Purification and characterization of Flavohemoglobin

A flavoheme enzyme

Bernt Wu

Master thesis at the Department of Biosciences]

UNIVERSITETY OF OSLO

June 2015

Abstract

Flavohemoglobin (FHb) is a flavoheme enzyme that catalyzes dioxygenation of nitric oxide (NO) to nitrate NO_3^-. NO is a diatomic radical gas produced by enzymatic and non-enzymatic oxidation of reduced nitrogenous compounds. NO acts as a membrane-permeable signal molecule in mammals. In addition it is toxic at higher concentrations and is utilized by the immune system against invading pathogens. FHb is involved in protection against the cytotoxic effects of NO (termed nitrosative stress) and is found in a range of bacterial species and some fungi. FHbs are not found in higher eukaryotes, but is employed by a range of pathogens, and is a potential therapeutic target.

In this thesis we sought to investigate the FHb from *Bacillus cereus* by determination of the X-ray crystal structure as only a few structures of FHb from different organisms have been obtained.

The protein was successfully cloned and expressed and a purification protocol has been developed, with ammonium sulfate precipitation, anion exchange chromatography and gel filtration. A high yield with high purity was obtained, but with heme and FAD at substoichiometric levels. Crystallization screening resulted in some small needles grown from precipitate, although they were non-reproducible and no crystal structure was obtained.
A heme assay was performed in order to identify the heme/protein ratio and protein assays were performed to calculate protein concentration. Testing of reconstitution with heme and FAD was performed with promising results.

Acknowledgements

The work presented in this thesis was performed in the laboratory of Professor K. Kristoffer Andersson at the Department of Biosciences, University of Oslo.

First of all I would like to thank my supervisor Dr. Hans-Petter Hersleth for making this master project possible, for being available with feedback, advice and support. Thanks to the group leader Professor K. Kristoffer Andersson for additional feedback and ideas for this thesis. I would also like thank the PhD students Marie Lofstad, Marta Hammerstad and Ingvild Gudim for advice in the lab, and my fellow master student Susanne Monka for trying, failing and occasionally succeeding, together,

Thanks to Christian Thon for taking a part of my workload so I could focus fully on the thesis and Christopher Fjeldstad for taking time during the holiday for proof-reading, Finally I would like to thank my family for support, and the dog for getting me out of the house to clear my head.

Contents

1 Introduction

This thesis will start with an introduction on flavoheme proteins, more specifically about the heme and flavin cofactors, followed by an introduction to nitric oxide. Thereafter flavohemoglobin (FHb), which is the focus of this thesis, will be introduced with some focus on structure, catalytic mechanism and an overview of the research field. At the end of the introduction a more specific description of the work carried out in this thesis will be given.

Heme and flavin cofactors

Biological cofactors are utilized in order for enzymes to enable a wide variety of reactions necessary for all aspects of life. In this section the characteristics of two cofactors will be presented, heme and flavin.

1.1.1 Heme

Heme-proteins are widely used throughout the biosphere, and the heme group is a central molecule of life [1, 2] that is involved in a range of reactions including: catalysis, electron transfer and immunological defense, in addition to oxygen transport and storage,

Heme is a prosthetic group consisting of an organic ring structure (protoporphyrin) and a single iron ion (figure 1.1). The four nitrogen atoms chelate the iron ion. Iron consists mainly in three oxidation states in biological systems: the reduced ferrous (Fe^{2+}) form with oxidation number +2, the ferric (Fe^{3+}) form with oxidation number +3 or the ferryl (Fe^{4+}) form with oxidation state +4 [3]. There are various types of heme, which differ in their protoporphyrin. The most abundant type, heme b, consists of a protoporphyrin IX. The other iron porphyrins differ in the substitutions pointing out from the ring system. Ferric (Fe^{3+}) protoporphyrin IX is called hemin and can be reduced by one electron to produce ferrous (Fe^{2+}) heme. Heme is usually used as a generic term for both the ferrous (Fe^{2+}) and ferric (Fe^{3+}) forms of iron protoporphyrin IX. Formerly however, heme only referred to the ferrous form and hemin to the ferric [1].

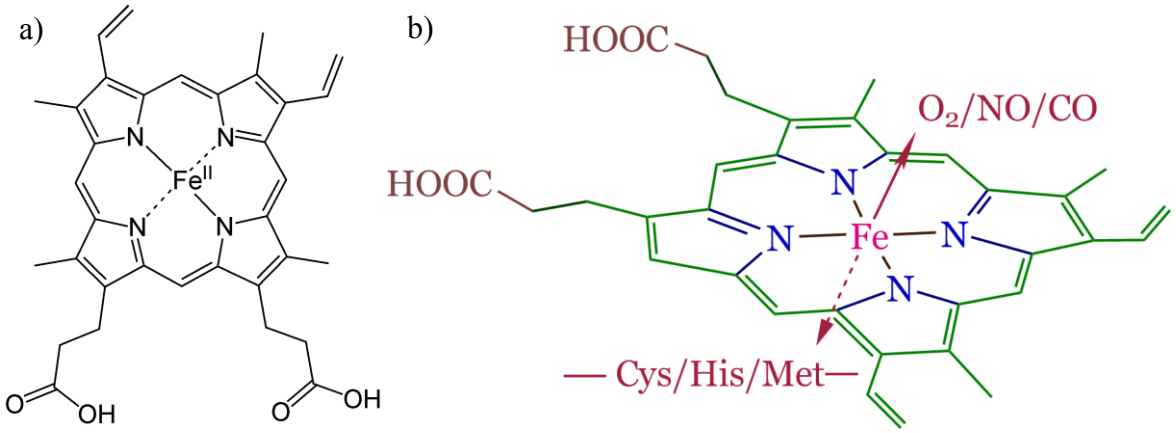


Figure 1.1: Heme cofactor: Heme consists of a protoporphyrin ring and a single iron ion. a) Heme b consists of protoporphyrin IX. b) The iron can coordinate two axial ligands, which may be Cys, His or Met residue in proteins or small molecules, including oxygen, nitric oxide and carbon monoxide. Taken from Zhang, Heme biology 2011 [1].

In proteins, the heme group is sequestrated in the protein structure in order to prevent irreversible conversion of the ferrous iron to the ferric state. The iron atom in heme can have up to six coordination bonds, four to the nitrogen atoms in the flat porphyrin ring and two perpendicular to the porphyrin. In heme proteins the iron may be 4, 5 or 6-coordinated with the axial ligands often being either cysteine, histidine, tyrosine or methionine and a small molecule [1].

1.1.2 Flavin

Flavoproteins are found in bacteria, archaea and eukaryotes, with the number of genes encoding flavin-dependent proteins varying from ~0.1% to 3.5% [4]. Flavoproteins catalyze a range of different reactions, and the flavin cofactors are said to compete for the title of master of versatility. Flavin-dependent reactions include dehydrogenation, oxidation, monooxygenation, halogenation, reduction of disulfide and other types of bonds, and light-sensing. Flavin-dependent enzymes include oxidoreductases (>90%), transferases (4.3%), lyases (2.8%) isomerases (1.4%) and ligases (0.4%) [4].

Flavins are a family of yellow-colored compounds with the basic structure of 7,8-dimethyl-10-alkylisoalloxazine (figure 1.2), which is a fused ring structure. Flavins are derived from riboflavin, commonly known as vitamin B2 [5]. Flavoproteins catalyze oxidation-reduction reactions. In most flavoproteins flavin is found in the form as flavin adenine dinucleotide

(FAD), or flavin mononucleotide (FMN), (figure 1.2). The isoalloxazine ring can undergo reversible reduction by accepting either one or two electrons, (figure 1.3).

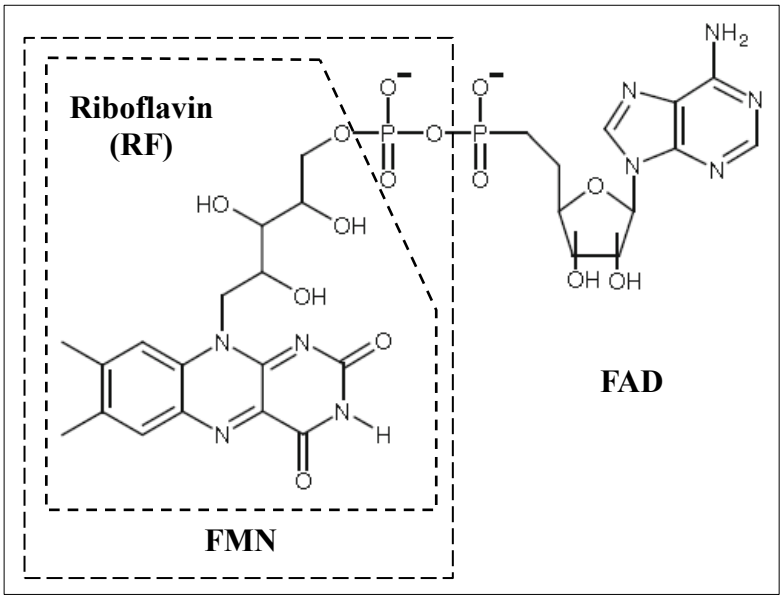


Figure 1.2: Structure of flavins: Structure of riboflavin (RF), flavin adenine dinucleotide (FAD) and flavin mononucleotide (FMN)

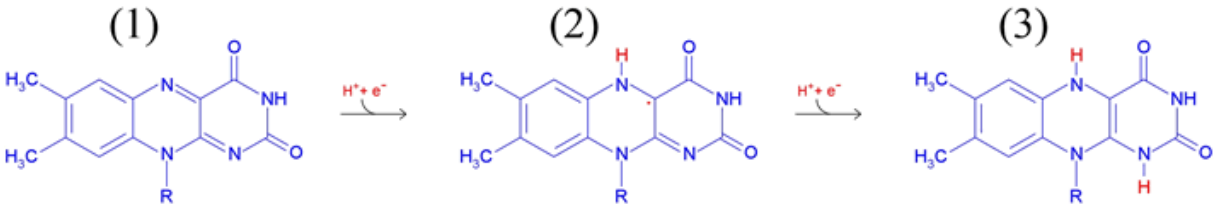


Figure 1.3: Flavin has three reduction states: Fully oxidized semiquionone form and fully reduced. Flavin can accept either one or two electrons and can be used in a range of reactions. The R-group differs depending on the type of flavin. Taken from Biochemistry dictionary, Gonzaga University [6].

The three redox states are: oxidized, one-electron reduced (semiquionone) and two-electron reduced. The three reduction states differ in color because of the conjugated bonds in the ring system. The fully oxidized form is yellow with absorption peaks at 445, 375, 265 and 220 nm. The semiquionone form can exist in either a neutral blue form or an anionic red form with a pKa of 8.5 [5]. The blue and red forms have absorption maxima at 500-600 nm and 370-400 nm respectively. The semiquionone form has a low stability in aqueous solution and will be protonated, followed by reduction to the fully reduced form. The fully reduced form is colorless with absorption maximum near 360 nm [7].

Flavin has the potential to transfer single electrons, hydrogen atoms and hydride ions making it more versatile than other redox cofactors. Other redox cofactors usually catalyze either one- or two-electron transfer processes exclusively. Flavoenzymes are therefore important mediators between one- and two-electron processes [5, 7].

In flavoproteins the flavin is tightly associated with the peptide, compared to NAD/NADP (nicotinamide adenine dinucleotide (phosphate)). The standard reduction potential of a flavin nucleotide depends on the protein with which it is associated and can be in the range from approximately -400 mV to +60 mV for the flavin. This is caused by local interactions with the protein. Similarly stabilization of the semiquinone form varies among flavoprotein [7].

Nitric Oxide and nitrosative stress

Nitric oxide (NO) or nitrogen monoxide is diatomic radical gas produced by enzymatic and non-enzymatic oxidation of reduced nitrogen compounds (figure 1.4). Properly, nitric oxide should be written as NO· to indicate the unpaired electron, but for the purposes of this thesis will be written as NO [8]. In order to stabilize the unpaired electron NO can react with other species containing an unpaired electron, or by interacting with some transition metals, in particular iron.

