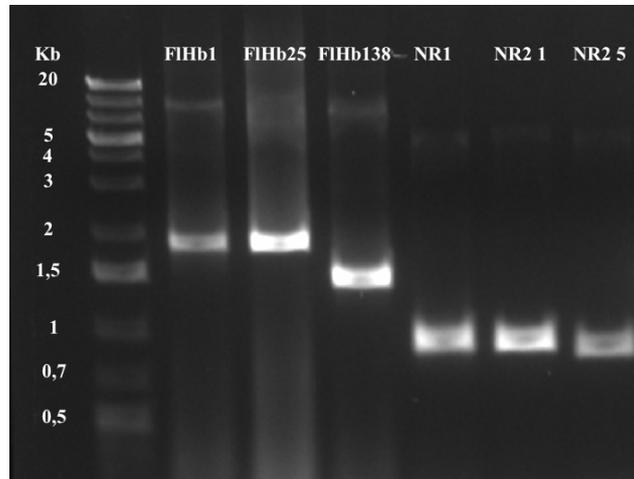


Figure 4.1: Electrophoresis in 1% agarose gel stained with GelGreen. Bands representing PCR amplification of *FIHb1*, *FIHb25*, *FIHb138*, *NR1*, *NR2 1* and *NR2 5*.

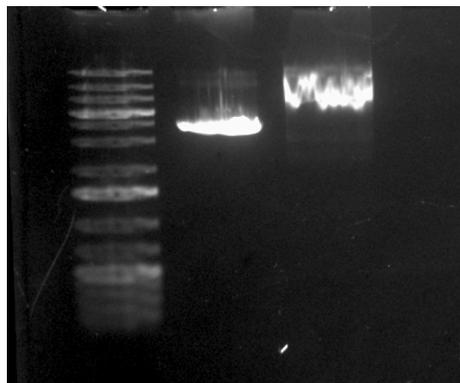
The bands from the gel were then cut out from the gel and purified with PCR cleaning kit, resulting the final concentrations of the genes after the cleaning which were rather low (see Table 4.2)

Table 4.2: Concentrations of the amplified genes after they are purified from 1 % agarose gel.

Gene	Concentration (ng/μL)
FIHb 1	9.8
FIHb 25	14
FIHb 138	6.5
NR1	4.8
NR2 1	7
NR2 5	10.7

The vector, pETNKI-his-3C-LIC-amp, was linearized, cut out of the gel and purified. Figure 4.2 shows a shift in the bands, which indicates that the linearization was successful. The linearized vector seems larger than the circular plasmid on the gel even though the size of the plasmid is the same. The circular plasmid can be observed in four different forms: circular

relaxed (nicked), circular supercoiled and circular single stranded (115). The conformational changes of the plasmid will affect their movement during gel electrophoresis (116). Supercoiled plasmid will move faster than relaxed plasmid because it is tightly coiled and more compact resulting in less frictional resistance from the gel. Linear plasmid on the other hand will cause more frictional resistance and therefore move slower on the gel, resulting in appearance at higher kB than supercoiled plasmid on the gel (116). The concentration of the vector after the cleaning was 17-30 ng/ μL .



*Figure 4.2: Electrophoresis in 1% agarose gel stained with GelGreen. The band shift indicates linearization of the vector pETNKI-his-3C-LIC-amp. The first line is the ladder, the second line is the vector without *kpn1*-treatment and the third line is the *kpn1*-treated vector.*

The amplified genes and vectors were separately treated with T4 DNA polymerase to create single stranded overhangs, mixed together followed by an annealing step. Mixtures were then transformed into competent cells, TOP10, and plated on agar plates containing ampicillin. Colonies from the plates were analyzed with colony PCR to determine the presence or absence of the insert. Colonies with insert are shown on the pictures below (Figure 4.3 and 4.4).

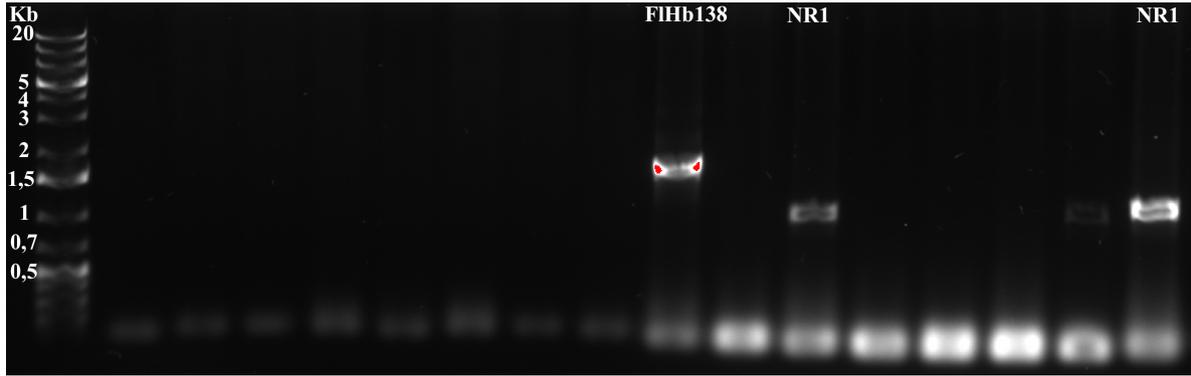


Figure 4.3: Electrophoresis in 1% agarose gel stained with GelGreen. Result from colony PCR. FHMP138 and NR1 are presented on the gel, while the rest of the bands are empty vectors.

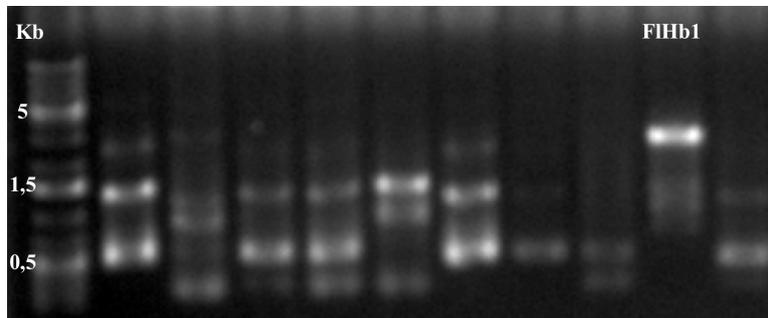


Figure 4.4: Electrophoresis in 1% agarose gel stained with GelGreen. Result from colony PCR with FHMP1 presented on the gel, while the rest of the bands are empty vectors or non-clear cloning result.

4.1.3 Sequencing

Positive colonies from the cloning process were sent for sequencing. The sequence results received back were positive for FIHb1, FIHb138 and NR1. The sequence results received back were compared with the sequences in Figure 2.5 and 2.6 using SnapGene® Viewer and ExPASy translate tool (104).

The cloning and sequence result for the FIHb25 construct and NR2 were negative.

4.2 Protein expression

4.2.1 Expression of flavohemoglobin assemblage A with pET14b-gFIHb

During the expression of FIHb (GL50803_15009) with the vector pET14b-gFIHb, aliquots from different time intervals were collected for analysis using SDS-PAGE. Samples from both soluble (supernatant) and mixed soluble and insoluble (lysate) fractions at each time

interval are presented on the SDS-PAGE (Figure 4.5). The SDS-PAGE shows clear, thick bands at approximately 54 kilodalton (kDa). The bands in Figure 4.1 indicate that 17 hours of incubation was enough for the *E.coli* Rosetta cells to express the protein well.

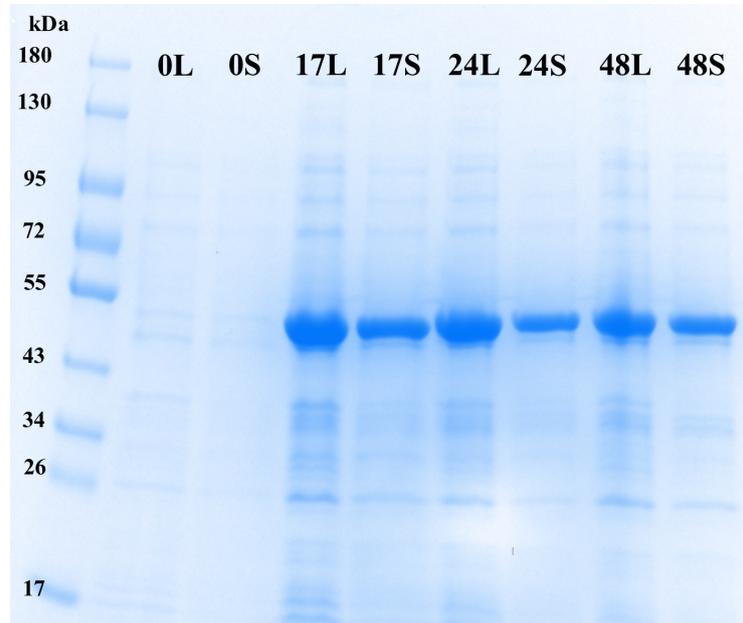


Figure 4.5: SDS-PAGE gel showing protein expression of pET14b-gFIHb. First line is the ladder, second and third one is at time 0 respectively lysate (L) and supernatant (S) sample. Fourth and fifth lines respectively show lysate and supernatant after 17 hours. Sixth and seventh lines respectively show lysate and supernatant after 24 hours and the eight and ninth respectively show lysate and supernatant after 2 days.

4.2.2 Expression of flavohemoglobin assemblage A with pJ401-gFIHb

FIHb (GL50803_15009) with the vector pJ401-gFIHb was expressed in both LB and TB (101) medium (all other conditions were the same as for pET14b-gFIHb) to find out which medium that would give the best expression. The protein is expressed at approximately 54 kDa (Figure 4.6). FIHb is presented both in the lysate (Figure 4.6) + and – from the SDS-PAGE (Figure 4.7) is from the same original plasmid, but those who are marked with + was sequenced in advance of the expression. From the SDS- PAGE (Figure 4.6) it looks like TB is the best choice for expressing FIHb in Rosetta (DE3).

Bands corresponding to the right size on the SDS-PAGE gel were sent for PMF, and the result confirmed that we had expressed the right protein.

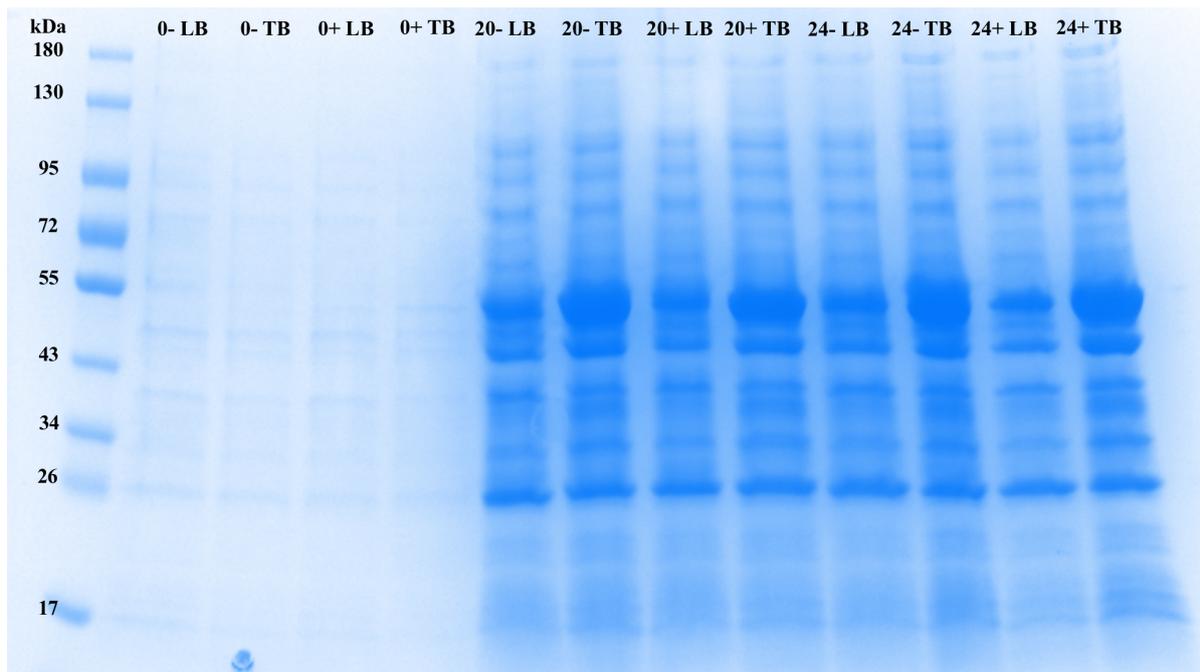


Figure 4.6: SDS-PAGE showing samples collected during expression of pJ401-gFlHb in both LB and TB media. Samples are collected before IPTG-induction (0) and after IPTG-induction after 20 and 24 hours. All the samples are from the lysate. + is sequenced pJ401-gFlHb and – is non-sequenced pJ401-gFlHb.

4.2.3 Expression of flavohemoglobin138

FlHb138 was expressed using Rosetta (DE3) cells as host in TB media. Aliquots were taken right before IPTG induction (0 h), and after 17 h and 24 h after IPTG induction, for analysing using SDS-PAGE. Samples from both supernatant (soluble) and lysate (cell lysate containing both soluble and insoluble proteins) fractions at each time interval, are presented on the SDS-PAGE (Figure 4.7). The protein is mostly represented in the lysate fraction (L), meaning that the protein might not be soluble in current buffer conditions or the cell lysis was not successful. The size of FHMP138 is 51 kDa with the His₆-tag and the cleavage site attached to the protein, which fits well with the thickest, bands shown on the SDS gel (Figure 4.7).

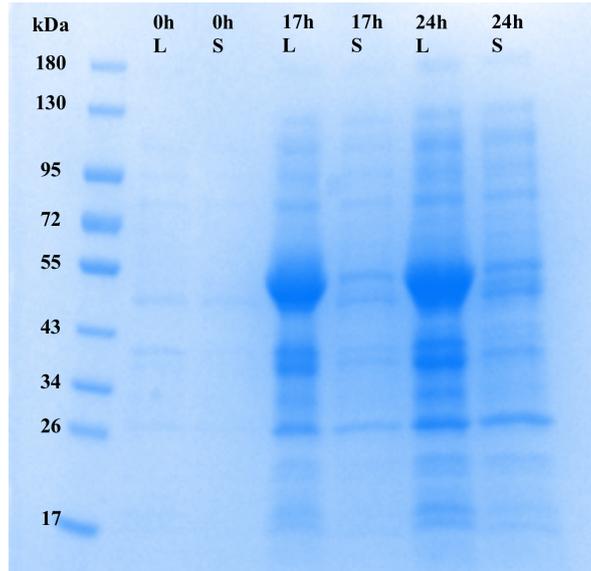


Figure 4.7: SDS-page gel showing protein expression of FIHb138. First line is the ladder, second and third one is at time 0 respectively lysate (L) and supernatant (S) sample. Fourth and fifth lines respectively show lysate and supernatant after 17 hours. Sixth and seventh lines respectively show lysate and supernatant after 24 hours.

4.2.4 Expression of nitroreductase 1

Expression test for NR1 was done in two different ways. In both cases, aliquots from different time intervals were collected for analysis using SDS-PAGE. Samples from the expression with Rosetta (DE3) cells in TB media, is shown on the left hand side from the SDS gel (Figure 4.8). Samples were collected right before IPTG induction (0 hours), after 17 hours and after 24 hours. The protein is expressed after 17 hours, but a higher expression level is reached after 24 hours. The protein is mostly in the lysate fraction, meaning that the protein is not that soluble in this case. NR1 size is 32 kDa. From the SDS gel (Figure 4.8) the thickest bands at approximately 35 kDa represents the NR1.