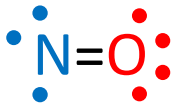


Figure 1.4: Nitric oxide: The structure of nitric oxide consists of a nitrogen atom and an oxygen atom with an unpaired electron.

NO acts as a signaling molecule in a range of processes in the human body including the nervous system and as a regulator of blood pressure and blood flow to different organs at nanomolar concentrations [8]. Today, other gases acting as signal molecules have been discovered (e.g. ethylene regulating ripening of fruit) but the discovery that such a structurally simple free radical could act as a signal molecule resulted in a major paradigm shift in the field of cell signaling [9]. In fact Furchgoff, Ignarro and Murad were awarded the Nobel Prize in Physiology or Medicine in 1998 for the discovery of nitric oxide as a signaling molecule [10]. In mammals nitric oxide is generated by nitric oxide synthases (NOS) and disruption of

NO signaling is linked to hypertension, neurodegeneration, stroke and heart disease [11]. In addition nitric oxide plays a role in cancer, both as a signal molecule and as a toxin [12].

At higher (micromolar) concentrations exogenous NO also demonstrates a broad range of cellular toxicities termed "nitrosative stress". Some of these toxicities are due to NO directly, but it seems that NO serves as a precursor for a range of reactive nitrogen species (RNS) such as dinitrogen trioxide (N_2O_3), nitrogen dioxide (NO_2^*) and peroxynitrite ($ONOO^-$).

The cytotoxic effects of NO are utilized by our immune system and so protection from NO is important for virulence of certain pathogens [13]. FHbs are found in various pathogens and protect against nitrosative stress by conversion of nitric oxide to nitrate (NO_3^-).

In addition FHb can be used to study NO biology by heterologous expression in plants [14, 15] and mammalian cells [16]. *E. coli* FHb have been shown to be active in mammalian cells. In addition FHb blocked the growth suppressing effects of exogenous and endogenous NO, and was able to blunt certain NO-dependent signaling pathways [16].

Flavohemoglobin (FHb)

1.1.3 Introduction

FHb is flavoheme enzyme containing both a flavin and a heme cofactor (figure 1.1 and 1.2) Several FHbs have been characterized to date [17-24] and are found in bacteria, fungi and protozoa. The *E. Coli* FHb, *hmp* is the most studied. The protein is a unique combination of an N-terminal globin domain and a C-terminal oxidoreductase domain (figure 1.5). The globin domain contains heme cofactor while a FAD cofactor is found in the oxidoreductase domain. FHb is an NO-dioxygenase and catalyzes the dioxygenation of NO by oxidizing NAD(P)H yielding nitrate (NO_3^-) (Scheme I). Dioxygenases incorporate both oxygen atoms from an oxygen molecule (O_2) into the substrate, in contrast to monooxygenases which introduce a hydroxyl group with water as a byproduct.

Scheme (I): $2NO + 2O_2 + NAD(P)H \rightarrow 2NO_3^- + NAD(P)^+ + H^+$

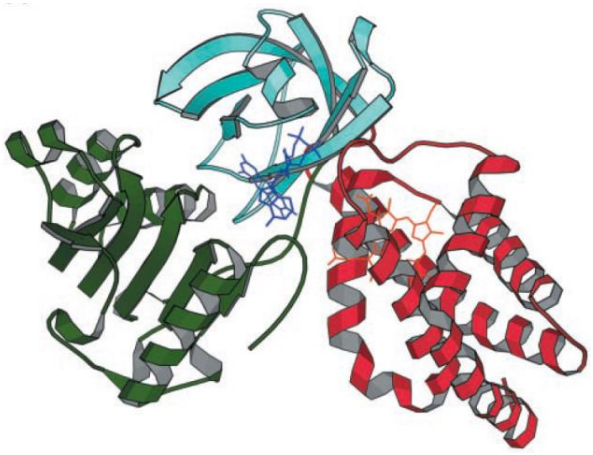


Figure 1.5: Heart shaped structure of *E. coli* FHb: The N-terminal globin domain is shown in red, the reductase domain is divided into a NAD-binding domain, shown in green and a FAD-binding domain shown in blue. Taken from Ilari *et al.* J. Biol. Chem. 2002 [25]

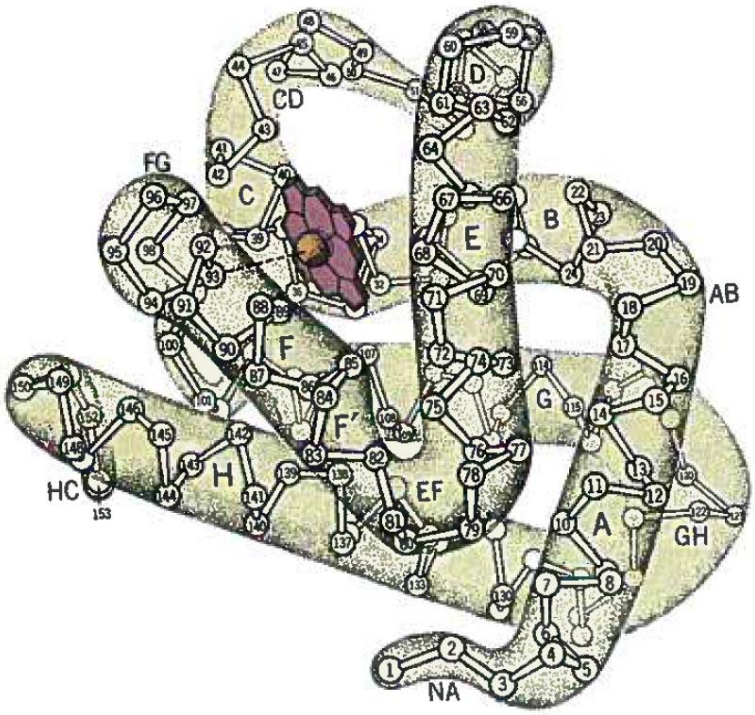


Figure 1.6: Classical globin fold. The classical globin fold (illustrated by myoglobin) consists of 6 helices with segments from A to H. A heme molecule is found in a hydrophobic pocket within the molecule, protected from solution. Taken from Garrett & Grisham, Biochemistry Third edition [26]

The Globin domain

All globin domains are structurally related to vertebrate hemoglobins (Hbs) and myoglobins (Mbs), with two types of globin folds, classical (3/3) and truncated (2/2). The classical globin fold consists of six helices consisting of eight helical segments from A to H (figure 1.6). The heme is found in a hydrophobic pocket and is five-coordinated with the iron ion ligated to the proximal histidine (His(F8)). In addition a conserved phenylalanine (Phe(CD1)) appears to be necessary to keep the heme in correct orientation within the pocket. Apart from His(F8) and Phe(CD1) a wide array of amino acid residues can be found within the heme pocket. [27]

The globin domain of FHb (~5-140 aa) is a classical globin fold with an unusually long H-helix and the D-helix is substituted by a large loop region. The heme pocket geometry shows little resemblance to that found in vertebrate Hbs and Mbs, apart from the conserved residues Phe(CD1) and His(F8).

The architecture of distal heme pocket is found unusual with Leu57(E11) isopropyl side chain closest to the heme (figure 1.7). The sidechains of Tyr29(B10) and Gln53(E7) were proposed to be involved in heme iron ligand stabilization and are found over 5 Å away. It has been suggested that a conformational change upon ligand binding will rotate the leucine sidechain away for the tyrosine hydroxyl group to interact with the ligand [28]. Additional side chains within 5Å are Phe43(CD1), Ile61(E15) and Val98(G8), and the backbone of Gln53 that with Leu57 fill the first distal shell.

The proximal heme pocket consists of a proximal histidine residue (His85(F8)) (figure 1.7). The histidine residue is involved in a hydrogen bonding network comprising of His85(F8), Tyr95(G5) and Glu135(H23) which is not found in vertebrate Hbs. The hydrogen bonds impart a rigid orientation of the imidazole ring of His85 with respect to the heme plane. Other conserved amino acids are Asn44(CD2), Leu012(G8) and Tyr124(H11) and Lys84(F15) [27]

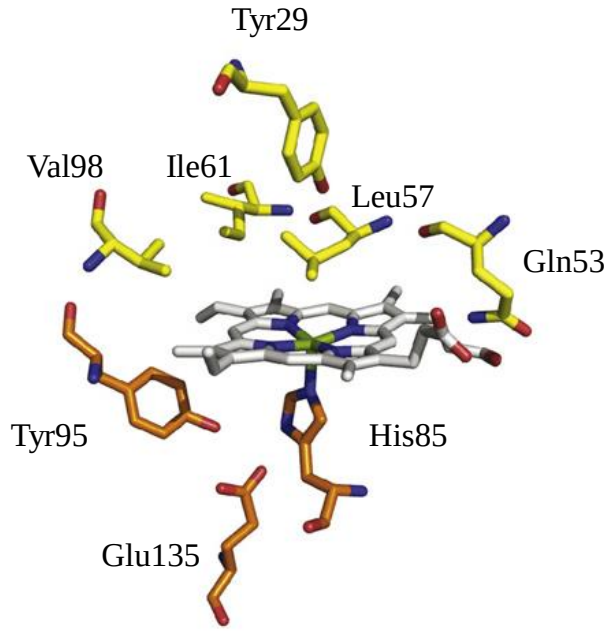


Figure 1.7: Heme-binding site of ferric unligated FHb from *E. coli* (HMP). The Heme group shown together with a selection of amino acid residues located within 5 Å of the heme. Taken from Forrester & Foster, Free Radic. Biol. Med. 2012 [13].

The FAD- and NAD(P)H-binding domain

The flavin binding domain (oxidoreductase) consists of two separate subdomains, a FAD-binding domain (−150-245 aa) and a NAD-binding domain (~250-400 aa). The FAD-binding domain consists of a six-stranded antiparallel β-barrel with Greek Key topology. A small α-helix is capping the β-barrel at the bottom and a long loop connecting Fβ2, Fβ3 β-sheets on the top. The NAD-binding domain consists of a five-stranded parallel β-sheet flanked by two helices on one side and by a helix and a loop on the other. The two subdomains form an oxidoreductase module, which structurally belongs to the ferrodoxin reductase family (FNR). [27],

The *re*-side geometry of the isoalloxazine-binding motif (ring structure of flavin) has identical orientation of the conserved residues Gln205, Tyr206, Ser207 and Tyr188 with respect to the isoalloxazine plane (figure 1.8). The conserved residues Val269 and Thr272 are also found on this side of the flavin. On the *si*-side of the flavin of *E.* coli Cys389, Phe390, Gly391 and Pro392 are found, the latter three being conserved. In *R. eutropha* with a bound phospholipid

the same conserved residues are observed but with a different orientation which might be caused by a conformational change caused by lipid binding (figure 1.7) [27].

The conserved residues in the flavin binding domain correspond to the following residues found in ferrodoxin reductases: Phe390, Gly270, Glu388, Thr272, Pro273, Gly186, Gln187, Tyr188, and Ser232. The ferrodoxin reductases are flavin binding proteins highly specialized in electron transfer from NADH to a flavin moiety and to a high-redox potential electron acceptor [27].

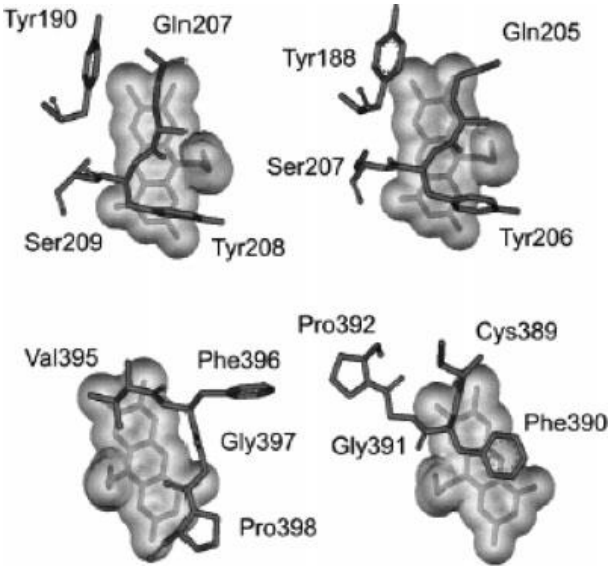


Figure 1.8: FAD binding site of FHb from, *R. eutropha* and *E. coli*. The FAD-binding site of *R. eutropha* (left) and *E. coli* (right) shown together with a selection of amino acid residues located within 5 Å. Both sides of the isoalloxazine ring are shown with the *re*-side on top and *si*-side on bottom. Taken from Bonamore & Boffi IUBMB Life, 2008 [27].

1.1.4 Catalytic mechanism

Two mechanisms have been postulated for the dioxygenation of NO: a "dioxygenation" mechanism and a "nitrosylation" mechanism. The first step in the dioxygenation mechanism involves O_2 binding to ferrous (Fe^{2+}) heme, while NO binds first in the nitrosylation mechanism (figure 1.9) [13]. Both mechanisms will result in generation of NO_3^- and oxidation of heme to the ferric state (Scheme II). Regeneration of ferrous heme is proposed to involve an initial two-electron reduction of FAD by hydride transfer from NAD(P)H. Reduced FAD then passes a single electron to the heme and ferrous heme is generated with FAD in the semiquinone state (FAD_{SQ}). NO dioxygenation may then proceed and FAD_{SQ} can transfer the second electron to the ferric heme.

Scheme (II): $NO + O_2 + Fe^{2+} \rightarrow NO_3^- + Fe^{3+}$

Scheme (III): $NAD(P)H + 2\ Fe^{3+} \rightarrow NAD(P)^+ + 2\ Fe^{2+}$

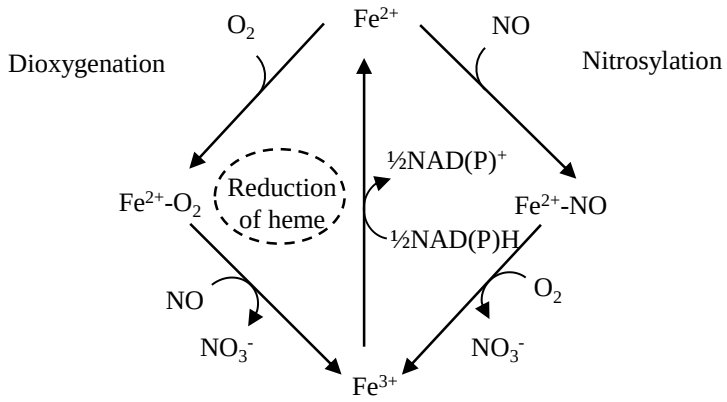


Figure 1.9: Reaction scheme for FHb. The dioxygenation mechanism (left) with O_1 binding first, followed by NO. The nitrosylation mechanism involves NO binding first (right) and regeneration of ferrous heme is performed by NAD(P)H (middle).

The first step in the dioxygenation model starts by O_2 binding to ferrous heme, which is axially ligated by a proximal histidine residue followed by NO binding forming a ferrous-oxy complex (Figure 1.10). In the next step the ferrous-oxy complex reacts with NO to form a transient Fe-ONOO intermediate, which is believed to undergo rapid isomerization and release of nitrate.

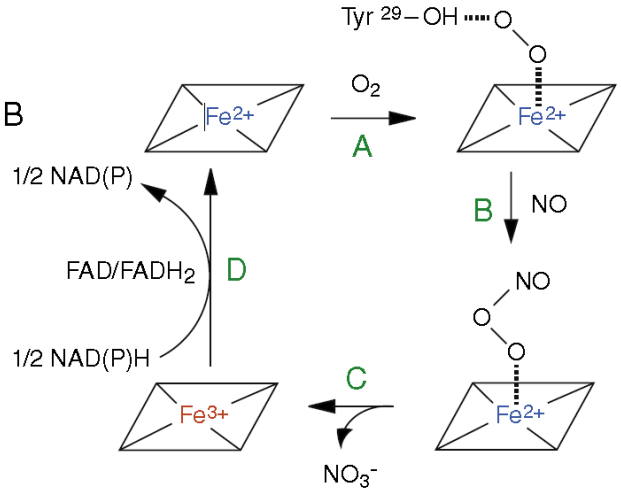


Figure 1.10: Dioxygenation mechanism for FHb. The dioxygenation mechanism starts with O_2 binding to ferrous (Fe^{2+}) heme. The ferrous oxy-complex is stabilized by Tyr29. Nitric oxide reacts with the complex and a transient Fe-ONOO intermediate is formed. Spontaneous isomerization follows with a release of nitrate. Ferric (Fe^{3+}) is reduced. Taken from Forrester & Foster. Free Radic. Biol. Med. 2012 [13].

The nitrosylation model starts with NO binding to form a ferrous nitrosyl species. An electron transfer from heme would generate a ferric-nitroxyl equivalent, which in turn would react with O_2 to form NO_3^-.

Both mechanism share the "push-and-pull" mechanism in which a proximal imidazole (His85) imparts an electron rich character to Fe^{2+} heme which allows electron transfer to either O_2 or NO_2

Research supporting the nitrosylation mechanism is the observation of a higher affinity of ferrous FHb for NO than O_2 under conditions where the remaining substrate is at saturating conditions. The K_M values are 0.28 μM for NO and 100 μM for O_2. In addition deoxygenated ferrous-nitrosyl FHb undergoes rapid generation of NO_3^- upon exposure to even micromolar concentrations of O_2. Arguing for the dioxygenation mechanism is that mutation of Tyr29 to Phe impairs both O_2-binding and catalysis with only marginal effects on NO affinity. [13]

For the dioxygenation mechanism a highly conserved Tyr29 has been postulated to stabilize the O_2-binding, but has been found over 5 Å from the heme in crystal structures for *E. coli*. Recently a computer simulation with *E. coli* FHb by Ferreira *et al.* reveals that stabilization of the ferrous-oxy complex is by H-bonding with Tyr29 mediated by water [29].

NO reductase-activity

In addition to NO dioxygenase activity FHb also has the ability to reduce NO to N_2O under anaerobic conditions (Scheme IV). It operates at approximately 1% the rate of aerobic dioxygenation. The physiological role of FHb's NO reduction ability remains unclear, and it's possible this function is overshadowed by other enzymes performing NO reduction in most bacteria [13].

Scheme (IV): $2NO + 2H^+ + 2Fe^{2+} \rightarrow N_2O + H_2O + 2Fe^{3+}$

Mühlig *et al.* conducted a recent knockout study of *Salmonella* Typhimurium, a serovar (sub-species) of *Salmonella enterica*. The bacterium mainly deploys three enzymes to detoxify NO: FHb (HmpA), flavorubredoxin (NorV) and the periplasmic cytochrome c nitrite reductase (NrfA) [30]. Experiments were performed in LB-medium mimicking conditions during fermentation of raw sausages. $NaNO_2$ salt is used for fermentation and exerts nitrosative stress. The FHb knockout showed decreased growth compared to the wild type and

the other knockout mutants under both high and low O_2 concentrations. This study indicates that NO reductase activity may be important for some systems.

1.1.5 Other studies

Several knockout studies of FHbs in different organisms have been performed and shown NO hypersensitivity, as well as decreased virulence/pathogenicity. Forrester & Foster have comprised a list of knockout studies [13]. Studies of *Bacillus subtilis* (a relative to *B. cereus*) show NO hypersensitivity after knock out of FHb [13], in addition to poor long term survival under anaerobic conditions in the presence of nitrate [31],

As mentioned it has been proposed that FHb was generated as a fusion of genes. Studies show that it is necessary that the two domains are in a single polypeptide to function properly. Individual expression of the two domains fails to complement FHb deletion *in vivo* [13]. Kaur, *et al.* 2002 made a chimeric protein by fusing cDNA encoding the single domain *Vitreoscilla* hemoglobin with the C-terminal reductase domain from *A. eutrophus* FHb. This chimeric protein had NO dioxygenase activity and was able to rescue NO-sensitivity in FHb-null *E. coli*. It is therefore interesting that there are bacteria, like *Mycobacterium tuberculosis*, that deploy a single globin domain (truncated hemoglobin) to protect against nitrosative stress. These proteins have an NO-consumption rate that are a magnitude slower than FHb [13].

Aims of the master project

As described earlier FHbs are important in protection against nitrosative stress. Nitric oxide is utilized by the immune system against pathogens and knockout mutants exhibit NO hypersensitivity, and for pathogenic species, less virulence. FHbs are not found in higher eukaryotes and is thus a promising therapeutic target. In addition FHb could be used as tool to study NO signaling in e.g. mammals.

Although FHb proteins from several bacterial relatives have been characterized, only a couple of structures from different organisms have been obtained without either a phospholipid bound (as in *R eutropha* [17]) or imidazole antibiotics [24, 25]. An additional crystal structure of FHb could contribute to our understanding of the catalytic mechanism, as only two structures of the heme pocket without ligands have been obtained.

This master thesis is focuses on three main aims for the study of FHb from *B. cereus*.

1. **Develop a purification protocol for FHb**
   Express protein using a host organism, develop a suitable purification protocol and purify sufficient amounts of protein for further analysis.

2. **Determine the X-ray crystal structure of FHb**
   Screen for crystallization conditions of FHb and determine the structure of the protein using X-ray crystallography.

3. **Biochemical characterization**
   Use spectroscopic techniques to investigate the properties of FHb.

2 Material and methods

Molecular biology methods

The gene for the protein to be investigated was purchased from GenScript, precloned into pEt22b(+) cloning vector with the restriction enzymes XbaI and BamHI (see Appendix 4).

2.1.1 Transformation of recombinant vector into competent cells

Transformation is a process in which exogenous genetic material is introduced into competent bacterial cells. The genetic material is taken up directly through the membrane, incorporated into host DNA and expressed. Recombinant plasmid pET22b(+) was purchased from GenScript and transformed into BL21(DE3) *E. coli* cells (Invitrogen).

The purchased plasmid was diluted with Milli-Q water to a concentration of 50 ng/μL for transformation. After thawing the *E. coli* cells on ice, 1-3 μL plasmid solution was added to 10 μl cells. The cells were kept on ice for 5 minutes, and then incubated at 42°C for 45 seconds before being returned to the ice for 2 minutes. 125 μL LB medium was added to the cells and shaken at 37°C for an hour at 225-250 rpm (Incubator Shaker Series 25, New Brunswick Scientific). The transformation reactions were plated on LB-plates with 100 μg/mL ampicillin and left over night at 37°C. The LB-plates were inspected for growth of colonies the next day. Two control reactions were used, positive control with 1 μl control plasmid (for ampicillin resistance) and a negative control without plasmid.

2.1.2 Making bacterial freeze stocks

Bacterial stocks are made for long term storage of plasmids within an appropriate host. Glycerol stocks can be stored for several years. An over-night culture of 5 mL LB-medium (100 μg/mL Amp) was inoculated and incubated over night at 30°C in a shaker at 225-250 rpm (Incubator Shaker Series 25, New Brunswick Scientific). After growth is observed, 200 μL 60 % glycerol solution is added to 800 μL of the overnight culture. Finally the stock is frozen with liquid nitrogen and stored at –80°C

To recover bacteria from the stocks, only a tiny part of the still frozen stock is streaked on LB-agar plates (100 μg/mL Amp). The plates are grown overnight at 37°C until colonies are visible.