Samples from the expression in autoinduction media (ZYM-5052 autoinduction media) (109) is shown on the right hand side from the SDS-gel (Figure 4.8). The protein was expressed during the weekend (for 64 hours), and is presented mostly in the lysate fraction in this case as well.

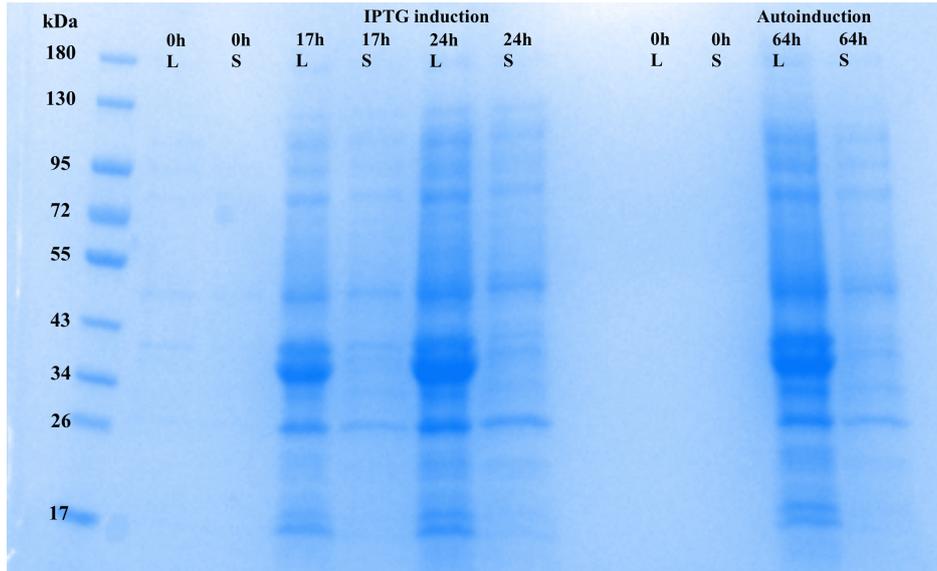


Figure 4.8: SDS-PAGE showing samples from expression of NR1, with IPTG induction on the left hand side and with autoinduction on the right hand side. L = lysate and S=supernatant. Samples from IPTG induction were collected right before IPTG was added, which is set to time 0, and then after 17 and 24 hours. Samples from autoinduction were collected right before the temperature was lowered, which is set to time 0, and then after the weekend (64 hours). The protein of interest, NR1, is shown as thick blue bands at approximately 34 kDa.

Bands corresponding to the size of the thickest bands from Figure 4.8 were sent for PMF. Results from PMF confirmed that the expressed protein was NR1.

4.3 Purification

4.3.1 Affinity purification of flavohemoglobin from assemblage A

His-tag purification using Ni-NTA column was successful for both the pJ401-gFIHb and the pET14b-gFIHb vector. The Ni-NTA resin turned from a blue color to a reddish brown color when the protein bound to the resin. The red color is likely to origin from the heme bound in the protein. The eluted protein solution had an intense reddish brown color. Samples from the purification steps were collected and are shown on the SDS-PAGE (Figure 4.9). The first line represents the lysate (L), the second line represents the supernatant (S), the third line represents the flow trough (FT) and the fourth line represents the elute (E). FIHb is presented in all of the lines, but it is the only one presented in E. FIHb is shown as a thick band at the SDS-PAGE at approximately 54 kDa.

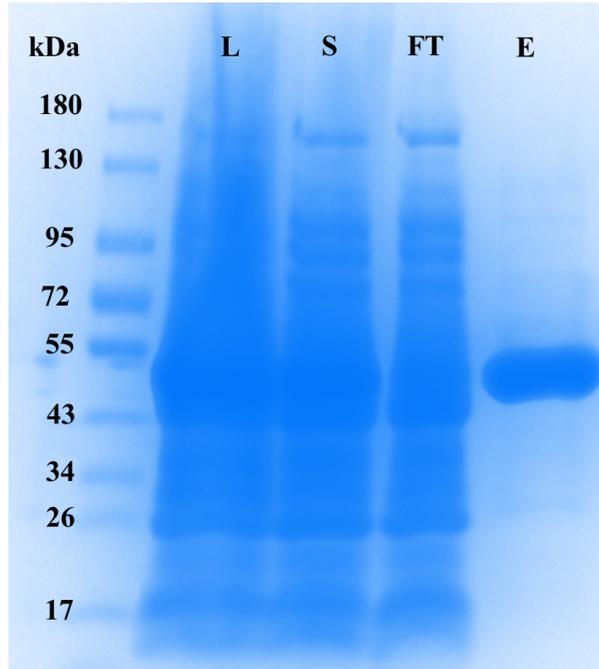


Figure 4.9: SDS-PAGE showing samples from the steps during His-tag affinity purification of FIHb assemblage A. L=lysate, S=supernatant, FT=flow trough, E=elute.

4.3.2 Affinity purification of flavohemoglobin138

For the His-tag affinity purification on FIHb138, the Ni-NTA resin turned from a blue color to a red color similar with FIHb from assemblage A. The samples collected from the purification steps are shown on SDS-PAGE (Figure 4.10). In this case, the elute was not as pure as for FIHb assemblage A. FIHb138 is presented as the thickest band in all of the lines, except wash number four (W4), at approximately 55 kDa from the gel (Figure 4.10) The size of FIHb138 is approximately 54 kDa.

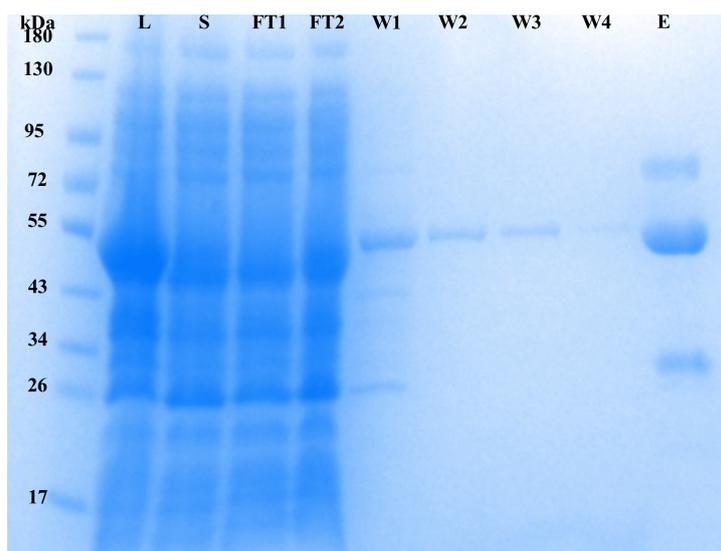


Figure 4.10: SDS-PAGE showing samples from the steps during His-tag affinity purification of FIHb138. L=lysate, S=supernatant, FT=flow trough, W=wash, E=elute. FT1 was collected and run trough the column, resulting in FT2. The column was washed four times before the protein was eluted.

4.3.3 Affinity purification of nitroreductase 1

His-tag affinity purification was used to purify NR1. The samples collected from the purification steps are shown on SDS-PAGE (Figure 4.11). The size of NR1 is 32 kDa without the His-tag and cleavage site on. The size of the protein will be larger because the His-tag and the cleavage site are attached to the protein. NR1 is presented at approximately 45 kDa on the SDS-PAGE (Figure 4.11).

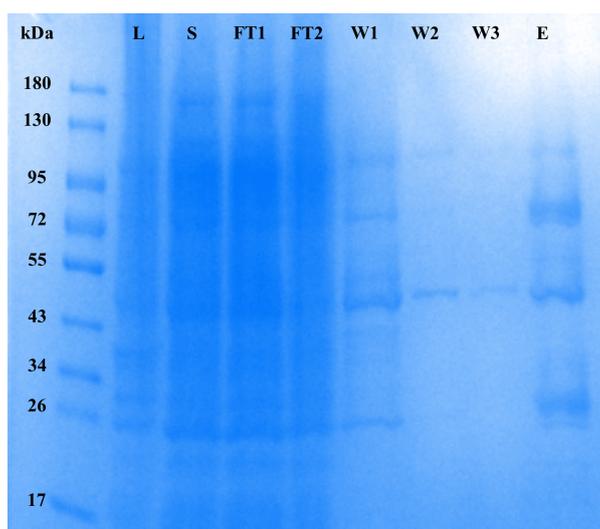


Figure 4.11: SDS-PAGE showing samples from the steps during His-tag affinity purification of NR1. L=lysate, S=supernatant, FT=flow trough, W=wash, E=elute. FT1 was collected and run trough the column, resulting in FT2. The column was washed three times before the protein was eluted.

4.3.4 Size exclusion chromatography

SEC was performed with FIHb from assemblage A and from assemblage B (with Superdex™ 200 increase 10/300). The elution profiles from both constructs are shown in Figure 4.13. The green line represents FIHb from assemblage A and the blue line represents FIHb138 from assemblage B. The two highest peaks represent FIHb from each assemblage. For FIHb assemblage A (green line), a smaller peak was eluting earlier than the main peak. Fractions from both the small peak and the main peak were collected in separate tubes. The fractions from the smaller peak were colorless, while the fractions from the main peak had a red color. The red color was however weaker than the protein solution had before SEC.

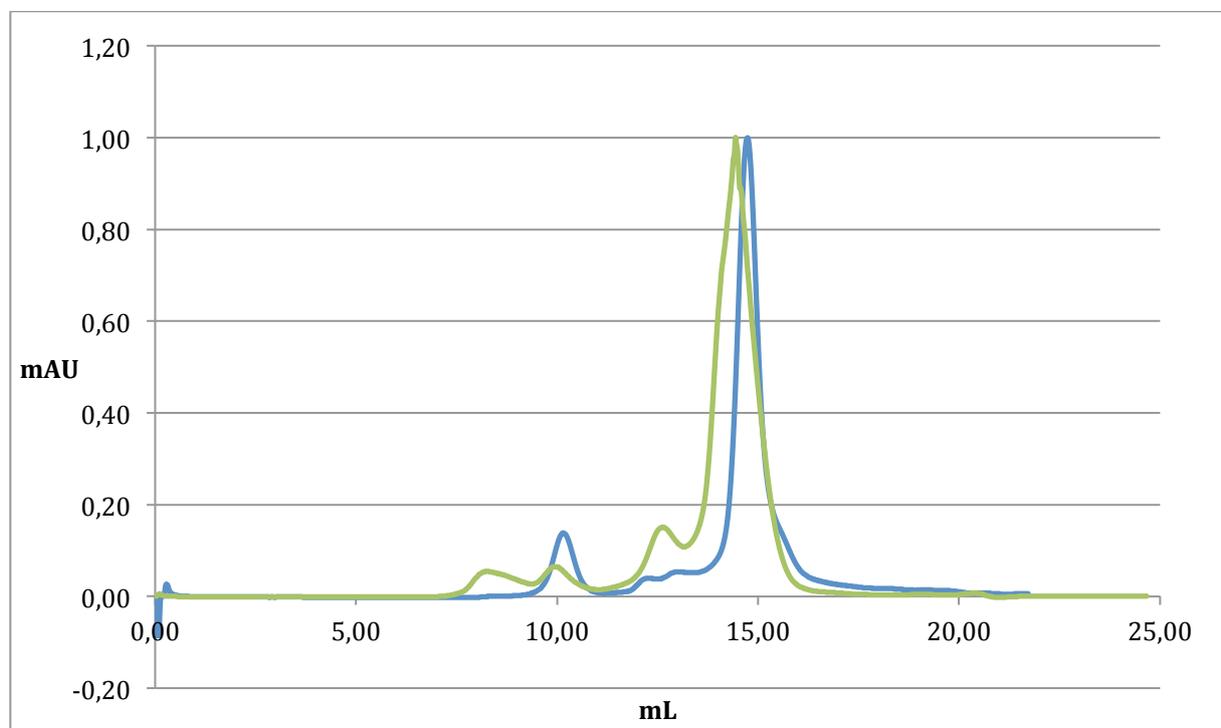


Figure 4.12: Elution profiles from SEC showing the elution profile from FIHb138 (blue) and elution profile from FIHb assemblage A (green) at UV280. Absorbance (mAU) is shown in a relative scale to fit the two peaks.

The fractions from the two peaks from assemblage A (green line in Figure 4.12), which were close to each other were collected and analyzed on SDS-PAGE. The bands representing each fraction were sent for PMF. Results from PMF, showed that the two bands were the same protein. One colorful variant of FIHb and one colorless variant of FIHb.

4.4 His-tag cleavage

The His₆-tag was not cleaved off with neither of the proteases TEV nor thrombin for FIHb from assemblage A. The cleavage of FIHb138 assemblage B with 3C protease seemed to work. The cleavage is shown at SDS-PAGE (Figure 4.13). There is small shift in the bands comparing the band representing FIHb138 sample without 3C protease and the band representing FIHb138 with 3C protease. The cleaved sample seems to be a few kDa smaller than the original FIHb138.

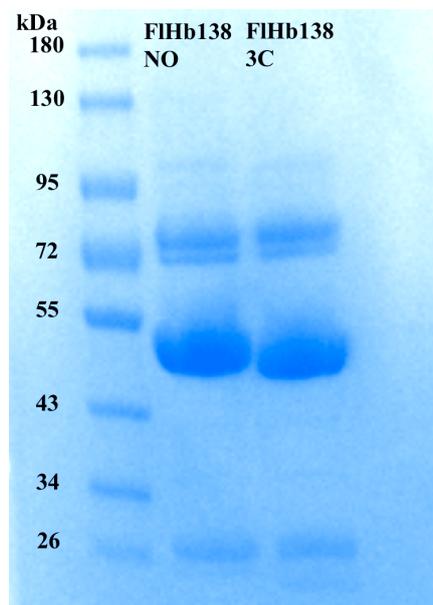


Figure 4.13: SDS-PAGE showing His₆-tag cleavage with 3C protease. The first line is the ladder, the second line is FIHb sample without 3C protease and the third line is FIHb with 3C protease.

4.5 Characterization of flavohemoglobin assemblage A

4.5.1 Light scattering

SEC-MALS was used to measure the molecular weight. The light scattering (LS) curve from SLS and the estimated molecular weight (Mw) are shown in Figure 4.14. The curve shows three peaks, whereas the first peak is probably aggregates, the second tiny peak (bump) is the colorless fraction and the third high peak is the FIHb. From the measurement, the molecular weight of FIHb was estimated to be 59.6 kDa. The theoretical Mw (calculated from ProtParam) Mw is 54.2 kDa with the His₆-tag and cleavage site attached to the protein. This shows that FIHb elutes as a monomer. Mw is shown as a red line in Figure 4.14.

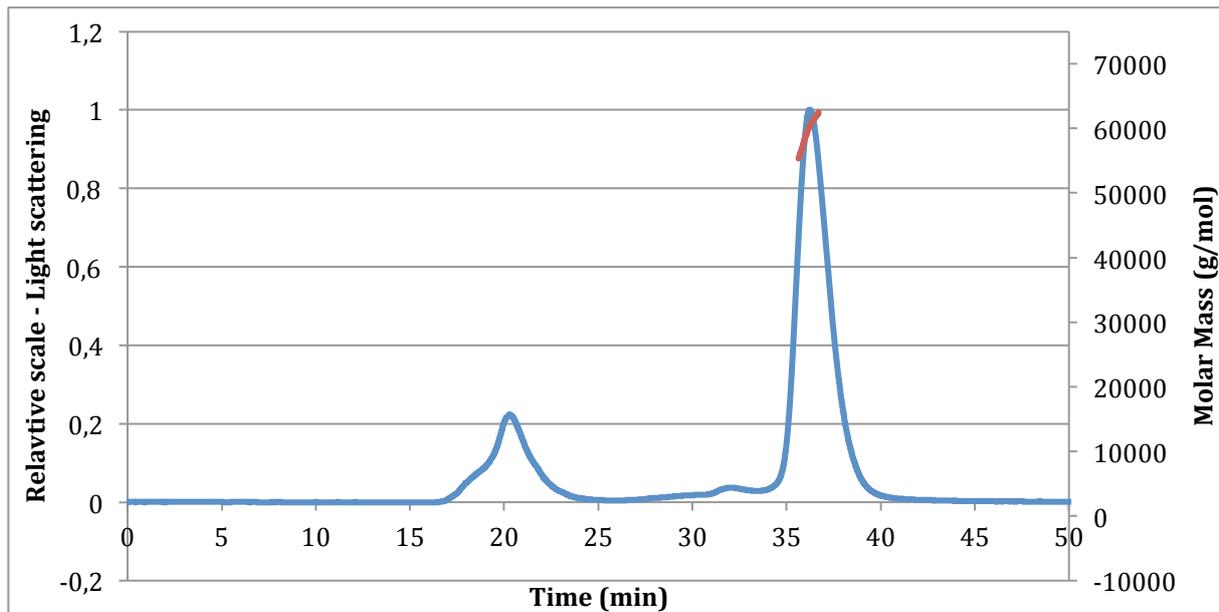


Figure 4.14: Curve from SEC-MALS, light scattering signal vs. time (min.). Estimated molecular weight for flavohemoprotein marked with red thick line (5.96×10^4 g/mol).

The second experiment of SEC-MALS was performed by Erik Hallin at Diamond before SAXS. The measurement was done from only affinity purified sample to have higher amount of the colorless variant of FIHb to see if the two species have different properties. The LS curve is presented in Figure 4.15 together with the estimated Mw. The two peaks in the curve represent respectively the colorless and the colorful variant of FIHb from assemblage A. The estimated Mw for the colorless variant of FIHb is shown as a red line in Figure 4.15. It was estimated to 128.2 kDa. The estimated Mw for the colorful variant of FIHb is shown as a green line in Figure 4.15 and it was estimated to 62.7 kDa.

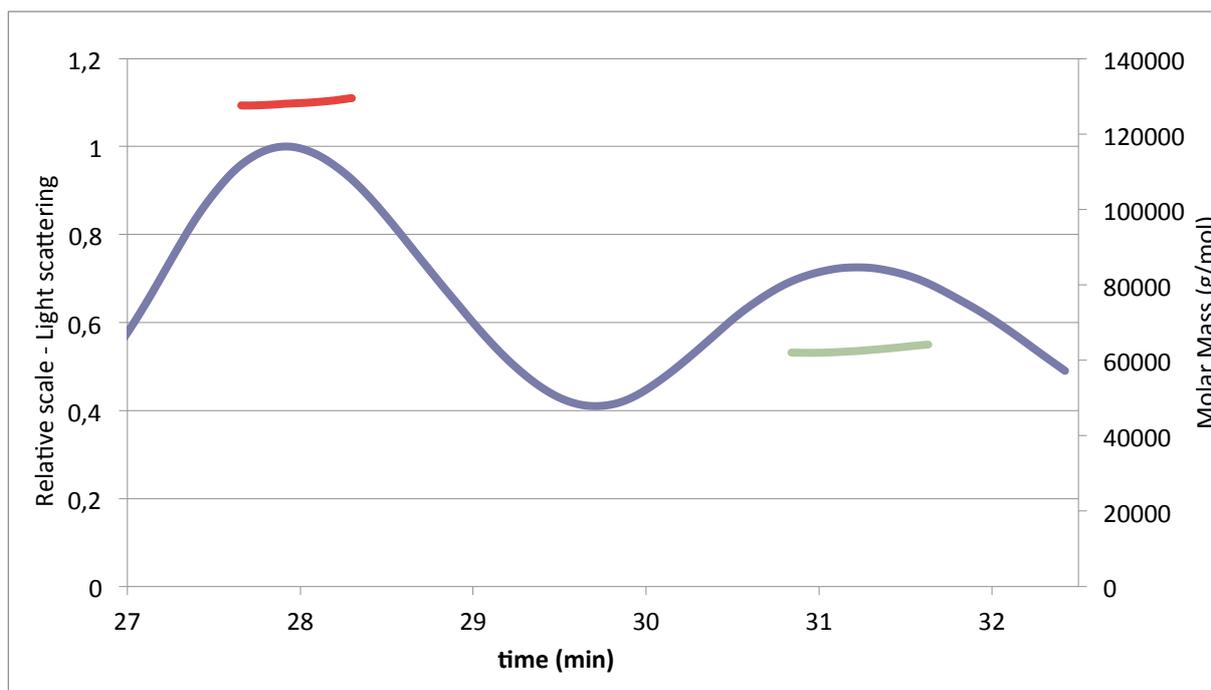


Figure 4.15: Light scattering curve from SEC-MALS in relative scale. The first peak represents the colorless FIHb from assemblage A, and the second peak represents the colorful variant of FIHb from assemblage A. The red line represents the M_w for the colorless variant (1.28×10^5 g/mol), and the green line represents the M_w for the colorful variant (6.27×10^4 g/mol).

4.5.2 Circular dichroism

CD was performed with FIHb from assemblage A. The CD spectrum of FIHb is shown in Figure 4.16. The shape of the curve illustrates the secondary structure of the protein. In this case it illustrates a combination of α -helices and β -sheets. CD was performed with changes in temperature as well, to estimate the T_m . This is shown in Figure 4.17. There is a transition between the line at 40 °C (orange) and 45 °C (blue) in Figure 4.17. This indicates that the T_m of FIHb is at approximately 45 °C, meaning that the protein unfolds at this temperature.

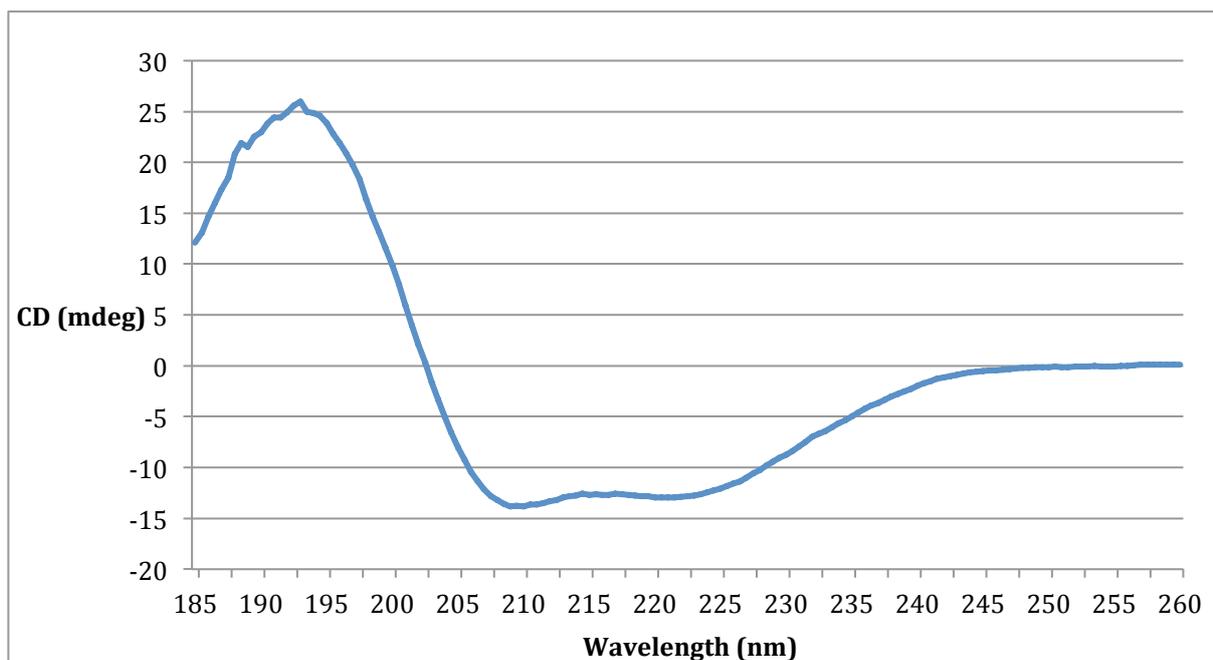


Figure 4.16: CD spectrum of FlHb from assemblage A.

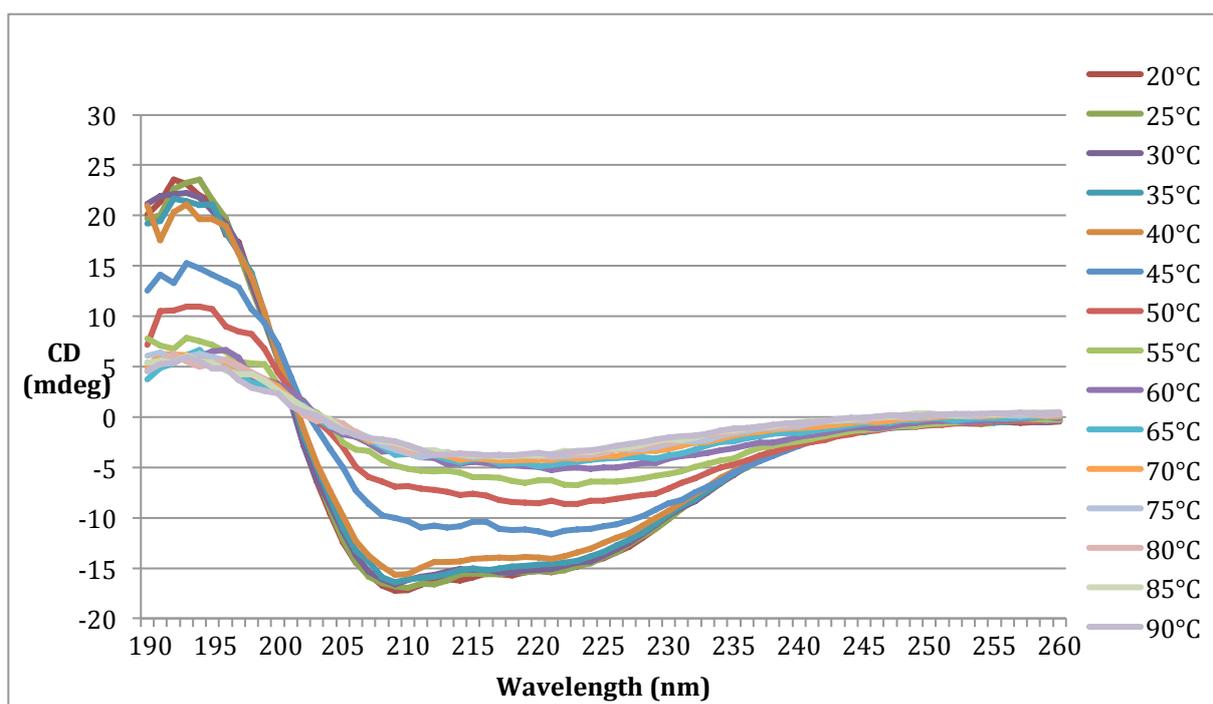


Figure 4.17: CD spectra with melting curve of FlHb from assemblage A. The melting point is around 45 °C (blue line).

Due to the observation of two variants of FIHb (colorful and colorless), samples from each of the variant were run on CD to compare the spectra. The spectra are shown in Figure 4.18. The spectrum of colorful FIHb is shown in blue, and the spectrum of colorless FIHb is shown in red in Figure 4.18. The spectra overlay well and no changes in the secondary structure can be seen.

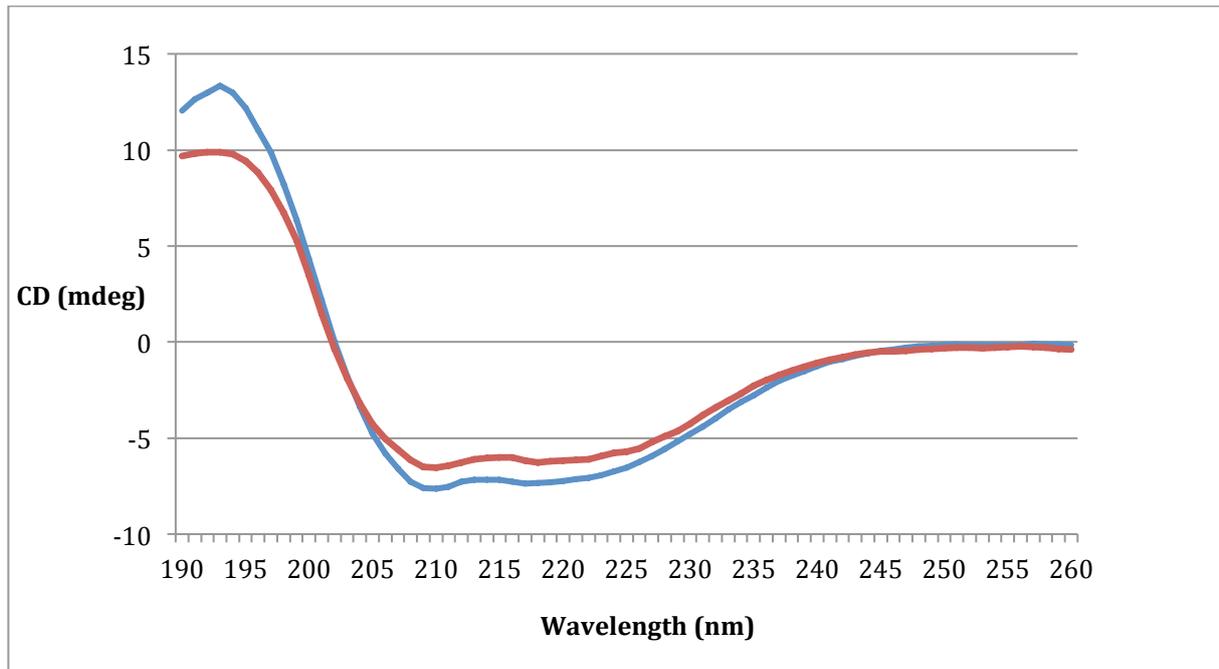


Figure 4.18: CD spectra of colorless (red) and colorful FIHb (blue) FIHb.

4.5.3 Thermoflour

The melting curve representing each well from the salt/buffer screen (Table 3.3) is presented in Figure 4.19. Each color line represents the T_m for each condition in each well. There were no large differences in the T_m s for each of the conditions. The melting temperature was between 42 °C and 44 °C in most of the conditions from Table 3.3. But the wells G9, G10, H9 and H10 (see Table 3.3) had slightly higher T_m (48 °C) values compared to the other wells. These wells had 20 mM buffer with pH 7.5, higher amount of salt (250-500 mM) and 10% glycerol. Based on these results, FIHb seems to be stable with higher concentration of salt and glycerol. We then decided to stay with the buffer (20 mM HEPES pH 7.5) we had used so far, with the addition of glycerol (5%) hoping that it would stabilize the protein.