Protein methods

2.1.3 Over-expression of recombinant gene in *E. coli* cells

Bacteria from freeze stocks were grown on LB-agar plates (100 μg/mL Amp) overnight at 37°C. Immediately after colonies were observed the plates were removed and stored in room temperature (20°C) for slow growth. An over-night culture of 200 mL LB-medium (100 μg/mL Amp) was inoculated and incubated over night at 30°C in a shaker at 225-250 rpm. The over-night culture was diluted by transferring 50 mL to 1 L TB-medium (100 μg/mL Amp) (See Appendix 3) and incubated at 30°C in a shaker at 225-250 rpm. After about 8 hours the culture was diluted 20 times into new flasks with 1 L TB-medium (100 μg/ml Amp) each, usually 12 flasks.

The cells were grown until OD_{600} reached 0.8 and, the heme precursor, δ-aminolevulinic acid (ALA) was added to a concentration of 0.2 μM and expression was induced with 1.0 mL 1 M IPTG after cooling the medium on ice to 20°C or less. The culture was incubated overnight at 20°C at 225-250 rpm. Bacteria were harvested by centrifugation at 6500 *g* the following day.

2.1.4 Protein purification

Lysis of bacterial cells

For cell lysis an AB X-press® system [32] was used. Frozen bacterial cells were lysed by high pressure and mechanical stress. Frozen bacterial cells were added to a precooled X-Press column (-20°C) and mounted on a stand. Using mechanical pressure the bacterial cells were pushed through a tiny hole of the X-press column repeatedly, which breaks the cells, freeing its contents.

Precipitation of DNA

The crushed cells were then dissolved in Buffer A (Appendix 3) containing ½ protease inhibitor pill (cOmplete ULTRA Tablets, Mini, EasyPack by Roche) at 4°C and centrifuged at 48 000 *g* for 1 hour at 4°C to remove cell debris. A 10% streptomycin sulfate solution (pH 7.5) was added slowly, drop by drop to the supernatant, to a final concentration of 2 % streptomycin sulfate in order to precipitate DNA. The solution was then centrifuged at 48 000 *g* for 30 minutes to remove the precipitated DNA. The supernatant was used for ammonium sulfate precipitation.

Ammonium sulfate precipitation

Ammonium sulfate ($(NH_4)_2SO_4$) precipitation is a method to purify proteins, as different proteins precipitate at different concentrations of ammonium sulfate. After addition of ammonium sulfate the precipitated proteins are separated from proteins in solution by centrifugation. When the process is repeated with sequentially higher ammonium sulfate concentrations the precipitates/pellets are essentially different fractions containing different proteins. The process can be optimized by varying salt concentration to remove proteins that precipitate before and after the protein of interest.

A procedure of adding ammonium sulfate in 3 steps was used, where each step was followed by centrifugation at 48 000 *g*. First 0.21 g/mL salt was added and spun, followed by 0.09 g/mL (30 g/mL in total). This was repeated until the ammonium sulfate concentration was 0.43 g/mL yielding three pellets.

Chromatography

All chromatography procedures were performed using the Äkta Purifier-system from GE Healthcare. This Äkta system can measure the absorbance and conductivity of the protein samples. Absorbance is measured in order to observe and separate proteins from each other while conductivity is an indicator of the salt content. The Äkta system can be operated manually or automatically.

Most of the columns used were packed by the group, column material and empty columns have been supplied by GE healthcare. During elution absorbance was measured at 280, 400 and 450 nm. All proteins absorb light at 280 nm as the aromatic amino acids (tyrosine and

tryptophan, and to a less degree phenylalanine) absorb light at this wavelength. Heme has a Soret peak (absorption maximum) at approximately 400 nm, and flavins have a maximum at 450 nm. Fractions with high absorbance at 400 nm and 450 nm were collected in the following chromatography procedures.

The following columns were used

XK 26 column with 70 mL Sephadex 25 (packed by the group)

XK 26 column with 70 mL HPQ Sepharose (packed by the group)

XK 16 column with 120 mL Superdex 200 (packed by the group)

Prepacked Superdex 200 10/300 GL (GE Healthcare)

HiTrap Desalting column, 5 mL (GE Healthcare)

Desalting

The precipitated protein was resolved in Buffer A, and a desalting column (Sephadex 25) was used in order to lower the salt concentration of the sample by separating the low molecular weight components (salt) and the high molecular weight molecules (proteins). A low salt concentration is essential for a successful ion-exchange chromatography, since salt will compete with protein for binding to the anion-exchange columns. Buffer A was used; the Äkta system was operated manually by starting to collect the protein fraction when an increase in absorbance at 400 nm was observed and stopping the collection when absorbance at 400 nm decreased and before the conductivity rose.

Anion-exchange chromatography

Anion-exchange is a method for separating proteins by the charges on the protein surface. Proteins will bind to the column by electrostatic interactions. The charges on the protein surface depend on the pH as some of the amino acid residues will get protonated or deprotonated affecting the total charge. During anion exchange chromatography proteins with negative charges on the protein surface bind to the column material, the column material itself is positively charged. Protein elution is performed by increasing the salt concentration as the negative ions of the salt will compete with the proteins for binding.

Different salt gradients can be used for elution, including simple linear gradients and step-gradients. A slack gradient will result in higher resolution, separating proteins of similar charge. The step gradients used consist of a series of linear gradients with “plateaus” (flat segments with no increase in salt concentration) to elute proteins that bind slightly stronger or weaker than the protein of interest.

For ion-exchange a XK 26 column with 70 mL HPQ-Sepharose was used with buffer A and buffer B. The procedure was optimized for the protein so that other proteins that bind poorly will elute early and proteins that bind tighter will be washed out by increasing the salt concentration. Both linear and step gradients were used for purification, more details are found in the Results-section.

Gel-filtration chromatography

Gel filtration chromatography is a type of size exclusion chromatography with an aqueous mobile phase and a stationary phase. Gel filtration separates protein based on protein size. The gel material consists of beads with pores that allow smaller molecules access to a larger volume than larger molecules. Larger molecules will have less volume available and will pass through the column quicker than smaller molecules. Protein samples from anion-exchange chromatography procedure were purified by gel filtration chromatography. 200 µl of concentrated protein was loaded on a Superdex 200 column (XK16) run with Buffer C separating the proteins with respect to size. The fractions with high 400 and 450 nm values were collected.

Sodium dodecyl sulfate-polyacrylamide gel electrophoresis (SDS-PAGE)

Samples from different purification steps can be analyzed by sodium dodecyl sulfate (SDS) polyacrylamide gel electrophoresis (SDS-PAGE). SDS is an anionic detergent that binds to proteins, causing it to linearize and become negatively charged. The charge of the protein becomes proportional to the mass, and proteins are separated by mass on the gel by an electric field.

For analysis protein samples were mixed with NuPAGE® LDS sample buffer (Life technologies) with a volume of a quarter of the sample volume and incubated at 96°C for 5 minutes, vortexed and applied to a 10- or 15-well NuPage 4-12% Bis-Tris gel

(Life technologies). The gel was run for 35 minutes at 200 V with NuPage® MOPS SDS Running Buffer (Life technologies). 10 μL of Novex® Sharp Pre-Stained Protein Standards (Invitrogen) ladder was used.

After gel electrophoresis the gel was either stained with Brilliant Blue R250 (Sigma) or InstantBlue (Expedeon). For Brilliant Blue staining the gel was soaked in 0.1% Coomassie Brilliant Blue R250, 40% EtOH (ethanol), 10% HAc (Acetic acid) and heated for 20 seconds in a microwave before being transferred into destaining solution and destained overnight. InstantBlue Coomassie staining was performed by leaving the gel in solution until bands were visible, ca. 15 minutes.

Ultra filtration

Ultra filtration was used to concentrate samples from ion-exchange and size exclusion chromatography for crystallization. Amicon Ultra 30 K-15 (or -4) or Centricon Plus 30 K-70 tubes were used as describes by the manufacturer (Millipore). A cutoff of 30 kDa was used as FHb has a molecular weight of 48.8 kDa. A typical set up for the Amicon Ultra 30 K-15 was 4750 *g* at 4°C with a fixed rotor angle for 18 minutes.

2.1.5 Protein crystallization

SDS-PAGE and UV-vis was used to analyze the purity and indicate cofactor content of the samples. Several crystallization trials were performed on pure protein samples. Crystallization screening was performed using either an Oryx6- (Douglas Inst. Ltd.) or Mosquito robot (TTP Labtech) offered by the Structural biology Core Facility at Rikshospitalet Oslo University Hospital (Bjørn Dalhus). Both robots prepared 96 sitting drop experiments in crystal trays, with the Mosquito being considerably faster, and needing less protein, but with a smaller drop size (0.2 μL in contrast to 0.8 μL for the Oryx6 robot). The Oryx6- and Mosquito robot prepared drops in room temperature and at 4°C, respectively. The crystal trays were stored either in room temperature or at 4°C. The trays were monitored with a light microscope immediately after drop preparation, in order to check for crystal growth, and regularly for the next weeks and (less frequently) months.

Optimization of crystal conditions was performed manually with 2 μL sitting drops at room temperature under a light microscope **(**SMZ168 LED, Motic**)** with a 1:1 ratio between the protein solution and crystallization solution.

2.1.6 Reconstitution of heme and FAD

Reconstitution of heme-deficient FHb with heme was performed under reducing conditions (10 mM DTT) and with the presence of catalase as described by Gardner, NO dioxygenase assays [33], using degassed buffer instead of N_2-sparged buffer and in a smaller scale. FHb (0.5 to 0.75 mM) was prepared in 400 μl Buffer D (3000 U of catalase) and 20 mM hemin (dissolved in dimethyl sulfoxide) was added slowly to a final concentration equal to the concentration of FHb. Dithionite (~2 mg freshly dissolved in water) was added after 15 minutes of incubation at 20°C, followed by 60 minutes of incubation at 37°C. The solution was then run on a 5 ml HiTrap Desalting column (Pharmacia) with buffer D and a Superdex-200 column to separate the target protein from heme, catalase and reductants.

Initial testing was performed with 100 μM FHb in 100 μL buffer. The same procedure was also tested with FAD (dissolved in MilliQ-H_2O) instead of heme, and by addition of both heme and FAD to the same solution, e.g. 100 μM heme, 100 μM FAD and 100 μM FHb in 100 μL buffer.

Protein Assays

2.1.7 Ultraviolet-visible spectroscopy

UV-vis spectroscopy was performed to determine protein concentration and indicate heme and FAD content of FHb.

Protein concentration of FHb was determined using Agilent 8453 Diode array spectrophotometer. Protein samples were diluted until the absorbance measured was below 1.0. The absorbance at 405 nm was measured and protein concentration was calculated using the extinction coefficient 88 600 $M^{-1}cm^{-1}$ [34]. Indication of heme content was measured by comparing absorbance at 280 nm and 405 nm. All proteins absorb at 280 nm and heme has a Soret peak at ~400 nm.

2.1.8 Assay for Heme content

A small pinch of dithionite (2 mg) was dissolved in MilliQ-H_2O. A 300 μL sample of FHb containing 1.2 to 12 nmol of heme was added to 300 μL of 4.4 M pyridine/0.2 M NaOH solution and mixed. Absorbance was measured at 556 nm and 539 nm before and immediately after adding the dithionite solution and gently mixing the sample in a cuvette.

Potassium ferricyanide (100 mM) was then added stepwise at 3-μL aliquots to the cuvette and absorbance at 556 nm and 539 nm were recorded. Heme concentrations in the sample was calculated from the absorbance differences of the reduced minus oxidized heme at 556 nm and 539 nm. μM heme is given by the following formula: μM heme = 46.7 (ΔA556 – ΔA539) [35, 36]

2.1.9 FAD analysis

10 to 20 nmol protein was dissolved in 1 ml Buffer C and boiled for 3 minutes. The solution was then centrifuged at 20 000 g to remove the denatured protein. 0.8 ml of the supernatant volume was added to 1.2 ml Buffer A and FAD-content was measured by Absorbance at 450 nm with an extinction coefficient of 11 300 $M^{-1}cm^{-1}$.