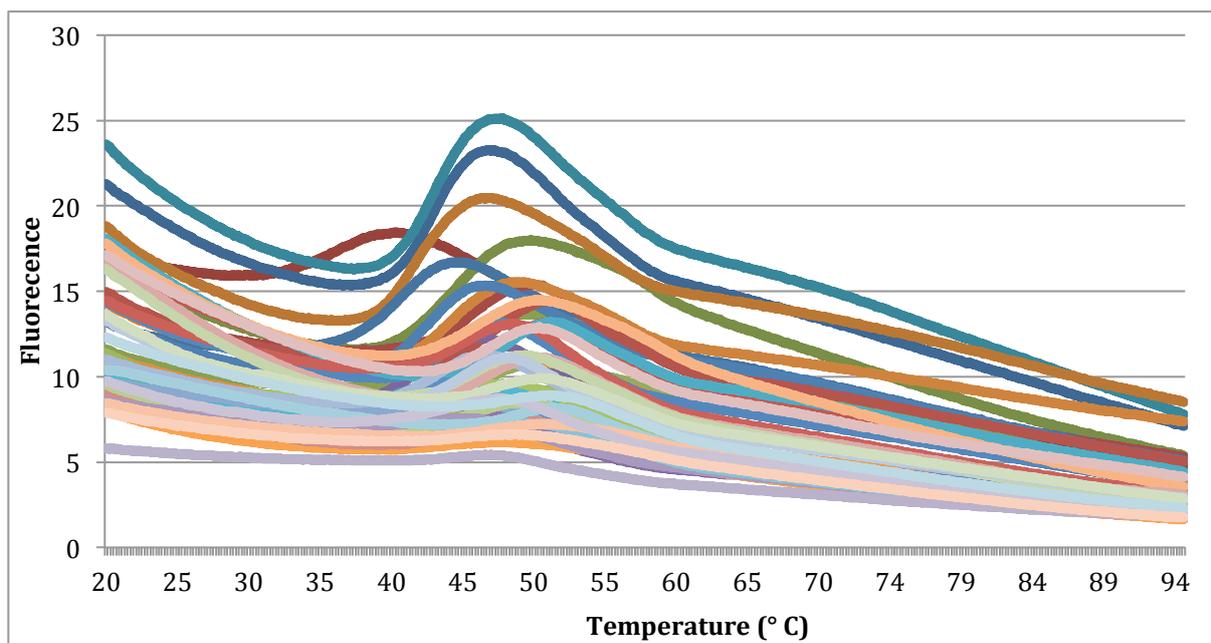


Figure 4.19: Melting curves from thermofluor. All conditions from the buffer/salt screen are represented in the figure showing that only minor changes were observed with this screening.

4.5.4 Small angle X-ray scattering

SEC coupled SAXS was performed with FIHb assemblage A (colorful and colorless variants). Collected SEC-SAXS frames were processed using a script written by Erik Hallin. The buffer region is defined from frames 50 to 150 (Figure 4.20). The buffer scattering was subtracted from the sample to obtain the protein scattering curve. The frames representing each peak of the protein were defined based on radius of gyration (R_g). R_g is linear in the regions of the frames. The first peak of the two highest peaks in Figure 4.20 is the colorless variant of the FIHb. SAXS data from this peak are defined from frames 338 to 348. The R_g values are linear for these frames shown in Figure 4.20 with red dots. The second peak of the two highest peaks in Figure 4.20 is the colorful variant of FIHb. SAXS data from this peak are defined from frames 367 to 378. The R_g values are linear for these frames shown in Figure 4.20 with green dots.

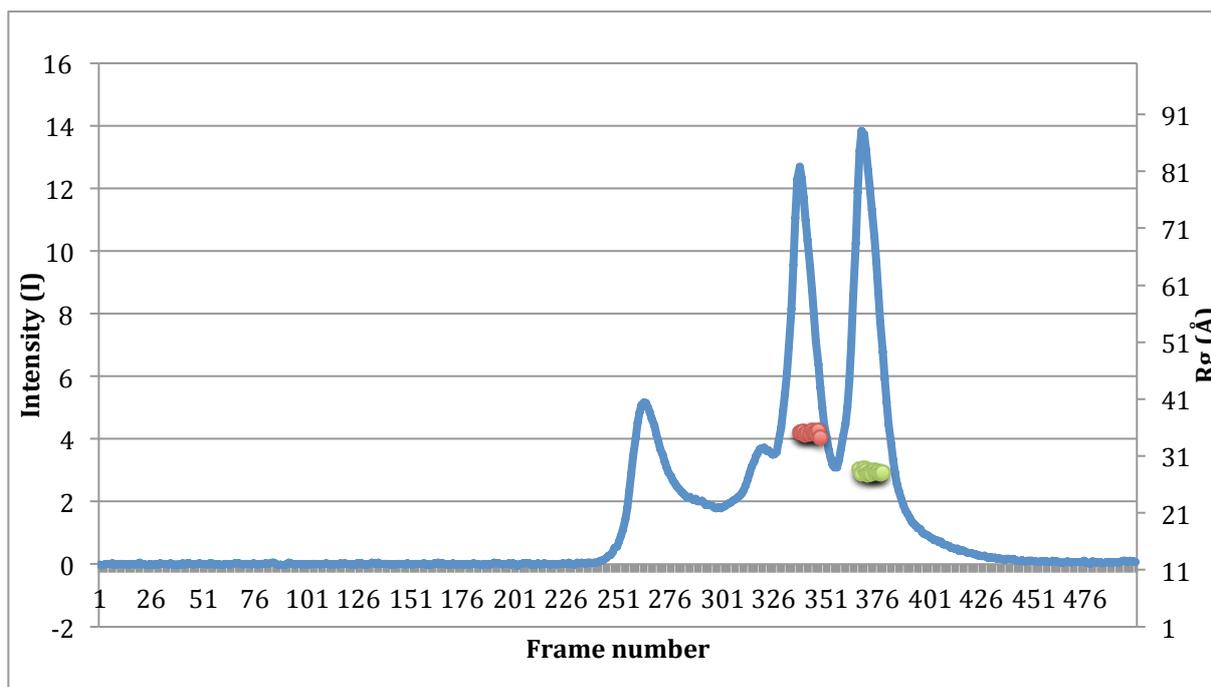


Figure 4.20: SEC-SAXS profile of FIHb containing two variants. The measured intensity is represented in the function of collected frames. The colorless variant is presented as the first peak of the two highest peaks close to each other, with its linear R_g region presented as red dots. The colorful variant is presented as the second peak of the two highest peaks close to each other, with its linear R_g region presented as green dots.

The scattering data from the colorless and the colorful variants are presented in Figure 4.21. A in Figure 4.21 is the SAXS curve (log plot) for the two variants. B in Figure 4.21 is Kratky plot of the colorful variant of FIHb (Kratky plot of the colorless variant looks similar). The curve has a bell-shape, which indicates that the protein is globular. 3D models based on the scattering data were created using GASBOR and the images were made using PyMol. The models of the overall shape of both the colorless and the colorful variants are shown aligned in Figure 4.21. The colorful variant has a more closed and globular shape compared with the colorless variant, which has a more open shape and the size seems to be larger. The M_w estimated to 94.7 kDa for the colorless variant and 53.6 kDa for the colourful variant, based on the porod volume presented in Table 4.3.

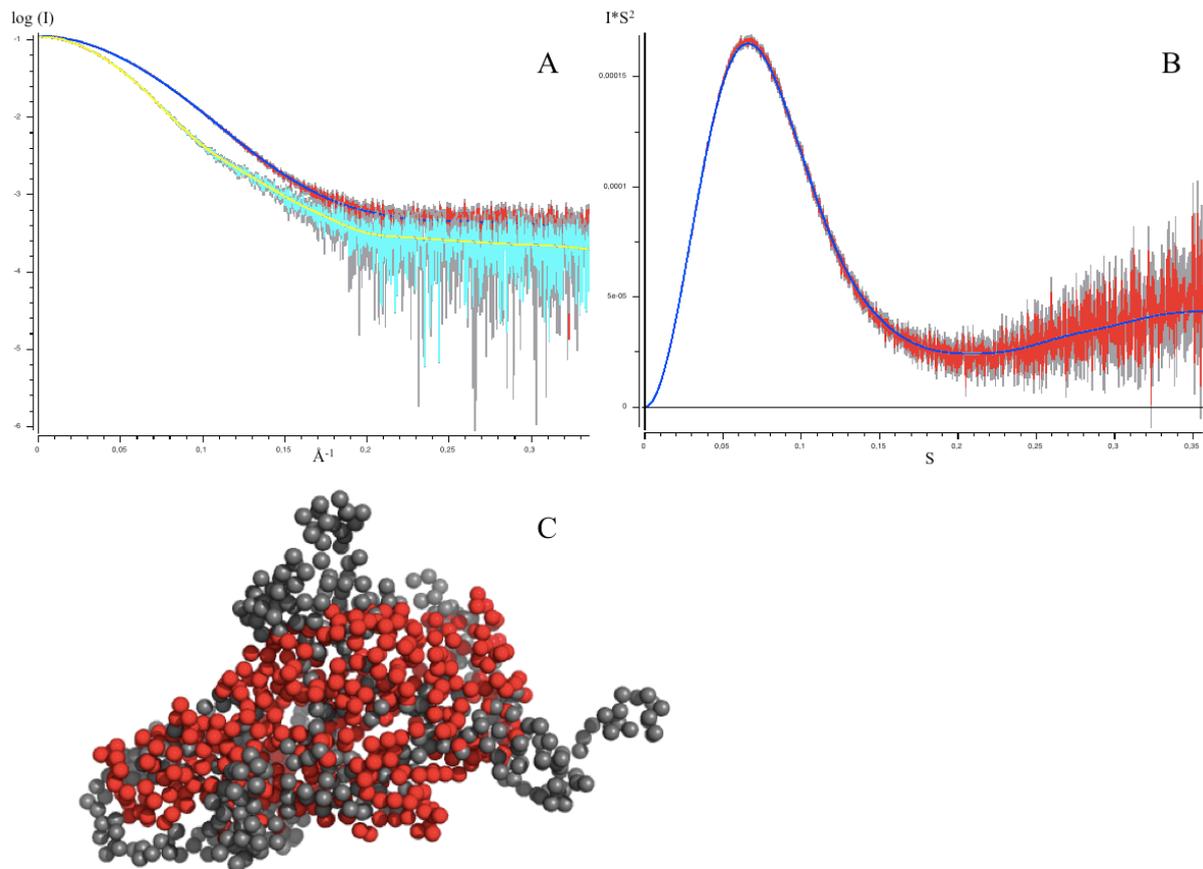


Figure 4.21: SAXS data from colorless and colorful variants of FIHb assemblage A. A) Logarithmic plot showing intensity in function s (inverse angstroms), of SAXS curved (raw data) with fitted data from colorless (turquoise raw data and yellow fitted line) and colorful (red raw data and blue fitted line) variants of FIHb. The plot is made with Primus. B) Kratky plot of colorful variant of FIHb (raw data is red and the line for averaged data is blue). The plot is made with Primus. C) Aligned overall shape of colorless (grey) and colorful (red) variants of FIHb. The model is made with GASBOR and created with PyMol.

Table 4.3: Results from Guinier and distance distribution analysis using Primus, molecular weight was estimated based on the porod volume.

	Colorless FIHb	Colorful FIHb
I_0	0.11	0.11
R_g (Å)	35.46	28.42
Porod volume	160964	91100.8
Molecular mass (kDa)	94.7	53.6

4.6 Protein crystallization

Protein crystals containing the characteristic reddish color of FIHb were not observed during crystal screening trials. Most of the conditions were prone to precipitate although lower concentrations of protein were tried in the crystallization experiments as well. However, square, transparent microcrystals were formed under the conditions presented in Table 4.4 (using PACT-premier). The protein solution of FIHb from assemblage A was only purified with affinity chromatography in advance. The concentration of the protein solution was 11 mg/mL. The microcrystals were colorless, but they were UV positive (see Figure 4.22) indicating them to be proteins. The microcrystals did not diffract when they were tested at the beam line P11 at DESY/PETRA III by Juha Kallio.

Table 4.4: Conditions where microcrystals were formed. Protein and the well solution were mixed 1:1 and the drop size was 600 nL.

Conditions in the protein solution	Condition in the well from PACT-premier
150 mM NaCl 20 mM HEPES pH 8 1 mM TCEP 5 % glycerol	0.2 M Na Sodium citrate (Na ₃ Cit) 20 %w/v PEG 3350 0.1 M BIS-TRIS pH 8.5

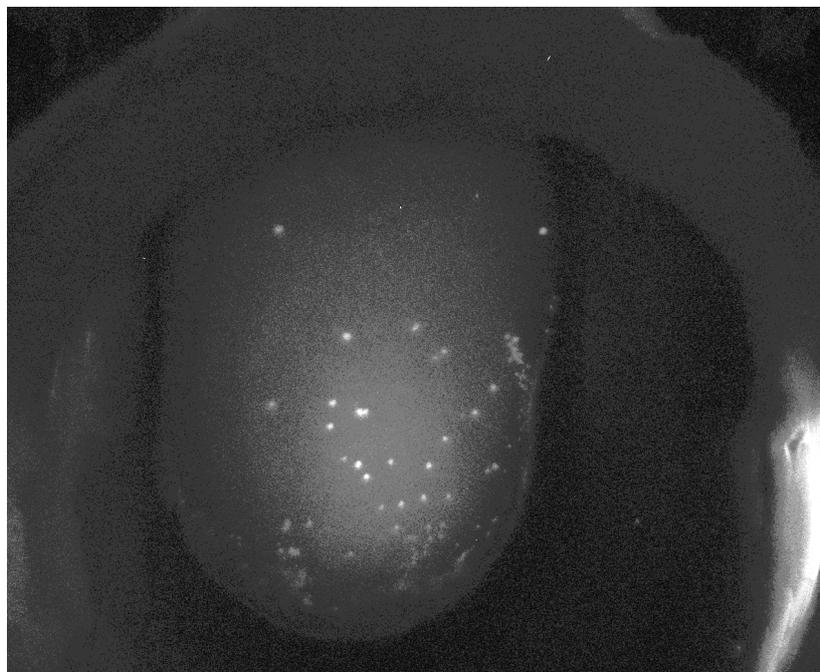


Figure 4.22: UV Fast image of microcrystals from FIHb from assemblage A (11 mg/mL). Image from ROCK IMAGER®.

5. DISCUSSION

FIHb from assemblage A was successfully expressed in *E.coli* and purified with both affinity purification and SEC. From the SEC elution profile, two variants of FIHb, one containing heme and one without heme, were observed. FIHb from assemblage A was characterized with CD, LS and SAXS. FIHb seems to consist of both α -helices and β -sheets as predicted. A solution 3D model for each of the variants was obtained by SAXS. The heme-containing (active) variant seems to be monomeric and have a globular shape, while the heme-less variant has a more open structural shape with a doubled size compared to the active variant. Stability tests were performed with thermofluor and CD, indicating that the T_m of FIHb is 40-45 °C and the protein is probably more stable in higher salt concentrations and addition of glycerol.

Two constructs of FIHb (FIHb1 and FIHb138) and NR1 from assemblage B were successfully cloned into pETNKI-his3C-LIC-amp vectors using competent cells. Further, FIHb138 and NR1 were for the first time expressed using *E.coli* as host cells, and purified with affinity purification. In addition, FIHb138 was purified with SEC. The His₆-tag was successfully cleaved off with 3C protease for FIHb138.

An overview of what have been done with each of the proteins is presented in Figure 5.1. The discussion is presented in a chronological order to get a better overview.

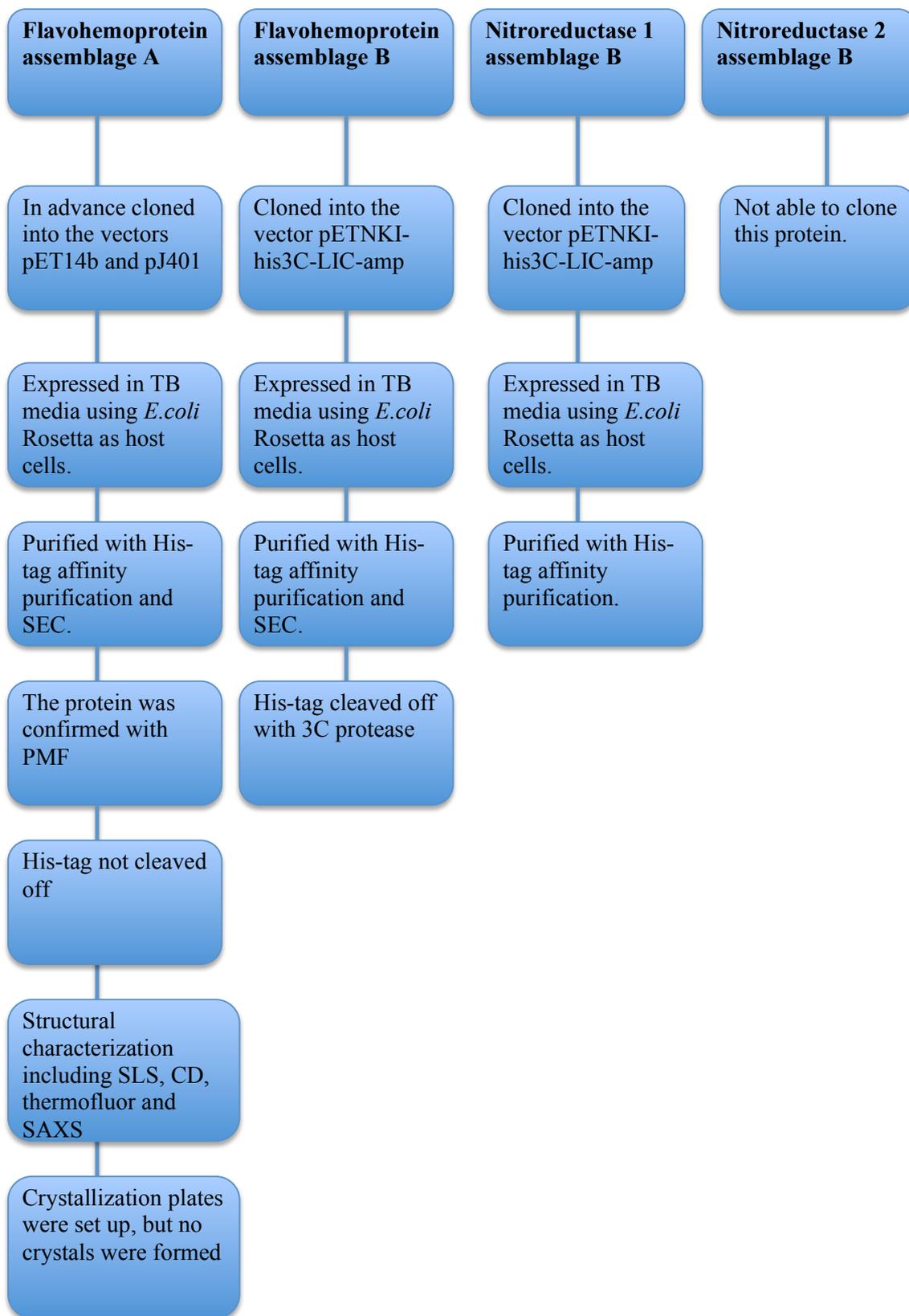


Figure 5.1: An overview of what have been done with each of the proteins in this study.

5.1 Genes used in this study

5.1.1 Flavohemoglobin

FlHb from assemblage A (GL50803_15009) is the reference genome of FlHb. This gene sequence is used in other studies as well (53, 117-119). The sequence of FlHb from assemblage B (GSB_151570) is from clinical isolates, which infects humans, and not from the reference genome used in the previous studies. FlHb from assemblage A and the construct FlHb138 from assemblage B have the same number of AA. They are aligned in Figure 5.2. The sequence identity is 72 %. The sequences are quite similar, indicating that their functions also should be similar. The structure and role of the long extended N-terminus of FlHb from assemblage B remains to be elucidated for the further research.

Ass. A	2	MTLSEDTLRAVEATAGLIAAQGIEFTRAFYERMLTKNEELKNIFNLAHQRTLQPKALLD	61
Ass. B	1	M LSEDT++AVEATA L+AAQG++FTRAFYERMLT+NEELK++FNL+HQRLRQPKALLD	60
Ass. A	62	SLVAYALNIRRINELYELKKGKGLPVPPEHWAELQGGFFSAAERVANKHTSFGIQPAQYQIV	121
Ass. B	61	SLVAYA +IR+INEL+EL+ +GLPVP E AELQGGFF+ AER+A+KH S GIQPAQYQIV	120
Ass. A	122	GAHLLATIEDRITKDKDILAEWAKAYQFLADLFIKREEEIYAATEGCKGGWRQTRTFRVE	181
Ass. B	121	GAHLLATIE+R+T DK ILA W+KAY FLA LF++REEEIY TE +GGWRQTR+FRVE	180
Ass. A	182	EKTRVNEIICKFRLVPAEEGAGVVEHRPGQYLAI FVRSPEHFQHQQIRQYSIISAPNSAY	241
Ass. B	181	EK ++ E I +FRLVPAE+G V H+PGQYLA+FVR P H+QIRQYSI SAPN Y	240
Ass. A	242	YEIAVHRDEKGTVSRYLHDYVSTGDLI FLRYLEADEQAPADTQASQEFQML	301
Ass. B	241	YEIAVHRD++ TVS YLHD+V+ GDLI :L Y E QA AD Q S E L	299
Ass. A	302	QSGAINFAAEKTMPIVLISGGIGQTPLLSMLRFLAQKEGKETARPIFWIHAAHNSRVRAF	361
Ass. B	300	GA+NFAAE+ PIVLISGGIGQTPLLSMLRFLA+KEG+ RPIFWIHAAH+SR RAF	359
Ass. A	362	KEEVDAIRETALPSLRVVTFLSEVRAT-DREGEDYDFAGRINLDRISELTKLEADNANPH	420
Ass. B	360	K EVD AI+ T LP LR TFLSEV T D++GEDYDFAGRI+LDR+ L +LEAD ANPH	419
Ass. A	421	YFFVGP TGFMTAVEEQLKTKSVPNSRIHFEMFGPFK	456
Ass. B	420	YFFVGP GFM AVE+QLK SVP RIHFEMFGPFK	455

Figure 5.2: Amino acid alignment of FlHb from assemblage A (ass. A) and FlHb138 from assemblage B (ass. B).

5.1.2 Nitroreductases

The gene sequence of NR1 (GSB_153178) and NR2 (GSB_22677), are from clinical isolates obtained from infected humans, and not from the reference genome. The NRs described in this study are hypothetical proteins, meaning that their existence have been predicted (in this case from clinical isolates), but there are lack of evidence that they are expressed in vivo (120). NR1 has now been expressed in this project.

5.1.3 Other metronidazole activating and O₂ detoxifying enzymes have been characterized

PFOR and TrxR are presented in Figure 1.3 and have shown to activate MTZ. These two enzymes have been characterized in previous studies. The crystal structure of TrxR from *Giardia* has been solved (PDB code 5M5J) (121). TrxR (strain ATCC 5803, WB clone C6) was expressed in *E.coli*. *In silico* studies with MTZ covalent bound TrxR was performed. From their results, it seems like MTZ has TrxR as the primary target (121). PFOR has been characterized (44), but the crystal structure needs to be solved. PFOR was expressed anaerobic in *Giardia* trophozoites, purified and the activity was detected (44).

NADH oxidase and FDP (1.4) have shown play a major role in O₂ detoxifying in *Giardia*. NADH oxidase (GL50803_33769) has been expressed in *E.coli*, purified and biochemically characterized (122), but a deeper understanding of its structural and functional properties are required. For FDP on the other hand, the structure has been solved (PDB code 2Q9U) (123). The protein was expressed using *E.coli*. Their results show that FDP scavenge O₂ more efficiently than NO (123). One study (124) demonstrated that PFOR are involved in MTZ activation and resistance, TrxR reduces MTZ and a decrease in NADH oxidase FMN reduction activity was found.

More research is needed for the proteins that already have been characterized and the proteins used in this study, to get the full picture of how *Giardia* is able to detoxify MTZ, NO and O₂.

5.2 Cloning

The cloning was successful for both the one-step SLIC method (100) and the original T4 DNA polymerase exonuclease treatment from ACEMBL (98). The one-step SLIC would be beneficial with respect to time, and may be used in future studies.

All preparation before PCR run was done under normal laboratory conditions. This could lead to contaminations from the environment. The cloning of the FIHb25 construct of FIHb and the constructs of NR2 was not successful. There may be several reasons for the unsuccessful cloning, and contaminations from the environment may be one of them. Another reason could be low concentrations of the constructs. Table 4.2 shows that the concentrations of the amplified genes after gel purification with PCR cleaning kit were low, but that applies to all of the constructs.

5.3 Expression

The FIHb pellets from expression using pET14b-gFIHb were chosen for structural studies. This leaky vector has slower expression rate for the protein than the vector that require IPTG induction, resulting in the heme synthesis to better keep up with the expression rate. This leads further to more heme-containing FIHb (125).

5.3.1 Time, temperature and obtained protein concentration

Expression of FIHb (GL50803_15009) with the vector pET14b-gFIHb seemed to best after 17 h from Figure 4.5. pET14b-gFIHb is a leaky vector, which can be favourable to express properly folded proteins and especially for heme synthesis to keep up with the protein expression (126). Lower protein concentrations, and slower rates of expression could also improve transcription and translation and favour proper folding of the proteins (126). FIHb seemed to be properly folded in this case, even if the time of incubation was 42 h, but the level of expressed FIHb was highest after 17 h. The expression could have been stopped after 17 h to get a better yield of the protein, even if the yield was high after both 24 h and 42 h. 24 h was chosen for larger scale expression because the difference between 17 h and 24 h was minimal. Better yield could also have been obtained by lowering the temperature (127), as a reduction in expression temperature is associated with increase in soluble proteins (128). Lower temperatures could have been used for the expression of FIHb138 and NR1, because

they were mostly presented in the lysate fractions of the samples collected during expression tests (Figure 4.7 and 4.8). This was not done, because of time limitation of the project. It should however be tried in further research.

5.3.2 Expression with addition of IPTG

The expression for pJ401-gFIHb failed the first times. The cells did not express FIHb. The cells were growing rather slowly while incubating and the pellets did not turn pink like they did for pET14b-gFIHb, and no clear thick bands were seen on the SDS-PAGE gel. The difference between the expression of FIHb using pJ401-gFIHb compared with pET14b-gFIHb was the addition of IPTG. IPTG from a new stock for pJ401-gFIHb was tried. After this, expression of FIHb using pJ401-gFIHb worked well as presented in the result section (Figure 4.6).

The IPTG concentration used in this study was 1 mM. 1 mM IPTG has been used to express FIHb in other studies as well (129, 130). Also as low concentration as 0.1 mM have been used to express FIHb (131). The optimal concentration of IPTG for induction can vary from protein to protein. If the concentration is too low, the expression will result in low levels of expressed protein. If the concentration is too high, the protein can aggregate and be insoluble (132). Different concentration intervals between 0.1 mM and 1 mM IPTG should be tried during expression of new FIHb constructs in further studies to find the optimal expression conditions.

5.3.3 Cell lysis

In this project, the pellets collected from expression were lysed with lysozyme (0.1 mg/mL) in addition to sonication. Lysozyme is an enzyme with the ability to break the bacterial cell walls (133). High concentrations (2.5 mg/mL) of lysozyme will result in a decreased lysozyme activity, meaning less lysed cells (134).

Sonication uses ultrasonic waves to generate localized areas of high pressure, resulting in cell break down (135). Sonication can lead to unfolding and aggregation of the proteins because of the heat generated and the mechanical forces applied (136). To avoid high temperatures, the sonication was performed while the bacterial solution was kept on ice. The combination of lysozyme and sonication seemed to work well for the expressed FIHb from assemblage A and

B. The smaller samples collected during expression and analyzed on SDS-PAGE were only sonicated before analyzed with SDS-PAGE, because the samples were so small (1 mL). For FIHb from assemblage A, the protein was present with both samples before (L) and after sonication (S) (see Figure 4.5). For FIHb138 (Figure 4.7) and NR1 (Figure 4.8) the proteins were mostly represented in the samples that were not sonicated. This can be a result of poor solubility in the condition used under the expression process or it can be a result of poor sonication. Alternative lysing methods should also be tested in future studies.

5.4 Purification

5.4.1 Affinity purification – binding to the resin

Purification using Ni-NTA affinity chromatography is a straightforward method. The purified product was pure for FIHb assemblage A (Figure 4.9), but both FIHb138 (Figure 4.10) and NR1 (Figure 4.11) had some impurities in the purified product. The impurities from NR1 were sent for PMF, and were recognized as enzymes (transferase/decarboxylase) from *E.coli*. *E.coli* was used as host cell for the expression of both FIHb and NR1, but it is not common that host cell proteins have histidine residues on their surfaces that can unspecifically bind Ni-NTA, especially when the expression level of the target protein is low (137). Some proteins of the host bind weakly to the resin, but can be removed by including a low concentration of imidazole (20 mM) for competitive binding (137). Although this was attempted, the impurities were still there. One possible explanation can be that proteins from the host cell can potentially cluster on the surface of the folded protein and mimic a His-tag, or proteins with natural ion binding sites will also bind to the resin and affect its selectivity. In this cases, host impurities will be co-eluted with the protein of interest (137).

In all the cases of using Ni-NTA affinity chromatography, the protein was in the flow through and in some cases also in the washing buffers. This indicates that not all of the protein was bound to the resin, even though the Ni-NTA resin has a high capacity (60 mg of protein per mL of resin) according to the manufacture.

To reduce unspecific binding, other metal affinity resins could also be tested. For example resin loaded with cobalt requires two consecutive histidines to meet the requirements of metal coordination and that increases the specificity to the his-tagged proteins (138).

5.4.2 Affinity purification – chemical environment

Different salt concentrations and pH were tested for NR1. It is common to use NaCl (0.3-0.1 M), and a slightly basic pH (137). NaCl is added to minimize ionic interactions between the protein and the resin (137). A reducing agent is added as well (0.5 mM TCEP) to prevent oxidation of the protein and to keep the surface exposed cysteine free of forming intramolecular disulfide bonds (139). In the first test purification of NR1, lysis buffer 1, wash buffer 1 and elution buffer 1 with pH 7 were used (see appendix 1). This was not successful. The second and third test purifications were done simultaneously with respectively lysis buffer 2, wash buffer 2 and elution buffer 2, and lysis buffer 3, wash buffer 3 and elution buffer 3. The buffers for the second purification test had an increased salt concentration (see appendix 1) and a pH of 5. The theoretical isoelectric point (pI) value of NR1 is 6.25. When the pH has the same value as the pI, the protein has a net charge of 0 (56). The selected pH during the purification should therefore be at least one unit above or under the pI of the protein to keep the protein charged (140, 141). Increasing the salt concentration will increase the ionic strength and therefore increase the solubility of the protein and prevent aggregation (142). The third test purification had the same increased salt concentration, but the pH was adjusted to 8.5. The purification with a pH of 5 seemed to be better for the solubility, but the pH is not ideal for the resin (according to the manufacture), and a last attempt using pH 6 was performed (all other conditions were the same as for the second purification). This was unsuccessful. Maybe due to the theoretical pI of 6.25 as the protein should be less soluble close to its pI value because the protein is not charged and therefore less soluble.