2.1.10 Bradford protein assay

Protein concentration was measured using a Quick Start Bradford protein assay (Biorad) and Modified Lowry protein assay, explained under [37].

Approximate protein concentrations were obtained through UV-vis spectroscopy by measuring absorbance at 280- and 405 nm with extinction coefficients of 44 900 $M^{-1}cm^{-1}$ (calculated with ProtParam [38]) and 88 600 $M^{-1}cm^{-1}$ [34], respectively.

The linear range of the Bradford assay with bovine serum albumin (BSA) is from 125 μg/mL to 1000 μg/mL. Protein concentration is determined by making a standard curve (e.g. linear regression) with a range of BSA concentrations. Concentration of the target protein is then calculated from absorbance of the sample by utilizing the standard curve.

The 1 mL assay with disposable cuvettes was performed to determine protein concentration. A prediluted BSA standard was used (GBiosciences) with the following concentrations: 100 μg/mL, 200 μg/mL, 300 μg/mL 500 μg/mL 800 μg/mL and 1000 μg/mL. In order to determine protein concentration, 1 mL of Coomassie Brilliant Blue G-250 dye (Biorad) was added to 20 μL of each standard and unknown sample solution and mixed. The samples were incubated for 5 minutes at room temperature and absorbance were measured at 595 nm and compared to the BSA standard curve.

2.1.11 Modified Lowry protein assay

Modified Lowry assay was used for protein determination. Similarly to the Bradford assay, protein concentration is calculated from a standard curve. The same prediluted BSA standard for the Bradford assay is used for this method. The BSA standard is not linear within the range, and a point-to-point curve is preferable to a linear fit, if plotted by hand, according to the manufacturer [39].

1X Folin-Ciocalteu Reagent (Thermo) was prepared by diluting with MilliQ-water. BSA was used as a standard from 100 μg/mL to 1000 μg/mL (GBiosciences). 0.2 mL of each standard and unknown sample replicates were prepared. At 15 second intervals 1 mL of Modified Lowry Reagent (1x) (Thermo) is added to each tube, mixed and incubated at room temperature for 10 minutes. 100 μL of 1X Folin-Ciocalteu Reagent was added after exactly 10 minutes, maintaining the 15-second interval between tubes. Test-tubes were covered and incubated for 30 minutes at room temperature. Absorbance was measured at 750 nm and concentrations were calculated from the BSA standard. The spectrophotometer was blanked with a sample of water. “Blank samples” without protein, were prepared as protein samples, and measured at 750 nm and subtracted from each standard and unknown protein sample.

3 Results and discussion

Expression

Plasmids ordered from (GenScript Inc.) were successfully transformed into competent BL21(DE3) cells. FHb was expressed and SDS-PAGE was used to identify over-expressed proteins which showed over-expression after induction with IPTG (Figure 3.1a). A total number of three over-expressions were performed. Because of low cofactor content of the pure protein, discussed later (see UV-vis and heme assay), the amount of ALA was increased from 0.2 μM to 0.25 μM and 0.3 μM riboflavin was added (Figure 3.1b) for the last run.

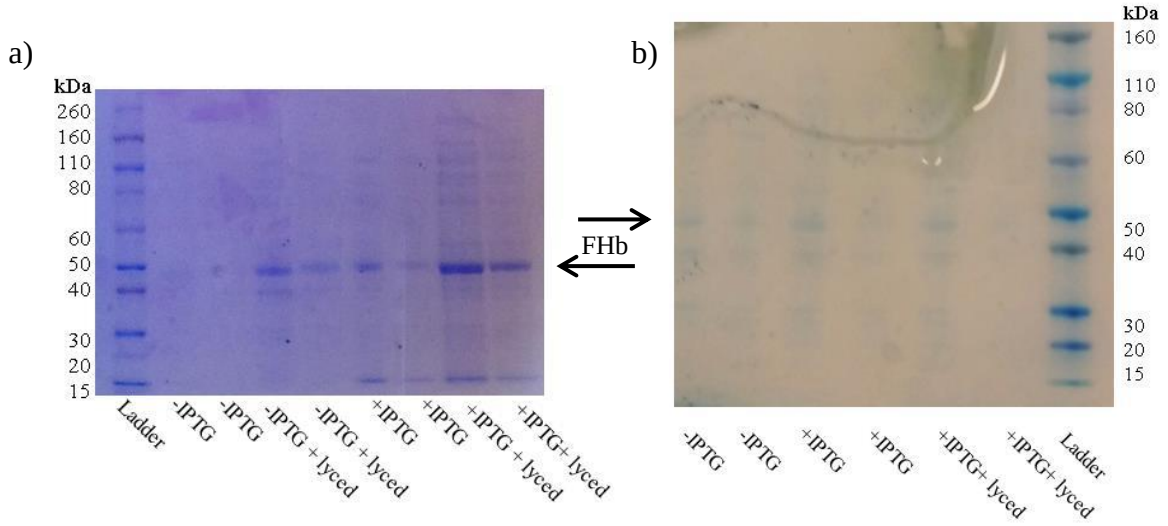


Figure 3.1: SDS-PAGE analysis of protein expression: Coomassie stained polyacrylamide gel showing expression of proteins before and after induction with IPTG. a) Over-expression with 0.2 μM ALA, b) Over-expression with 0.25 μM ALA and 0.3 μM riboflavin. Cells are either lysed or not, as indicated on the figure. Lysed cells show proteins in the cell cytoplasm and indicate that the protein is soluble. Each sample was applied twice at different concentrations the second (right) are diluted 2-fold.

Both gels have clear bands between 40 kDa and 50 kDa, which indicate that FHb is overexpressed (MW 44.8 kDa). A band of similar size is present even before induction. This might be due to another protein of similar size or some expression of FHb even before induction. The FHb band is clearly present after cell lysis, which indicates that the protein is successfully overexpressed and soluble inside the cells. After overexpression a yield of ~150 g bacteria is obtained for 12 L of TB-medium or ~12.5 g per L of cell culture.

Purification

In order to obtain pure protein samples for crystallization a 3-step purification procedure was established and optimized consisting of ammonium sulfate precipitation, anion exchange chromatography and gel filtration.

3.1.1 Ammonium Sulfate precipitation

The first step of purification is ammonium sulfate precipitation; figure 3.2 shows an SDS-PAGE analysis of this step. A band of expected size (~45 kDa) is present in all of the pellets/fractions but in higher concentration in the last two (0.30 g/ml and 0.43 g/ml). There is no band at ~45 kDa in the supernatant in figure 3.2 indicating that the protein has precipitated and is found in the pellets. Holo-FHb contains both heme and FAD and has an orange color. FHbs from other organisms have shown that heme and FAD content varies and both cofactors can be lost during purification, resulting in substoichiometric heme and FAD [33]. The last precipitation had the strongest reddish color (figure 3.3) indicating a higher content of heme. Further purification was performed with the pellets precipitated with 0.30 g/mL and 0.43 g/mL ammonium sulfate either separately or by combining the two (mix-fraction).

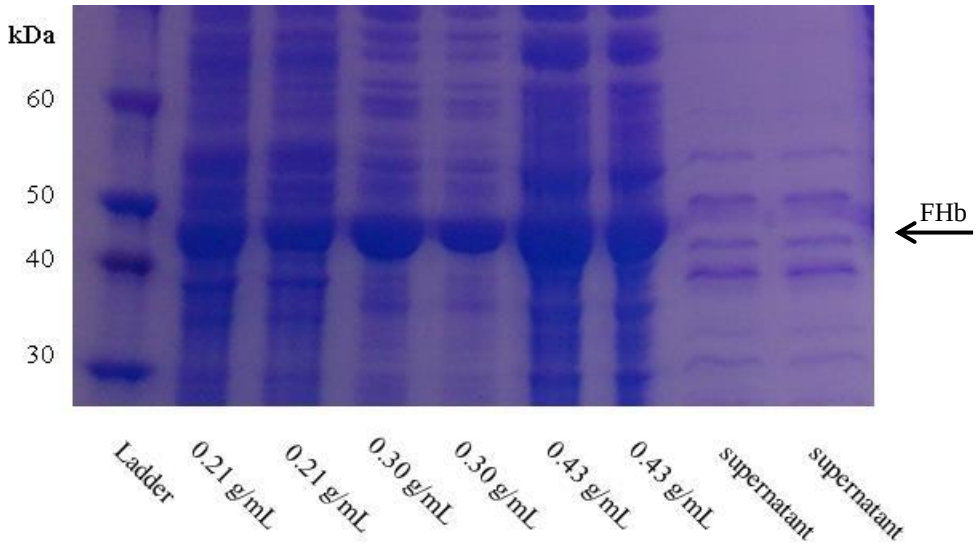


Figure 3.2: SDS-PAGE analysis of ammonium sulfate precipitation: The gel shows protein samples after resuspension of pellets from ammonium sulfate precipitation. Precipitation is performed in three steps with 0.21 g/mL, 0.30 g/mL and 0.43 g/mL ammonium sulfate (see Methods). Supernatant after addition of 0.43 g/mL and centrifugation is also shown. Each sample is applied to the gel twice with the first (left) with 2-times more protein.

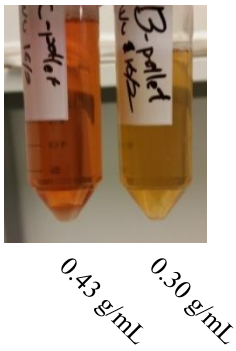


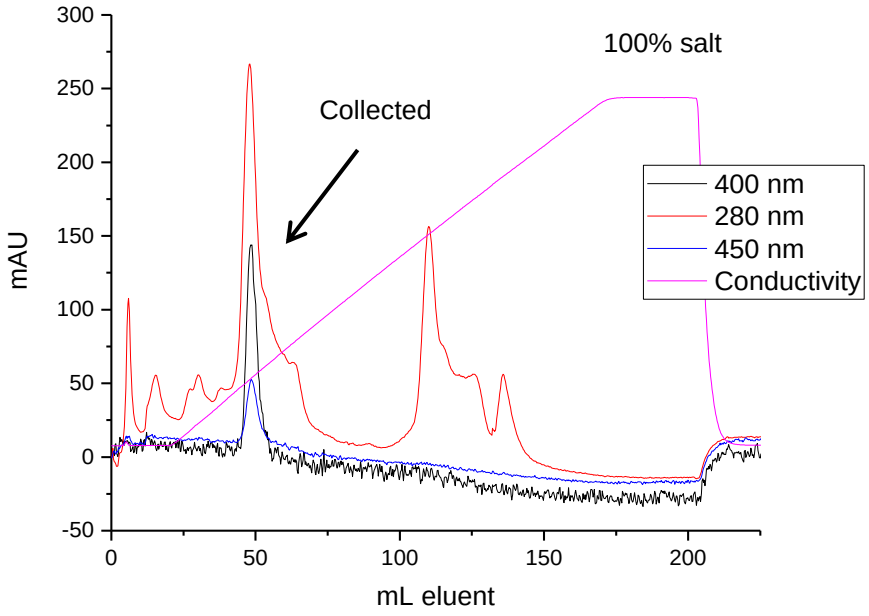
Figure 3.3: Picture of protein sample after resuspension of the pellets obtained after ammonium sulfate precipitation. The 0.43 g/mL sample has a stronger reddish color than the 0.30 g/mL sample.

3.1.2 Anion Exchange chromatography

The second step of purification is anion exchange chromatography (IEX). Optimization of the purification procedure was mainly performed on this step. An overview of the different programs/gradients used is found in table 3.1. A small scale test was performed (figure 3.4) with a gradient from 0-100 % salt at pH 7.5, with the 0.43 g/mL and 0.30 g/mL fractions from ammonium sulfate precipitation. Both fractions were tested as the fractions had different color, indicating a difference in cofactor content and with other proteins present two different IEX-procedures might be needed.