Suitable chemical environment should be extensively screened during proper solubilization screening were different compositions of lysis buffers are used in parallel, to see what kind of conditions could possibly affect the solubility of the protein sample.

5.4.3. Freeze the purified protein

After purification of the protein of interest, it is important to store the protein solution in a proper way. The storage method of choice depends on the stability of the protein and the storage time (143). For short term (up to 24 h) storage, the protein solution can be kept at 4 °C. If the protein is stored for weeks or months, it is necessary to freeze the protein (-20 °C). And further, if the protein solution is stored for even longer terms (months to years), it is recommended to store the protein at -80 °C (143).

In this project, the protein solution of FIHb assemblage A, was first stored at -20 °C. When the protein solution later was thawed, the proteins were precipitated. This means that FIHb cannot be stored in the freezer, and other storage options for the proteins were needed. Other storage options for the proteins is to lyophilize the proteins and then store them at either 4 °C or -20 °C, but not all proteins are stable during freeze-drying processes (144). Another option would be to store them in a solution consisting of 50 % glycerol at -20 °C. Snap freezing with nitrogen, with addition of 5-50 % of glycerol and storage at -80 °C could also be an option (143). Snap freezing with liquid nitrogen with addition of 5 % glycerol was chosen in this project and was successful.

FIHb protein solutions with different concentrations (2-18 mg/mL) and volumes (200-500 µL) were stored at -80 °C in separate eppendorf tubes. When thawing the protein solutions later, the protein solutions with the highest concentration (18 mg/mL) were precipitated. Diluted protein solutions are more prone to inactivation and loss of protein, and therefore protein solutions in more concentrated form is recommended (145). Based on the observation that protein solutions with higher concentrations were precipitated, we diluted the following protein solutions to lower concentrations (5-10 mg/mL), to allow storage. Smaller volumes (200 µL) were favourable in order to make the freezing process with liquid nitrogen faster.

5.5 His-tag cleavage

The His₆-tag of FIHb138 was successfully cleaved off using 3C protease. In this case, the cleavage site is exposed to and recognized by the protease, resulting in a successful cleavage.

For the flavohemoprotein from assemblage A, the advantage of using pJ401-gFIHb compared with pET14b-gFIHb is that the His₆-tag can be cleaved off using TEV protease. TEV is a widely used protease to cleave the His₆-tag due to its high specificity (146). In theory this enzyme works well for cleaving the tag off. This was not the case in this project. Using pET14b-gFIHb, the His₆-tag can be cleaved off using thrombin. Cleavage with thrombin is not as specific as for TEV, and the cleavage can occur at other sequence sites in the protein (147). In this project it seemed that also the thrombin cleavage failed. Considering that the cleavage trials in both cases failed, the explanation may be related to the folding of FIHb

construct. The cleavage site may be buried or blocked due to the protein folding, so that the site recognized by protease is not accessible.

Cleavage of small affinity tags such as His₆ is not required to grow protein crystals, and may not affect the crystallization process (148). Cleavage of the His₆-tag is still wanted, because it is not a part of the original protein if interest.

5.6 Colorful and colorless variants of flavohemoglobin from assemblage A

The two variants of FIHb from assemblage A, were first observed from the SEC elution profile. The two peaks were both conformed to be FIHb from PMF. Fractions from the main peak had a reddish brown color, while fractions from the smaller peak were colorless. The reddish brown color comes from the heme bound inside the protein. Considering the one variant of the protein, is missing the reddish brown color, it is likely that the heme is not bound to this variant. The heme-containing one, is the one if interest, because it is the active FIHb.

From the SAXS model (Figure 4.21), the overall shape of the two variants indicates that the colorless variant had a more open structure compared to the colorful variant. The colorful variant had a globular, more closed structure. The FIHb structure from *Alcaligenes eutrophus* is given in Figure 1.5, where the heme pocket is illustrated. It is likely that heme is bound at the approximately same place in the structure of FIHb from *Giardia*. Based on the structural differences between the colorful and the colorless variant, and the lack of heme in the colorless variant, one suggestion would be that the heme keeps the structure more globular. When the heme is released, FIHb might undergo conformational changes. Based on our results we can suggest that colorless variant might form a dimer in solution. From the Mw estimated from both SAXS and SEC-MALS, it seems like the size of the colorless variant is doubled compared to the colorful variant. One explanation of the doubled size can be that two protein molecules from the colorless variant are bound, forming a dimer. One possibility is that when the heme is released, and the structure opens up, it might require another FIHb molecule to bind and cover the exposed more hydrophobic surface. Probably only the colorful variant is active of which would be requested for the further research.

5.6.1 Addition of heme

The color change of FIHb has been seen in other studies as well from other organisms, caused by loss of both cofactors; heme and FAD (54, 130, 131, 149). Heme reduction may be achieved by an associated FAD-containing reductase domain (149). The heme content of FIHb can be increased by addition of heme and FAD before gel filtration (54). If the heme is lost during gel filtration, reconstitution of heme and FAD after purification may help to obtain more heme bound to FIHb (130).

Heme was not added in this project due to lack of time to find a suitable protocol for this. Heme is challenging to use because of its solubility. Heme is soluble in diluted ammonia, NaOH and other strong organic bases (150). By dissolving heme in such a strong base, it will probably change the chemical environment of the protein even with small additions of heme. Another problem with heme is that it is expensive and the amount of heme required to keep the heme inside FIHb is not known. A third problem with heme is that we do not know if it will get stuck in the column during gel filtration or affect the purity of the protein sample. One possibility of doing this would be to add the heme before gel filtration followed by buffer exchange and then gel filtrate.

5.7 Exposure to oxygen

All the experiments in this project were performed under aerobic conditions, while *Giardia* is a microaerophilic parasite. In reality, giardial FIHb and NRs would not be exposed to the same amount of O₂ as during the laboratory work in this project. Protein expression, purification and structural studies are more challenging to perform under anaerobic conditions, and were therefore not done in this project. FIHb detoxifies NO using O₂ as a co-substrate (54). FIHb from *E.coli* detoxifies NO in aerobic conditions, but not in anaerobic conditions (151). Based on this, and the fact that FIHb was expressed and characterized under aerobic conditions it does not look like O₂ is harmful to the protein. But it would be of interest to try to use anaerobic growth conditions in addition to anaerobic purification steps, to compare the proteins under each condition, to see if O₂ has an effect on the protein. Activity assays would also be of interest to perform for aerobic and anaerobic expressed FIHb, to compare the enzymatic activity with or without the absence of O₂.

5.8 Crystallization

5.8.1 Optimization of crystal conditions

Screening for crystallization conditions was done with different concentrations of FIHb from assemblage A. The concentration varied from 5-12 mg/mL, with minimal change in the protein behavior in the corresponding drops on the plate. The proportion of precipitating drops and clear drops were quite similar for each of the concentration variations of the protein. When the drops are clear, it means that the protein-solvent interactions are stronger than the protein-protein interactions (152). When the drop on the other hand is producing precipitants, it means that the protein-protein interactions are stronger than the protein-solvent interactions causing the solubility of the protein to decrease too much (152). Crystals are formed when the concentration of protein is increased until the solubility is exceeded and the protein comes out of solution (56).

Once some crystals are formed, the optimization of the crystal conditions can begin (93). We obtained some microcrystals that were UV-positive (Figure 4.22), but unfortunately the microcrystals did not diffract. Crystals were grown under the conditions presented in Table 4.4. Each of the factors from the Table 4.4 could be systematically varied to find the optimal crystallization condition for this hit. Although it might not be worth of optimizing because the crystals were colorless, and were probably not formed from the active FIHb.

5.8.2 Homogeneity

Homogeneity is important for successful crystal growth. The protein solution must be purified to homogeneity or as close as possible to homogeneity (94). The observation of FIHb from assemblage A, separating into two different variants during SEC, indicates that the FIHb protein solution is not homogeneous. Even though fractions from the colorful peak from the elution profile were collected and used for crystallization trials, it seems like the protein is not stable over time. Crystallization trials were set up for both the colorful and the colorless variant of FIHb from assemblage A with the same conditions for both variants. The crystallization set up was done after SEC to separate the two variants. Resulting in no crystals. By separating the two variants, it is likely that the two separate protein solutions would be closer to homogeneity, than combining the two variants. Reconstitution of heme and FAD after purification may result in a homogenous solution (130).

5.8.3 Obtained flavohemoglobin crystals in literature

In the publication of crystallization of FIHb from *Alcaligenes eutrophus* (64) they describe the crystals as yellow-brown colored with a needle-shape. The crystals were formed in 18 °C under the crystallization conditions: 14 % PEG3350, 0.2 M NaCl and 0.05 M sodium citrate, pH 5. In a master thesis from University of Oslo (130) they were able to grow crystals of the active FIHb from *Bacillus cereus*. The crystals formed were red or orange needles under various conditions in room temperature. The crystals did however not diffract under X-ray. Based on these two papers, it is likely to suggest that the crystals from FIHb from *Giardia* should be colorful and have a needle-shape. Our obtained microcrystals were much smaller, square and colorless (transparent) compared to the FIHb crystals obtained in the literature. Based on this, it is likely that our microcrystals were not formed by FIHb, but by something else. The microcrystals might be formed by precipitants.

6. CONCLUSION

6.1 Flavohemoglobin

FIHb is a NO detoxifying enzyme found in *Giardia*. Both the full length (FIHb1) of FIHb from assemblage B and the construct (FIHb138) corresponding to FIHb from assemblage A, were successfully cloned into competent *E.coli* cells using pETNKI-his3C-LIC-amp as a vector and the method SLIC. FIHb138 was further expressed in Rosetta (DE3) cells in TB medium.

FIHb from *Giardia* assemblage A was received in two different vectors (pJ401-gFIHb and pET14b-gFIHb). The protein was expressed in *E.coli* and purified. From the SEC elution profile, it was observed two variants of FIHb, one containing heme and one without heme. FIHb was characterized with thermofluor, CD, SLS and SAXS. The structural studies gave the possibility to compare the two observed variants of FIHb. From the structural studies it seems like the heme-less variant has a more open structural shape and behaves as a dimer compared to the heme-containing active variant, which is monomeric and has a closed, globular shape.

To obtain more information about its function, it would be desirable to form crystals with high resolution. No relevant protein crystals were obtained in this study, and new trials to obtain crystals need to be done in order to solve the structure of FIHb.

6.2 Nitroreductases

NR1, are together with NR2, thought to be involved in the MTZ resistance mechanism of *Giardia*. NR1 was successfully cloned into competent *E.coli* cells using pETNKI-his3C-LIC-amp as a vector and the method SLIC. The protein was expressed using Rosetta (DE3) as host cells in TB media, and purified. Little new information about NR1 was obtained in this study. NR2 was not straightforward to clone into the NKI system. Further research needs to be done for both of the enzymes, to obtain more information about their functions.

7. FURTHER RESEARCH

To be able to understand the MTZ resistance mechanisms of *Giardia*, proteins in the pathways that are thought to play a role in this mechanism needs further investigations. Characterization of FIHb, NR1 and NR2, which are important in these pathways, was started. By characterizing and solving the structure of the proteins, the structure-function relations of the proteins can be discovered. From the structure, more of the proteins functions could be obtained. This project resulted in no diffracting protein crystals, but initial expression and purifications were discovered for the FIHb and NR1, which may form the basis for further research. The expression and purification for assemblage B FIHb and NR1 need to be further optimized to obtain more soluble proteins and increasing the yields. In addition, some structural characterization was done for assemblage A FIHb. These results can be used in comparison when characterizing the FIHb from assemblage B. Further research to obtain the crystal structure of FIHb is needed. To obtain crystals from the active variant (heme-containing), a way of adding heme to the protein solutions should be investigated to get heme to stay inside the protein. Additionally heme assays should be performed to quantify the heme concentration of FIHb and assess the ratio between concentrations of FIHb and heme. NADH oxidase activity could be performed to measure the enzyme activity of the FIHb after it can be stabilized to its active form.

No structural characterization was yet done for NR1 nor NR2 due to the time limit. This work should be continued in further studies.

REFERENCES

1. Adam RD. The *Giardia lamblia* genome. *Int J Parasitol.* 2000;30(4):475-84.
2. Adam RD. Biology of *Giardia lamblia*. *Clinical microbiology reviews.* 2001;14(3):447-75.
3. Miotti Pg Fau - Gilman RH, Gilman Rh Fau - Santosham M, Santosham M Fau - Ryder RW, Ryder Rw Fau - Yolken RH, Yolken RH. Age-related rate of seropositivity of antibody to *Giardia lamblia* in four diverse populations. (0095-1137 (Print)).
4. Craun GF. Waterborne giardiasis in the United States: a review. (0090-0036 (Print)).
5. McCormick BJ. Frequent symptomatic or asymptomatic infections may have long-term consequences on growth and cognitive development2014.
6. Nataro JP. Diarrhea among children in developing countries. *Advances in experimental medicine and biology.* 2013;764:73-80.
7. Rogawski ET, Bartelt LA, Platts-Mills JA, Seidman JC, Samie A, Havt A, et al. Determinants and Impact of *Giardia* Infection in the First 2 Years of Life in the MAL-ED Birth Cohort. *Journal of the Pediatric Infectious Diseases Society.* 2017;6(2):153-60.
8. Cacciò SM, Sprong H. Epidemiology of Giardiasis in Humans. In: Luján HD, Svärd S, editors. *Giardia: A Model Organism.* Vienna: Springer Vienna; 2011. p. 17-28.
9. Nygard K, Schimmer B, Sobstad O, Walde A, Tveit I, Langeland N, et al. A large community outbreak of waterborne giardiasis-delayed detection in a non-endemic urban area. *BMC public health.* 2006;6:141.
10. Feng Y, Xiao L. Zoonotic potential and molecular epidemiology of *Giardia* species and giardiasis. *Clinical microbiology reviews.* 2011;24(1):110-40.
11. Plutzer J, Ongerth J, Karanis P. *Giardia* taxonomy, phylogeny and epidemiology: Facts and open questions. *International journal of hygiene and environmental health.* 2010;213(5):321-33.
12. Robertson LJ, Hermansen L, Gjerde BK, Strand E, Alvsvag JO, Langeland N. Application of genotyping during an extensive outbreak of waterborne giardiasis in Bergen, Norway, during autumn and winter 2004. *Applied and Environmental Microbiology.* 2006;72(3):2212-7.
13. Morch K, Hanevik K, Robertson LJ, Strand EA, Langeland N. Treatment-ladder and genetic characterisation of parasites in refractory giardiasis after an outbreak in Norway. *The Journal of infection.* 2008;56(4):268-73.
14. Litleskare S, Rortveit G, Eide GE, Hanevik K, Langeland N, Wensaas KA. Prevalence of Irritable Bowel Syndrome and Chronic Fatigue 10 Years After *Giardia* Infection. *Clinical gastroenterology and hepatology : the official clinical practice journal of the American Gastroenterological Association.* 2018.
15. Litleskare S, Wensaas K-A, Eide GE, Hanevik K, Kahrs GE, Langeland N, et al. Perceived food intolerance and irritable bowel syndrome in a population 3 years after a giardiasis-outbreak: a historical cohort study. *BMC Gastroenterology.* 2015;15(1):164.
16. Lipoldova M. *Giardia* and *Vilem Dusan Lambl*. *PLoS neglected tropical diseases.* 2014;8(5):e2686.
17. Watkins RR, Eckmann L. Treatment of giardiasis: current status and future directions. *Current infectious disease reports.* 2014;16(2):396.

18. Ankarklev J, Jerlstrom-Hultqvist J, Ringqvist E, Troell K, Svard SG. Behind the smile: cell biology and disease mechanisms of *Giardia* species. *Nature reviews Microbiology*. 2010;8(6):413-22.
19. Tovar J, Leon-Avila G, Sanchez LB, Sutak R, Tachezy J, van der Giezen M, et al. Mitochondrial remnant organelles of *Giardia* function in iron-sulphur protein maturation. *Nature*. 2003;426(6963):172-6.
20. Ali V, Nozaki T. Iron-sulphur clusters, their biosynthesis, and biological functions in protozoan parasites. *Advances in parasitology*. 2013;83:1-92.
21. Mastronicola D, Falabella M, Forte E, Testa F, Sarti P, Giuffre A. Antioxidant defence systems in the protozoan pathogen *Giardia intestinalis*. *Molecular and biochemical parasitology*. 2016;206(1-2):56-66.
22. Heyworth MF. *Giardia duodenalis* genetic assemblages and hosts. *Parasite (Paris, France)*. 2016;23:13.
23. Paget TA, Macechko PT, Jarroll EL. Metabolic changes in *Giardia intestinalis* during differentiation. *The Journal of parasitology*. 1998;84(2):222-6.
24. Lauwaet T, Davids BJ, Reiner DS, Gillin FD. Encystation of *Giardia lamblia*: a model for other parasites. *Current opinion in microbiology*. 2007;10(6):554-9.
25. Esch K, Petersen C. Transmission and Epidemiology of Zoonotic Protozoal Diseases of Companion Animals 2013. 58-85 p.
26. Rendtorff RC. The experimental transmission of human intestinal protozoan parasites. II. *Giardia lamblia* cysts given in capsules. *American journal of hygiene*. 1954;59(2):209-20.
27. Gardner TB, Hill DR. Treatment of giardiasis. *Clinical microbiology reviews*. 2001;14(1):114-28.
28. Chester AC, MacMurray FG, Restifo MD, Mann O. Giardiasis as a chronic disease. *Digestive diseases and sciences*. 1985;30(3):215-8.
29. Faubert G. Immune response to *Giardia duodenalis*. *Clinical microbiology reviews*. 2000;13(1):35-54, table of contents.
30. Kotloff KL, Nataro JP, Blackwelder WC, Nasrin D, Farag TH, Panchalingam S, et al. Burden and aetiology of diarrhoeal disease in infants and young children in developing countries (the Global Enteric Multicenter Study, GEMS): a prospective, case-control study. *Lancet (London, England)*. 2013;382(9888):209-22.
31. Wolfe MS. Giardiasis. *Clinical microbiology reviews*. 1992;5(1):93-100.
32. Karin Leder M, FRACP, PhD, MPH, DTMHPeter F Weller, MD, MACP. Giardiasis: Epidemiology, clinical manifestations, and diagnosis: UpToDate; 2016 [updated Sep 27, 2016]. Available from: <https://www.uptodate.com/contents/giardiasis-epidemiology-clinical-manifestations-and-diagnosis>.
33. Verweij JJ, Blange RA, Templeton K, Schinkel J, Brienen EA, van Rooyen MA, et al. Simultaneous detection of *Entamoeba histolytica*, *Giardia lamblia*, and *Cryptosporidium parvum* in fecal samples by using multiplex real-time PCR. *Journal of clinical microbiology*. 2004;42(3):1220-3.
34. Organization WH. WHO Model List of Essential Medicines: WHO; 2017 [updated March 2017. 20th List:[Available from: <http://www.who.int/medicines/publications/essentialmedicines/en/>.
35. Leitsch D. Drug Resistance in the Microaerophilic Parasite *Giardia lamblia*. *Current tropical medicine reports*. 2015;2(3):128-35.
36. Brown DM, Upcroft JA, Edwards MR, Upcroft P. Anaerobic bacterial metabolism in the ancient eukaryote *Giardia duodenalis*. *Int J Parasitol*. 1998;28(1):149-64.
37. Ansell BR, McConville MJ, Ma'ayeh SY, Dagley MJ, Gasser RB, Svard SG, et al. Drug resistance in *Giardia duodenalis*. *Biotechnology advances*. 2015;33(6 Pt 1):888-901.

38. Müller J, Rout S, Leitsch D, Vaithilingam J, Hehl A, Müller N. Comparative characterisation of two nitroreductases from *Giardia lamblia* as potential activators of nitro compounds. *International Journal for Parasitology*. 2015;5(2):37-43.
39. Leitsch D. A review on metronidazole: an old warhorse in antimicrobial chemotherapy. *Parasitology*. 2017:1-12.
40. Raj D, Ghosh E, Mukherjee AK, Nozaki T, Ganguly S. Differential gene expression in *Giardia lamblia* under oxidative stress: significance in eukaryotic evolution. *Gene*. 2014;535(2):131-9.
41. Lu J, Holmgren A. The thioredoxin antioxidant system. *Free radical biology & medicine*. 2014;66:75-87.
42. Nabarro LE, Lever RA, Armstrong M, Chiodini PL. Increased incidence of nitroimidazole-refractory giardiasis at the Hospital for Tropical Diseases, London: 2008-2013. *Clinical microbiology and infection : the official publication of the European Society of Clinical Microbiology and Infectious Diseases*. 2015;21(8):791-6.
43. Nillius D, Muller J, Muller N. Nitroreductase (GINR1) increases susceptibility of *Giardia lamblia* and *Escherichia coli* to nitro drugs. *The Journal of antimicrobial chemotherapy*. 2011;66(5):1029-35.
44. Townson SM, Upcroft JA, Upcroft P. Characterisation and purification of pyruvate:ferredoxin oxidoreductase from *Giardia duodenalis*. *Molecular and biochemical parasitology*. 1996;79(2):183-93.
45. Townson SM, Hanson GR, Upcroft JA, Upcroft P. A purified ferredoxin from *Giardia duodenalis*. *European journal of biochemistry*. 1994;220(2):439-46.
46. Muller J, Wastling J, Sanderson S, Muller N, Hemphill A. A novel *Giardia lamblia* nitroreductase, GINR1, interacts with nitazoxanide and other thiazolides. *Antimicrobial agents and chemotherapy*. 2007;51(6):1979-86.
47. Muller J, Schildknecht P, Muller N. Metabolism of nitro drugs metronidazole and nitazoxanide in *Giardia lamblia*: characterization of a novel nitroreductase (GINR2). *The Journal of antimicrobial chemotherapy*. 2013;68(8):1781-9.
48. Muller J, Ley S, Felger I, Hemphill A, Muller N. Identification of differentially expressed genes in a *Giardia lamblia* WB C6 clone resistant to nitazoxanide and metronidazole. *The Journal of antimicrobial chemotherapy*. 2008;62(1):72-82.
49. Muller J, Sterk M, Hemphill A, Muller N. Characterization of *Giardia lamblia* WB C6 clones resistant to nitazoxanide and to metronidazole. *The Journal of antimicrobial chemotherapy*. 2007;60(2):280-7.
50. Upcroft P, Upcroft JA. Drug targets and mechanisms of resistance in the anaerobic protozoa. *Clinical microbiology reviews*. 2001;14(1):150-64.
51. Fabrizio Testa AG, Daniela Mastronicola, Elena Forte and Paolo Sarti. Enzymatic detoxification of O₂ and NO in the human parasite, *Giardia intestinalis*: A mini review. *Indian Journal of Biotechnology*. 2011;10:404-9.
52. Ma'ayeh SY, Knörr L, Svärd SG. Transcriptional profiling of *Giardia intestinalis* in response to oxidative stress. *International Journal for Parasitology*. 2015;45(14):925-38.
53. Andersson JO, Sjögren ÅM, Davis LAM, Embley TM, Roger AJ. Phylogenetic Analyses of Diplomonad Genes Reveal Frequent Lateral Gene Transfers Affecting Eukaryotes. *Current Biology*. 2003;13(2):94-104.
54. Gardner PR, Gardner AM, Martin LA, Salzman AL. Nitric oxide dioxygenase: an enzymic function for flavohemoglobin. *Proc Natl Acad Sci U S A*. 1998;95(18):10378-83.
55. Berg JM TJ, Stryer L. Protein Structure and Function. 2002. In: *Biochemistry* [Internet]. New York: W H Freeman. 5th edition. Available from: <https://www.ncbi.nlm.nih.gov/books/NBK21177/>.
56. Lesk AM. Introduction to protein science 3. ed. UK: Oxford university press; 2016.

57. Breda A VN, Norberto de Souza O, et al. Protein Structure, Modelling and Applications. 2006. In: Bioinformatics in Tropical Disease [Internet]. National Center for Biotechnology Information. Available from: <https://www.ncbi.nlm.nih.gov/books/NBK6824/>.
58. Tanner NK. Ribozymes: the characteristics and properties of catalytic RNAs. FEMS microbiology reviews. 1999;23(3):257-75.
59. Jeremy M. Berg JLT, Lubert Stryer Biochemistry. Biochemistry. 7th ed. New York: W. H Freeman and Company; 2012.
60. Schmitt E, Tanrikulu IC, Yoo TH, Panvert M, Tirrell DA, Mechulam Y. Switching from an induced-fit to a lock-and-key mechanism in an aminoacyl-tRNA synthetase with modified specificity. J Mol Biol. 2009;394(5):843-51.
61. Ilari A, Boffi A. Structural studies on flavohemoglobins. Methods in enzymology. 2008;436:187-202.
62. Bonamore A, Boffi A. Flavohemoglobin: structure and reactivity. IUBMB life. 2008;60(1):19-28.
63. Ilari A, Bonamore A Fau - Farina A, Farina A Fau - Johnson KA, Johnson Ka Fau - Boffi A, Boffi A. The X-ray structure of ferric Escherichia coli flavohemoglobin reveals an unexpected geometry of the distal heme pocket. (0021-9258 (Print)).
64. Ermler U, Siddiqui RA, Cramm R, Schroder D, Friedrich B. Crystallization and preliminary X-ray diffraction studies of a bacterial flavohemoglobin protein. Proteins. 1995;21(4):351-3.
65. Goodwin A, Kersulyte D, Sisson G, Veldhuyzen van Zanten SJ, Berg DE, Hoffman PS. Metronidazole resistance in Helicobacter pylori is due to null mutations in a gene (rdxA) that encodes an oxygen-insensitive NADPH nitroreductase. Molecular microbiology. 1998;28(2):383-93.
66. Guillen H, Curiel JA, Landete JM, Munoz R, Herraiz T. Characterization of a nitroreductase with selective nitroreduction properties in the food and intestinal lactic acid bacterium Lactobacillus plantarum WCFS1. Journal of agricultural and food chemistry. 2009;57(21):10457-65.
67. Tavares AF, Nobre LS, Melo AM, Saraiva LM. A novel nitroreductase of Staphylococcus aureus with S-nitrosoglutathione reductase activity. Journal of bacteriology. 2009;191(10):3403-6.
68. Leicester Uo. Recombinant DNA and genetic techniques Leicester: University of Leicester; [updated 24. april 2018. Available from: <https://www2.le.ac.uk/projects/vgec/schoolsandcolleges/topics/recombinanttechniques>.
69. EMBL. Cloning methods - cloning using restriction enzymes: EMBL; [24. april 2018]. Available from: https://www.embl.de/pepcore/pepcore_services/cloning/cloning_methods/restriction_enzymes/index.html.
70. Dmowski M, Jagura-Burdzy G. Active stable maintenance functions in low copy-number plasmids of Gram-positive bacteria I. Partition systems. Polish journal of microbiology. 2013;62(1):3-16.
71. Baneyx F. Recombinant protein expression in Escherichia coli. Current Opinion in Biotechnology. 1999;10(5):411-21.
72. Preston A. Choosing a cloning vector. Methods in molecular biology (Clifton, NJ). 2003;235:19-26.
73. Sigma-Aldrich. Protein Expression Systems: Sigma-Aldrich; 2005 [Available from: <https://www.sigmaaldrich.com/technical-documents/articles/biology/protein-expression-systems.html>].

74. Rosano GL, Ceccarelli EA. Recombinant protein expression in *Escherichia coli*: advances and challenges. *Frontiers in microbiology*. 2014;5:172.
75. Bujard H, Gentz R, Lanzer M, Stueber D, Mueller M, Ibrahim I, et al. A T5 promoter-based transcription-translation system for the analysis of proteins in vitro and in vivo. *Methods in enzymology*. 1987;155:416-33.
76. Novagen. pET System Manual USA & Canada: Novagen; 1999 [cited 2018 May]. Available from: <https://research.fhcrc.org/content/dam/stripe/hahn/methods/biochem/pet.pdf>.
77. Lodish H BA, Zipursky SL, et al. *Molecular Cell Biology*. 2000 [cited 30. april 2018]. In: *Molecular Cell Biology* [Internet]. New York: W. H. Freeman. 4th edition. [cited 30. april 2018]. Available from: <https://www.ncbi.nlm.nih.gov/books/NBK21498/>.
78. Green R, Rogers EJ. Transformation of chemically competent *E. coli*. *Methods in enzymology*. 2013;529:329-36.
79. Deutzmann R. Structural characterization of proteins and peptides. *Methods in molecular medicine*. 2004;94:269-97.
80. Greenfield NJ. Using circular dichroism spectra to estimate protein secondary structure. *Nature protocols*. 2006;1(6):2876-90.
81. Kelly SM, Jess TJ, Price NC. How to study proteins by circular dichroism. *Biochimica et Biophysica Acta (BBA) - Proteins and Proteomics*. 2005;1751(2):119-39.
82. Geerlof MGA. Protein characterization by static light scattering Munich, Germany: Helmholtz Zentrum München; 2006 [cited 2018 May 2018]. Available from: <https://www.helmholtz-muenchen.de/fileadmin/PEPF/PDF/SLS-protocol.pdf>.
83. Worldwide MI. Static Light Scattering technologies for GPC - SEC explained Worcestershire, UK: Malvern Instruments Worldwide; 2015 [Available from: <https://www.chem.uci.edu/~dmitryf/manuals/Fundamentals/SLS Technologies GPC-SEC Explained.pdf>].
84. Nettleship JE, Brown J, Groves MR, Geerlof A. Methods for protein characterization by mass spectrometry, thermal shift (ThermoFluor) assay, and multiangle or static light scattering. *Methods in molecular biology (Clifton, NJ)*. 2008;426:299-318.
85. Ahner K, Buchacher A, Iberer G, Josic D, Jungbauer A. Analysis of aggregates of human immunoglobulin G using size-exclusion chromatography, static and dynamic light scattering. *Journal of chromatography A*. 2003;1009(1-2):89-96.
86. Hong P, Koza S, Bouvier ESP. Size-Exclusion Chromatography for the Analysis of Protein Biotherapeutics and their Aggregates. *Journal of Liquid Chromatography & Related Technologies*. 2012;35(20):2923-50.
87. Reinhard L, Mayerhofer H, Geerlof A, Mueller-Dieckmann J, Weiss MS. Optimization of protein buffer cocktails using ThermoFluor. *Acta Crystallographica Section F: Structural Biology and Crystallization Communications*. 2013;69(Pt 2):209-14.
88. Allec N, Choi M, Yesupriya N, Szychowski B, White MR, Kann MG, et al. Small-angle X-ray scattering method to characterize molecular interactions: Proof of concept. *Scientific Reports*. 2015;5.
89. Kikhney AG, Svergun DI. A practical guide to small angle X-ray scattering (SAXS) of flexible and intrinsically disordered proteins. *FEBS letters*. 2015;589(19, Part A):2570-7.
90. Biosaxs. Small angle X-ray scattering Hamburg, Germany: Biosaxs; [cited 2018 May 2018]. Available from: <http://biosaxs.com/technique.html>.
91. Franke D, Kikhney AG, Svergun DI. Automated acquisition and analysis of small angle X-ray scattering data. *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment*. 2012;689:52-9.
92. McPherson A. A brief history of protein crystal growth. *Journal of Crystal Growth*. 1991;110(1):1-10.