A peak with absorption at 400 nm is observed at a salt gradient of ~20 % for both pellets. The peak is distinctly larger for the 0.43 g/mL and with a lower 280 nm peak. Indicating less protein but more heme cofactor as is expected by the color of the sample (Figure 3.3). The 0.43 g/mL fraction also contains more of other proteins and the "shoulders" are larger, and possibly less pure than the 0.30 g/mL sample. The 400 nm peaks were collected and analyzed by SDS-PAGE (figure 3.5) for assessing degree of purity and to aid optimization. The SDS-PAGE shows a band at the expected size for both ammonium sulfate fractions. Some impurities are observed but the protein is already quite pure after this step.

a)



b)

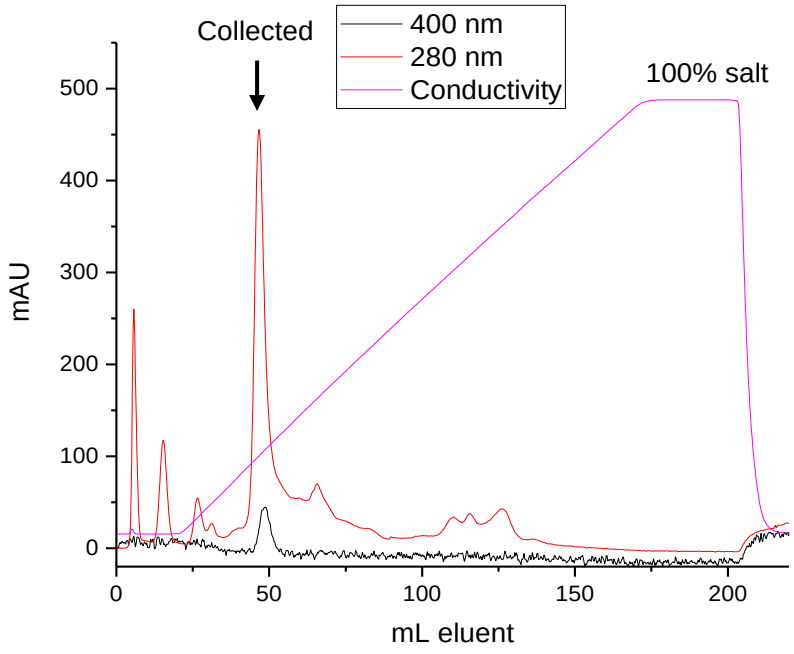


Figure 3.4: Small scale anion exchange chromatogram
Chromatogram from small scale IEX with a salt gradient of 0-100%. The black line, red line and blue line correspond to 400 nm, 280 nm and 450 nm, respectively. The pink line is measured conductivity. Protein with absorption at 400 nm is collected (indicated in the figure). a) Chromatogram for 0.43 g/mL pellet,
b) Chromatogram for 0.30 g/mL pellet

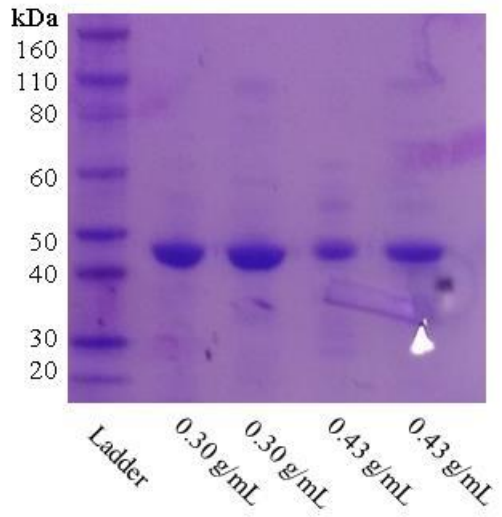


Figure 3.5: SDS-PAGE analysis of small scale IEX
The gel shows collected fraction after small scale IEX. Each sample is applied to the gel twice with the first (left) with 2-times more protein.

Table 3.1 Overview of anion-exchange experiments performed

Fraction	Gradient type	Gradient	Chromatogram	Gel
0.43 g/mL	Linear	0-100%	Fig. 3.4a	Fig. 3.5 and 3.10
	Linear	0-30%	Fig 3.6a	Fig 3.10
	Step	Step1	Fig 3.6b	-
0.30 g/mL	Linear	0-100%	Fig 3.4b	Fig 3.5 and 3-10
	Linear	0-50%	Fig 3.7b	Fig 3.10
	Step	Step1*	Fig 3.7a	-
Mix	Linear	0-30%	Fig. 3.9	Fig 3.10
	Step	Step1*	Fig. 3.8a	Fig 3.10
	Step	Step2**	Fig. 3.8b	Fig 3.10

Overview of anion exchange experiments performed. Step gradients consist of 6 steps of 2 column volumes (CV) each, unless specified otherwise. After the step gradient the column is washed and re-equilibrated with 5 CV and 5 CV, respectively. *Step1: 0-0%, 0-8%, 8-8%, 8-17% (8 CV), 17-17%, 17-30%
**Step2: 0-0%, 0-12%, 12-12%, 12-20% (10 CV), 20-20%, 20-50%

Large scale purification

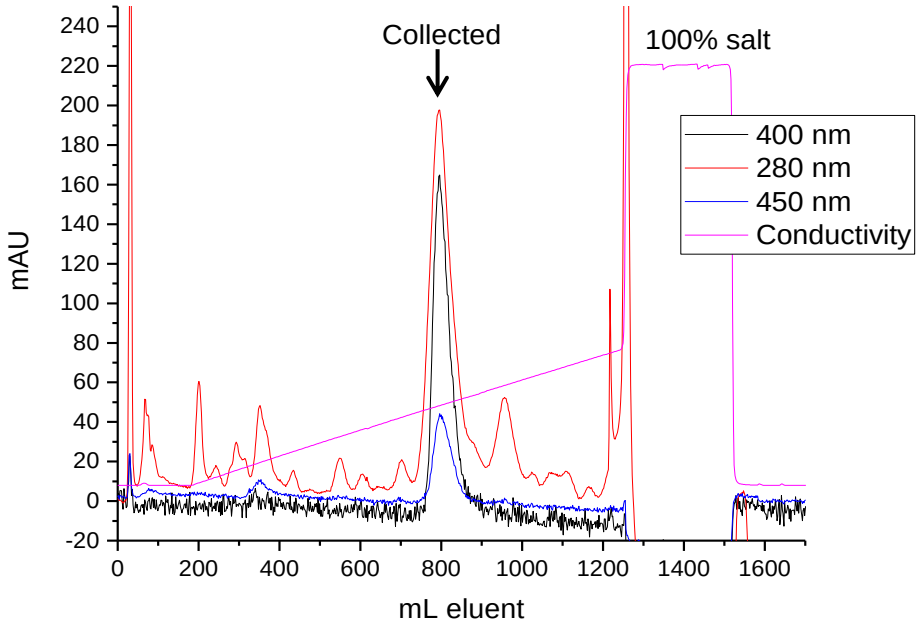
After the initial test of anion exchange, large scale purification was performed at pH 7.5. SDS-PAGE analysis was performed after chromatography and a selection of SDS-gels are shown in figure 3.10.

The 0.43 g/mL fraction

As the 0.43 g/mL-fraction was expected to have the heme cofactor this was purified first. Based on the chromatogram from the small scale tests the gradient range was reduced to obtain a better separation of FHb from the other proteins. This resulted in a linear gradient for large scale purification from 0-30% salt at pH 7.5 (figure 3.6a) and was the main purification method for this fraction/pellet. During purification of the 0.30 g/mL fraction a step gradient was tested, (Step1 found in table 3.1). The Step1 gradient was also tested on the 0.43 g/mL pellet (figure 3.6b)

The chromatogram for purification with a linear gradient (Figure 3.6a) has a similar peak as during initial testing but the "shoulders" are smaller. The protein has a high degree of purity as seen on the SDS-gel. However it is difficult to compare the PAGE-analysis between the initial test and the large scale purification, although the chromatogram shows a better separation. The step gradient in figure 3.6b looks promising with a broader peak and better separation but there is still some background. pH during the step chromatography is uncertain but probably around 7 as there were some problems with the pH-meter at the time, which was discovered later.

a)



b)

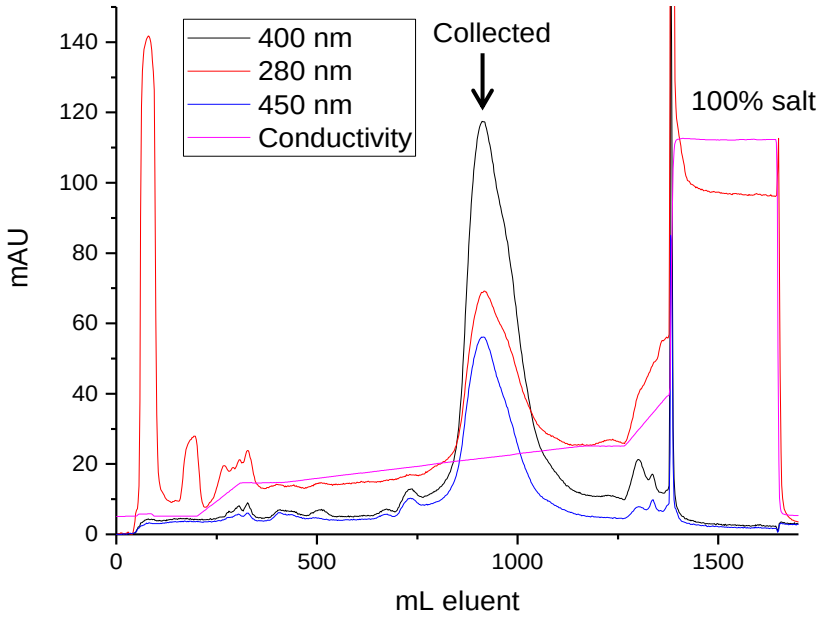
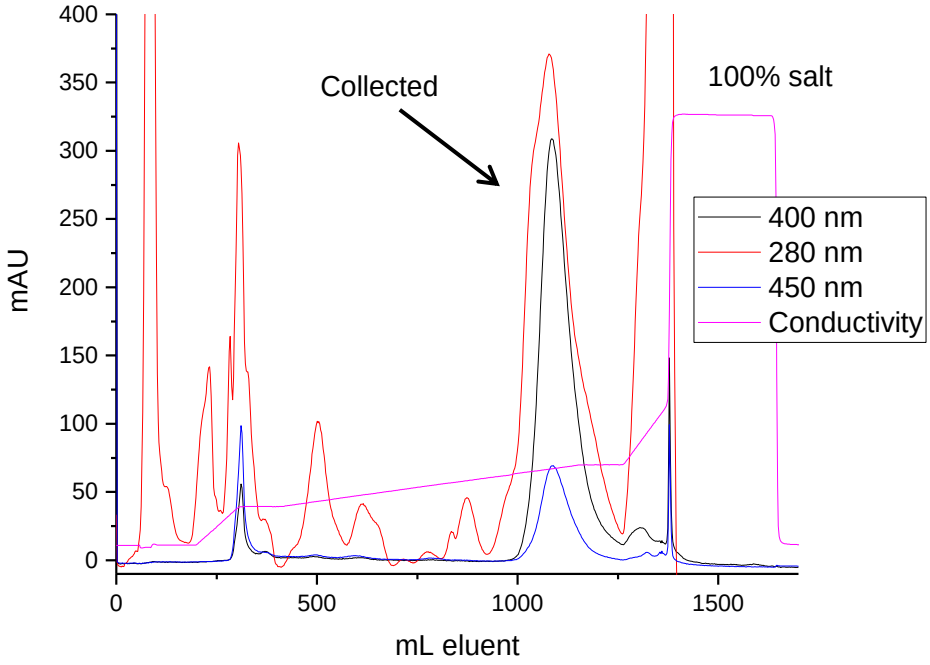