93. McPherson A, Gavira JA. Introduction to protein crystallization. *Acta Crystallographica Section F, Structural Biology Communications*. 2014;70(Pt 1):2-20.
94. Dessau MA, Modis Y. Protein Crystallization for X-ray Crystallography. *Journal of Visualized Experiments : JoVE*. 2011(47).
95. institute Nc. NKI Protein Facility LIC vectors. Netherland: Netherlands cancer institute
96. NKI. NKI - Crystallisation Construct Designer: West-Life Structures for life; 2017 [Available from: <https://ccd.rhpc.nki.nl>].
97. Michael R. Green JS. *Molecular cloning: a laboratory manual* 4th ed. Cold Spring Harbor, N.Y: CSH Press; 2012.
98. Yan Nie CB, Imre Berger. ACEMBL Expression System- User Manual. 2009. Grenoble, France: European Molecular Biology Laboratory. Vers. 09.11. [9-13]. Available from: <https://www.embl.fr/research/services/berger/ACEMBL.pdf>.
99. Scientific T. User Guide: FastDigest KpnI: ThermoFisher Scientific; 2012 [Available from: <https://www.thermofisher.com/order/catalog/product/FD0524>].
100. Jeong JY, Yim HS, Ryu JY, Lee HS, Lee JH, Seen DS, et al. One-Step Sequence- and Ligation-Independent Cloning as a Rapid and Versatile Cloning Method for Functional Genomics Studies. *Applied and Environmental Microbiology*. 2012;78(15):5440-3.
101. Michael R. Green JS. Terrific Broth. *Molecular Cloning - A laboratory manual*. 3. 4th ed. New York, USA: Cold Spring Harbor Laboratory Press; 2012. p. 1825.
102. Svergun DI, Petoukhov MV, Koch MH. Determination of domain structure of proteins from X-ray solution scattering. *Biophysical Journal*. 2001;80(6):2946-53.
103. P.V. Konarev AGK, A.V. Sokolova, D.I. Svergun1 & V.V. Volkov2, editor PRIMUS - a Windows-PC based system for small-angle scattering data analysis 2003: *Journal of Applied Crystallography*; 2003.
104. Bioinformatics SIO. ExpASy Translate tool Swiss Institute of Bioinformatics [cited 2018 March 2018]. Available from: <https://web.expasy.org/translate/>.
105. Bioinformatics SIO. ExpASy ProtParam tool: Swiss Institute of Bioinformatics; [cited 2018 March]. Available from: <https://web.expasy.org/protparam/>.
106. O'Sullivan D J, Klaenhammer TR. Rapid Mini-Prep Isolation of High-Quality Plasmid DNA from *Lactococcus* and *Lactobacillus* spp. *Applied and Environmental Microbiology*. 1993;59(8):2730-3.
107. Academy K. Polymerase chain reaction (PCR): Khan Academy; [cited 2018 23. january 2018]. Available from: <https://nb.khanacademy.org/science/biology/biotech-dna-technology/dna-sequencing-pcr-electrophoresis/a/polymerase-chain-reaction-pcr>.
108. Brown TA. *Gene cloning and DNA analysis - An introduction*. Manchester: Wiley-Blackwell; 2010 [cited 2018 25. january 2018]. Available from: [https://eclass.upatras.gr/modules/document/file.php/BIO276/Gene Cloning %26 DNA Analysis.pdf](https://eclass.upatras.gr/modules/document/file.php/BIO276/Gene%20Cloning%26%20DNA%20Analysis.pdf).
109. Studier FW. Protein production by auto-induction in high density shaking cultures. *Protein expression and purification*. 2005;41(1):207-34.
110. Laboratories B-R. *A Guide to Polyacrylamide Gel Electrophoresis and Detection* [cited 2017 14. September]. Available from: http://www.bio-rad.com/webroot/web/pdf/lsr/literature/Bulletin_6040.pdf.
111. Harris DC. *Quantitative Chemical Analysis*. Quantitative Chemical Analysis. 8th ed. New York: W. H. Freeman and Company; 2010. p. 502-29.
112. Duong-Ly KC, Gabelli SB. Chapter Nine - Gel Filtration Chromatography (Size Exclusion Chromatography) of Proteins. In: Lorsch J, editor. *Methods in enzymology*. 541: Academic Press; 2014. p. 105-14.

113. Waugh DS. An Overview of Enzymatic Reagents for the Removal of Affinity Tags. *Protein expression and purification*. 2011;80(2):283-93.
114. Saraswathy N, Ramalingam P. 13 - Protein Identification by Peptide Mass Fingerprinting (PMF). *Concepts and Techniques in Genomics and Proteomics*: Woodhead Publishing; 2011. p. 185-92.
115. Balagurumoorthy P, Adelstein SJ, Kassis AI. Method to eliminate linear DNA from mixture containing nicked circular, supercoiled, and linear plasmid DNA. *Anal Biochem*. 2008;381(1):172-4.
116. Aaij C, Borst P. The gel electrophoresis of DNA. *Biochimica et Biophysica Acta (BBA) - Nucleic Acids and Protein Synthesis*. 1972;269(2):192-200.
117. Morrison HG, McArthur AG, Gillin FD, Aley SB, Adam RD, Olsen GJ, et al. Genomic minimalism in the early diverging intestinal parasite *Giardia lamblia*. *Science (New York, NY)*. 2007;317(5846):1921-6.
118. Rafferty S, Luu B, March RE, Yee J. *Giardia lamblia* encodes a functional flavohemoglobin. *Biochem Biophys Res Commun*. 2010;399(3):347-51.
119. Mastronicola D, Testa F, Forte E, Bordi E, Pucillo LP, Sarti P, et al. Flavohemoglobin and nitric oxide detoxification in the human protozoan parasite *Giardia intestinalis*. *Biochemical and Biophysical Research Communications*. 2010;399(4):654-8.
120. Galperin MY. Conserved 'hypothetical' proteins: new hints and new puzzles. *Comparative and functional genomics*. 2001;2(1):14-8.
121. Brogi S, Fiorillo A, Chemi G, Butini S, Lalle M, Ilari A, et al. Structural characterization of *Giardia duodenalis* thioredoxin reductase (gTrxR) and computational analysis of its interaction with NBDHEX. *European Journal of Medicinal Chemistry*. 2017;135:479-90.
122. Castillo-Villanueva A, Mendez ST, Torres-Arroyo A, Reyes-Vivas H, Oria-Hernandez J. Cloning, Expression and Characterization of Recombinant, NADH Oxidase from *Giardia lamblia*. *The protein journal*. 2016;35(1):24-33.
123. Di Matteo A, Scandurra FM, Testa F, Forte E, Sarti P, Brunori M, et al. The O₂-scavenging flavodiiron protein in the human parasite *Giardia intestinalis*. *J Biol Chem*. 2008;283(7):4061-8.
124. Leitsch D, Burgess AG, Dunn LA, Krauer KG, Tan K, Duchene M, et al. Pyruvate:ferredoxin oxidoreductase and thioredoxin reductase are involved in 5-nitroimidazole activation while flavin metabolism is linked to 5-nitroimidazole resistance in *Giardia lamblia*. *The Journal of antimicrobial chemotherapy*. 2011;66(8):1756-65.
125. Londer YY, Pokkuluri PR, Tiede DM, Schiffer M. Production and preliminary characterization of a recombinant triheme cytochrome c7 from *Geobacter sulfurreducens* in *Escherichia coli*. *Biochimica et Biophysica Acta (BBA) - Bioenergetics*. 2002;1554(3):202-11.
126. Freigassner M, Pichler H, Glieder A. Tuning microbial hosts for membrane protein production. *Microbial cell factories*. 2009;8:69.
127. Sivashanmugam A, Murray V, Cui C, Zhang Y, Wang J, Li Q. Practical protocols for production of very high yields of recombinant proteins using *Escherichia coli*. *Protein science : a publication of the Protein Society*. 2009;18(5):936-48.
128. Cunningham F, Deber CM. Optimizing synthesis and expression of transmembrane peptides and proteins. *Methods (San Diego, Calif)*. 2007;41(4):370-80.
129. Frey AD, Farres J, Bollinger CJ, Kallio PT. Bacterial hemoglobins and flavohemoglobins for alleviation of nitrosative stress in *Escherichia coli*. *Applied and Environmental Microbiology*. 2002;68(10):4835-40.
130. Wu B. Purification and characterization of Flavohemoglobin- A flavoheme enzyme [Master thesis]. Oslo: UNIVERSITY OF OSLO; 2015.

131. Kobayashi G, Nakamura T, Ohmachi H, Matsuoka A, Ochiai T, Shikama K. Yeast flavohemoglobin from *Candida norvegensis*. Its structural, spectral, and stability properties. *J Biol Chem*. 2002;277(45):42540-8.
132. Fathi-Roudsari M, Akhavian-Tehrani A, Maghsoudi N. Comparison of Three *Escherichia coli* Strains in Recombinant Production of Reteplase. *Avicenna journal of medical biotechnology*. 2016;8(1):16-22.
133. Salton MR. The properties of lysozyme and its action on microorganisms. *Bacteriological reviews*. 1957;21(2):82-100.
134. Wilcox FH, Daniel LJ. Reduced lysis at high concentrations of lysozyme. *Archives of Biochemistry and Biophysics*. 1954;52(2):305-12.
135. Brown RB, Audet J. Current techniques for single-cell lysis. *Journal of the Royal Society, Interface*. 2008;5 Suppl 2:S131-8.
136. Stathopoulos PB, Scholz GA, Hwang YM, Rumfeldt JA, Lepock JR, Meiering EM. Sonication of proteins causes formation of aggregates that resemble amyloid. *Protein science : a publication of the Protein Society*. 2004;13(11):3017-27.
137. Cheung RC, Wong JH, Ng TB. Immobilized metal ion affinity chromatography: a review on its applications. *Appl Microbiol Biotechnol*. 2012;96(6):1411-20.
138. Block H, Maertens B, Spriestersbach A, Brinker N, Kubicek J, Fabis R, et al. Immobilized-metal affinity chromatography (IMAC): a review. *Methods in enzymology*. 2009;463:439-73.
139. Graslund S, Nordlund P, Weigelt J, Hallberg BM, Bray J, Gileadi O, et al. Protein production and purification. *Nature methods*. 2008;5(2):135-46.
140. Adhikari S, Manthena PV, Sajwan K, Kota KK, Roy R. A unified method for purification of basic proteins. *Anal Biochem*. 2010;400(2):203-6.
141. Widmann M, Trodler P, Pleiss J. The isoelectric region of proteins: a systematic analysis. *PloS one*. 2010;5(5):e10546.
142. Scientific T. Overview of Affinity Purification: ThermoFisher Scientific; [cited 2018 May]. Available from: <http://www.thermofisher.com/no/en/home/life-science/protein-biology/protein-biology-learning-center/protein-biology-resource-library/pierce-protein-methods/overview-affinity-purification.html>.
143. EMBL. PROTEIN PURIFICATION - STORAGE OF PURIFIED PROTEINS. In: Protein Expression and Purification Core Facility [Internet]. EMBL.
144. Bhatnagar BS, Bogner RH, Pikal MJ. Protein stability during freezing: separation of stresses and mechanisms of protein stabilization. *Pharmaceutical development and technology*. 2007;12(5):505-23.
145. Biotechnology P. Protein stability and storage USA: Pierce Biotechnology; 2003 [Available from: http://wolfson.huji.ac.il/purification/PDF/StorageProteins/PIERCE_ProteinStorage.pdf.
146. Parks TD, Leuther KK, Howard ED, Johnston SA, Dougherty WG. Release of proteins and peptides from fusion proteins using a recombinant plant virus proteinase. *Anal Biochem*. 1994;216(2):413-7.
147. Jenny RJ, Mann KG, Lundblad RL. A critical review of the methods for cleavage of fusion proteins with thrombin and factor Xa. *Protein expression and purification*. 2003;31(1):1-11.
148. Bucher MH, Evdokimov AG, Waugh DS. Differential effects of short affinity tags on the crystallization of *Pyrococcus furiosus* maltodextrin-binding protein. *Acta crystallographica Section D, Biological crystallography*. 2002;58(Pt 3):392-7.
149. Gardner PR. Chapter Twelve - Assay and Characterization of the NO Dioxygenase Activity of Flavohemoglobins. In: Poole RK, editor. *Methods in enzymology*. 436: Academic Press; 2008. p. 217-37.

150. Sigma-Aldrich. HEMIN ProductInformation: Sigma-Aldrich; 1999 [cited 2018 20.05.2018]. Available from: https://www.sigmaaldrich.com/content/dam/sigma-aldrich/docs/Sigma/Product_Information_Sheet/2/h2250pis.pdf.
151. Gardner AM, Gardner PR. Flavohemoglobin detoxifies nitric oxide in aerobic, but not anaerobic, Escherichia coli. Evidence for a novel inducible anaerobic nitric oxide-scavenging activity. *J Biol Chem.* 2002;277(10):8166-71.
152. Luft JR, Wolfley JR, Snell EH. What's in a drop? Correlating observations and outcomes to guide macromolecular crystallization experiments. *Crystal growth & design.* 2011;11(3):651-63.

APPENDIX 1

Lysis buffer 1

150 mM NaCl
20 mM HEPES pH 7.5
0.5 mM TCEP
10 mM Imidazole

Lysis buffer 2

400 mM NaCl
20 mM C₂H₇NO₂, pH 5
0.5 mM TCEP
10 mM Imidazole
5 % glycerol

Lysis buffer 3

400 mM NaCl
20 mM Tris, pH 8.5
0.5 mM TCEP
10 mM Imidazole
5 % glycerol

Wash buffer 1

500 mM NaCl
20 mM HEPES, pH 7.5
0.5 mM TCEP
30 mM Imidazole

Wash buffer 2

800 mM NaCl
20 mM C₂H₇NO₂, pH 5
0.5 mM TCEP
30 mM Imidazole
5 % glycerol

Wash buffer 3

800 mM NaCl
20 mM Tris, pH 8.5
0.5 mM TCEP
30 mM Imidazole
5 % glycerol

Elution buffer 1

150 mM NaCl
20 mM HEPES, pH 7.5
0.5 mM TCEP
400 mM Imidazole

Elution buffer 2

400 mM NaCl
20 mM C₂H₇NO₂, pH 5
0.5 mM TCEP
400 mM Imidazole
5 % glycerol

Elution buffer 3

400 mM NaCl
20 mM Tris, pH 8.5
0.5 mM TCEP
400 mM Imidazole
5 % glycerol

Gel filtration buffer 1

150 mM NaCl
20 mM HEPES, pH 8
0.5 mM TCEP
5 % glycerol

Gel filtration buffer 2

50 mM NaCl
25 mM phosphate buffer,
pH 7.5

Gel filtration buffer 3

150 mM NaCl
20 mM phosphate buffer,
pH 6.5
0.5 mM TCEP
5 % glycerol

Dialysis buffer 1

150 mM NaCl pH 8.5
20 mM HEPES
1 mM DDT

Dialysis buffer 2

20 mM NaF
20 mM Na/K phosphate

PBS buffer

1 tablet of phosphate
buffered saline dissolved
in 200 mL dH₂O, pH 7.4

TAE Buffer 1 X

Made by Ju Xu

SDS running buffer 1 X

Made by Ju Xu