Figure 3.6: Anion exchange chromatograms for 0.43 g/mL fraction
The black line, red line and blue line correspond to 400 nm, 280 nm and 450 nm, respectively. The pink line is measured conductivity. Protein with absorption at 400 nm is collected as indicated in the figure. The chromatograms show two different gradients: a) linear gradient from 0-30% salt and b) Step1.

a)



b)

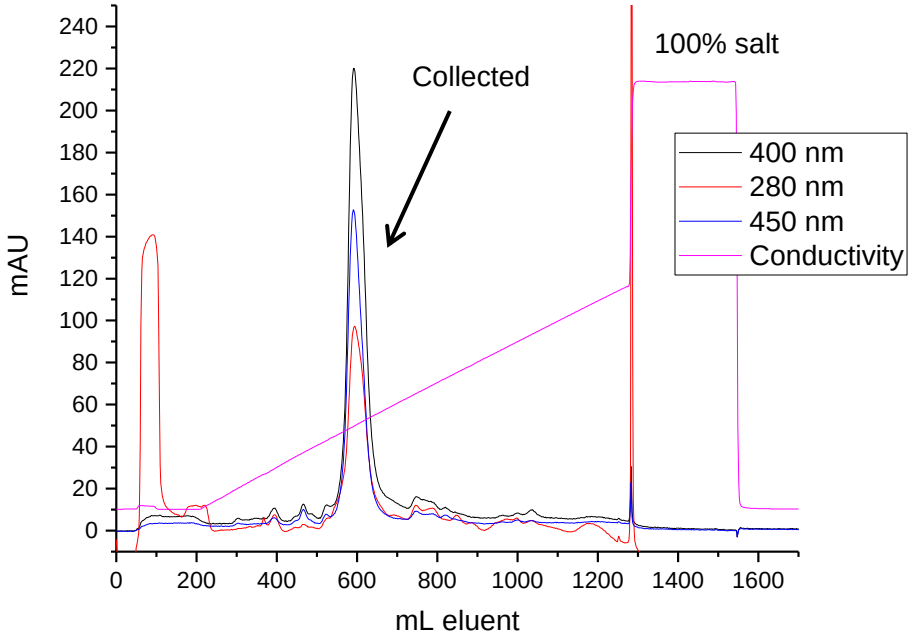


Figure 3.7: Anion exchange chromatogram for 0.30 g/mL-fractions

The black line, red line and blue line correspond to 400 nm, 280 nm and 450 nm, respectively. The pink line is measured conductivity. Protein with absorption at 400 nm is collected as indicated in the figure.
The chromatograms show two different gradients: a) Step1 and b) a linear gradient from 0-50% salt.

The 0.30 g/mL fraction

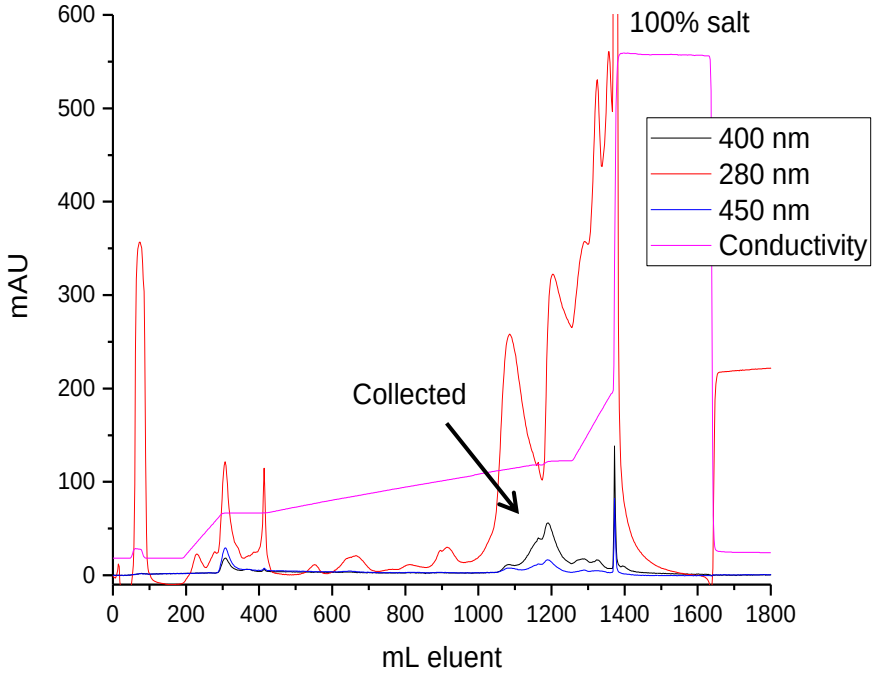
Even though the 0-30 % gradient worked for the 0.43 g/mL fraction, a step gradient was tested (Step1 in Table 3.1) for the 0-30 g/mL fraction to improve separation of proteins. The Step1 gradient is based on elution before 18% salt, as was the case during purification of 0.43 g/mL, with a slack gradient from 8% salt to 17% salt. FHb elutes at a quite high salt concentration compared to this range and there were some problems with FHb eluting late during the program (results not shown). In response to this a simple gradient from 0-50% salt was used for main purification of the 0.30 g/mL fraction (figure 3.7b). From the chromatogram with the linear gradient the protein is expected to be quite pure, which is confirmed by the gel (figure 3.10).

Mixing of ammonium sulfate -fractions

Purification of both ammonium sulfate fractions yield high purity after anion exchange and mixing of the pellets was tested. The result after mixing the pellets will be called the "mix-fraction" As the 0.43 g/mL fraction contains more heme as indicated by the color, the fractions could be separated in order to get protein with more cofactor for characterization and crystallization. However (as discussed later) the heme content varies and reconstitution after purification is needed anyway as crystals require a homogenous solution (either only apo- or holo-protein). In addition by mixing the pellets, which corresponds to removing an ammonium sulfate precipitation, it would be possible to simplify and "speed up" the purification procedure by needing fewer separate steps.

Both the 0-30% linear gradient and the Step1 gradient were tested. In addition another step gradient was created (Step2, see Table 3.1). Step2 was created with expected elution between 12% and 20% based on the observation from the Step1 gradient that the protein is eluting a bit late. It is worth noting that the sample from this expression had a lower heme content than previous runs, as the 405 nm / 280 nm ratio (discussed later) is consistently lower after purification and is also observed in figure 3.8. This also shows that the cofactor content can vary between different expression runs. An offset observed between the main 280 nm peak and 400 nm peak, as seen in Figure 3.8 for Step2, indicates that apo-FHb has a slightly lower affinity for the column. Both step gradients have two distinct flavin peaks at 450 nm within the main 280 nm peak. This is not observed for the other ammonium sulfate fractions.

a)



b)

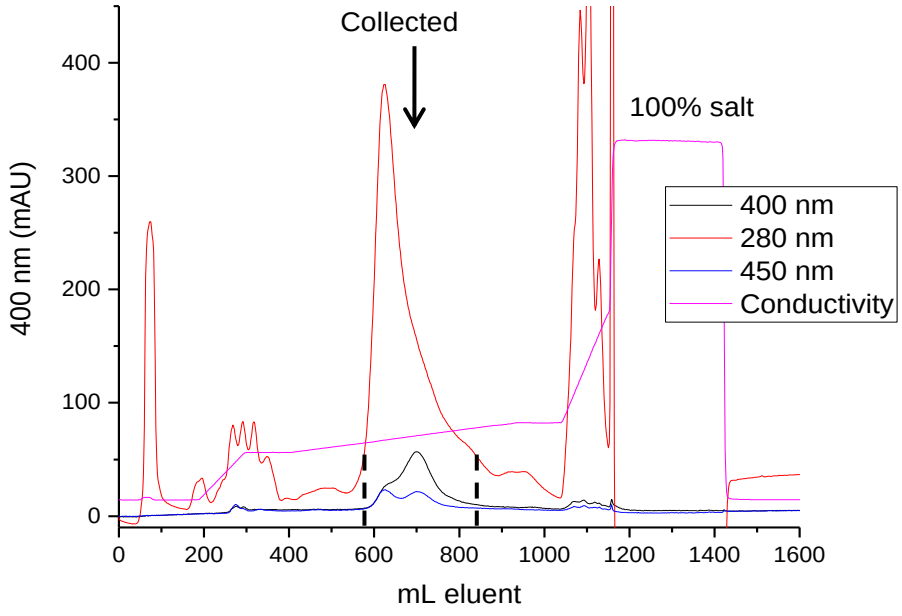


Figure 3.8: Anion exchange chromatogram with step gradients for the mix fraction.

Chromatograms for the mix fraction. The black line, red line and blue line correspond to 400 nm, 280 nm and 450 nm, respectively. The pink line is measured conductivity. Protein with absorption at 400 nm is collected as indicated in the figure. The chromatograms show two different gradients: a) Step1 and b) Step2.

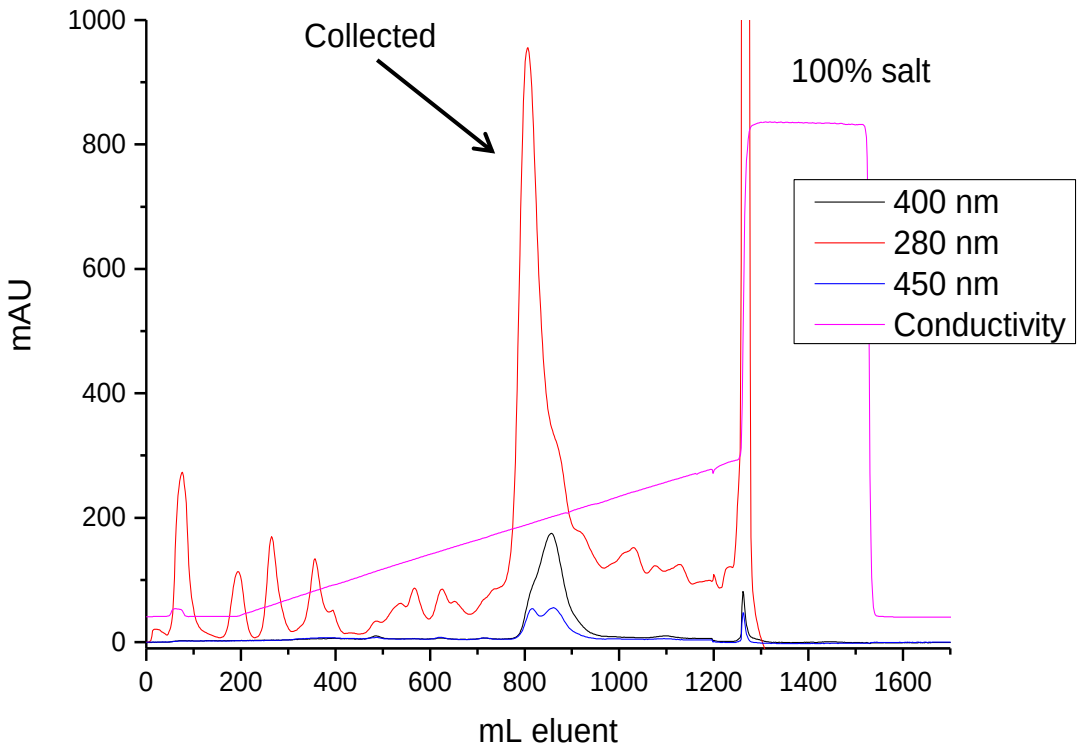


Figure 3.9: Anion exchange chromatogram with a linear gradients for mix fraction
Chromatogram of mix fraction with a salt gradient from 0-30%. The black line, red line and blue line correspond to 400 nm, 280 nm and 450 nm, respectively. The pink line is measured conductivity. Protein with absorption at 400 nm is collected as indicated in the figure.

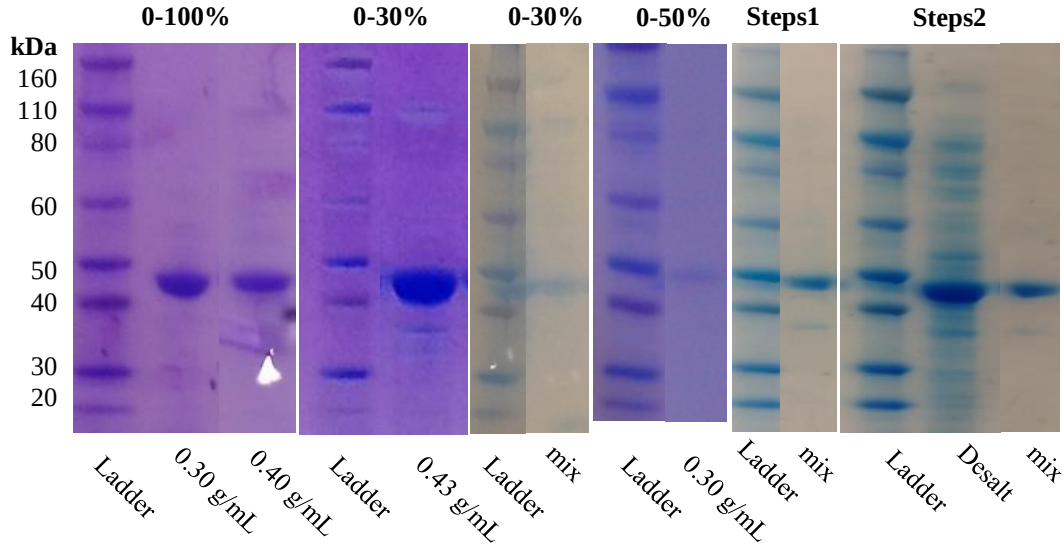


Figure 3.10: SDS-PAGE analysis of anion exchange chromatography
The figure shows a collection of different SDS-gels, divided by white lines. The top row indicates the salt gradient used for purification, while the bottom row indicates the ammonium sulfate fraction.

The chromatograms with a linear gradient (figure 3.9) and Step2 (figure 3.8b) show a single peak at 280 nm and 400 nm for FHb. The Step1 (figure 3.8a) 400 nm peak is found between two peaks at 280 nm, it is however likely part of the first peak.

SDS-PAGE analyses show that all methods yield a pure protein, but containing a varying fraction of apo-protein. For the mix Step2 seems to be the method giving the highest purity. The chromatograms are in agreement with this, with the Step2 showing less other protein as seen from the 280 nm detection curve, with Step 1 being the worst.

Both the linear gradients with 0-30 % or 0-50% salt are good for purifying the ammonium sulfate fractions. Arguing for separation of the two ammonium sulfate fractions is the fact that the 0.43 g/mL fraction contains more heme, and it could be beneficial to purify separately in order to have a higher heme/protein ratio. However this ratio is not very high (discussed later) and heme reconstitution should be performed anyway. In order to save time mixing is probably better and Step2 is the best choice as the background at 280 nm is lowest for this method. With this procedure it might be possible to separate fractions depending on the amount of cofactor in the protein. According to Gardner, NO dioxygenase assays [33] ammonium sulfate precipitation could decrease heme content. Using another method for crude protein purification it might be possible to separate apo- and holo-FHb with the Step2-gradient.

3.1.3 Gel filtration

The last step of purification is gel filtration. An overview of gel filtration chromatograms from the different anion-exchange programs are found in the appendix, as they are quite similar. A small contamination of another protein can be observed by a shoulder before the main peak (figure 3.11). The shoulder size varies, from a quarter of the size (height) and less. Occasionally an additional peak is observed at around 120 mL, but it never overlaps with the collected fraction and is always quite small (Appendix 6). SDS-PAGE analysis of the collected fraction shows a very pure protein. The purity increases for each purification step. A weak band at around 35 kDa is observed on approximately half of the gels, depending on the amount of sample applied, (figure 3.12). Some additional bands are observed in this particular instance. Protein yield after purification is around ~20 mg / 30 g bacteria (~8 mg per 1L cell culture)

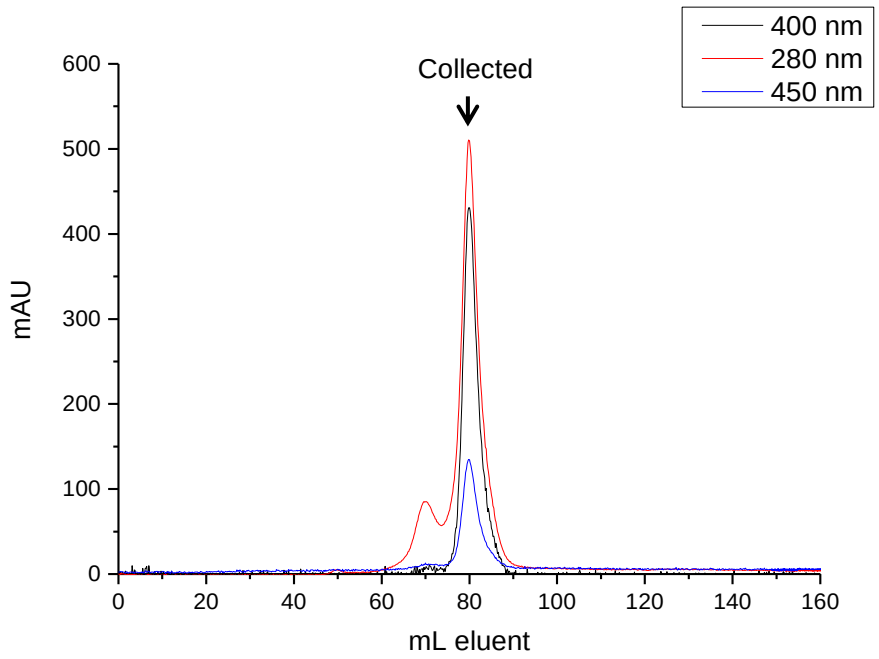


Figure 3.11: Chromatogram from gel filtration
Chromatogram of gel filtration. The black line, red line and blue line correspond to 400 nm, 280 nm and 450 nm, respectively. Protein with absorption at 400 nm is collected as indicated in the figure.

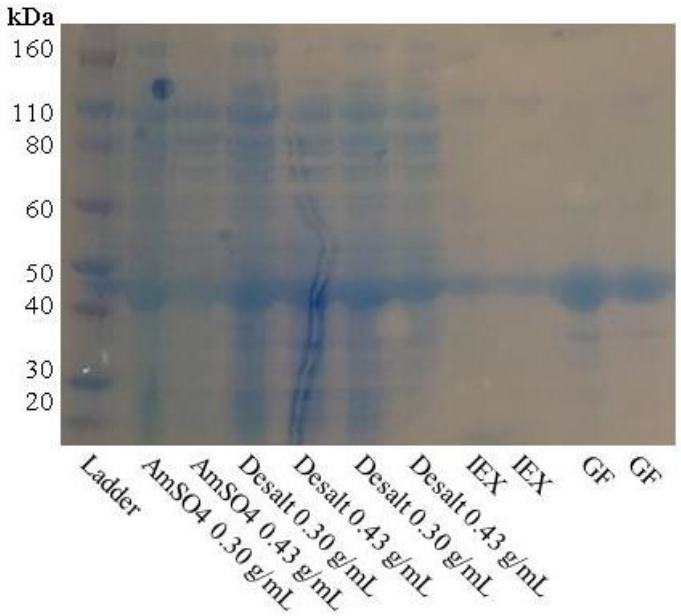


Figure 3.12: SDS-PAGE analysis of protein purification: The gel shows protein samples after ammonium sulfate precipitation, desalting, anion exchange experiments and a gel filtration step.

UV-vis spectroscopy

UV-vis spectroscopy was used to determine protein concentrations as described in the methods section. Calculated yield is found in table 3.2. The spectrophotometer was also used to give an indication of cofactor content, mainly heme.

3.1.4 Ammonium sulfate precipitation

Absorbance was measured after ammonium sulfate precipitation and desalting. The ratio between 405 nm (Soret peak) and 280 nm is found in Table 3.3. The results are from the expression with addition of ALA and riboflavin, and the ratio is expected to be higher for earlier purification experiments. When the ratio of purified protein is measured, the last expression with ALA and riboflavin has a lower 405 nm / 280 nm ratio, than previous purifications. The results show a ratio of 0.054 for the 0.30 g/mL pellet and 0.096 for the 0.43 g/mL-pellet indicating an almost twice the heme content, which is in agreement with the color of the solution (figure 3.3) and chromatograms from IEX (figure 3.4-3.9). The mix fraction has a ratio between the two pellets which is expected.

3.1.5 Gel filtration

The ratio between absorbance at 405 nm and 280 nm was also recorded after gel filtration as an indication of the amount of FHb containing the heme cofactor (table 3.3). The ratio shows a rather large variation and is found to be roughly between 0.15 and 0.70. As the 405 nm / 280 nm-ratio depends on both the initial cofactor content incorporated during the over-expression as well as the choice of purification procedures it is difficult to draw conclusions about the effect of using different gradients during IEX. The 0.43 g/mL sample has the highest ratio, which is expected. Similarly the 0.30 g/mL sample has a lower ratio. From table 3.3 it is clear that the mix-fractions have the lowest ratio. This is likely a consequence of a less successful over-expression with respect to cofactor incorporation. Induction and addition of cofactors were a bit late, and riboflavin was not dissolved properly at the time of addition.

Table 3.2: 405 nm / 280 nm ratio after desalt

Ratio	0.30 g/mL	0.43 g/mL	Mix1	Mix2
OD_{405}/OD_{280}	0.05423	0.09648	0.08317	0.07455

The table shows the ratio between 405 nm and 280 nm after ammonium sulfate precipitation and desalting. About half of the ammonium sulfate fractions were mixed before application to the desalting column, and had to be applied twice. The eluates are named Mix1 and Mix2.

Table 3.3: 405 nm / 280 nm ratio after gel filtration

$(NH_4)_2SO4$	ALA μM	Riboflavin μM	Gradient	Figure	405 nm / 280 nm
0.43 g/mL	0.2	-	Linear 0-30%	3.6	0.52-0.63
0.43 g/mL	0.2	-	Step1	3.6b	0.66
0.30 g/mL	0.2	-	Linear 0-50%	3.7b	0.37-0.38
0.30 g/mL	0.2	-	Step1	3.7a	0.44
Mix	0.25	0.30	Linear 0-30%	3.9	0.26
Mix	0.25	0.30	Step1	3.8a	0.16
Apo*	0.25	0.30	Step2	3.8b	0.05

The table shows the ratio between 405 nm and 280 nm after gel filtration with additives during over-expression, the ammonium sulfate fraction used, the gradient type during IEX and the applicable figure. *After IEX of a mix-fraction the 280 nm peak was collected in two fractions, the fraction consisting of apo-protein is shown.

3.1.6 FHb reduction by NADPH

UV-vis spectra of FHb after purification and after reduction by NADPH in GF-buffer (pH 7) under aerobic conditions were taken. After purification FHb is mainly in the ferric (Fe^{3+}) form, with a Soret peak at 405 nm, with additional smaller peaks at 542 nm and 572 nm. After addition of NADPH, the Soret peak moves to 411 nm, with clear peaks at 543 nm and 578 nm, respectively. The spectra are similar to that of flavohemoglobins from other organisms [18, 20-22, 34]. The reduced spectrum indicate that NADPH can reduce heme through FAD as an intermediate, like other FHbs and that it works as an electron donor. It would be interesting to see if reduction is similar with NADH. As oxygen is present it will likely bind to the reduced iron of the heme group as shown by the appearance of the peaks at 543 nm and 578 nm. A FAD peak is expected between 450 nm and 480 nm, but because of overlapping spectra with heme, it is difficult to observe, there is a slope change after reduction, which might be from FAD.

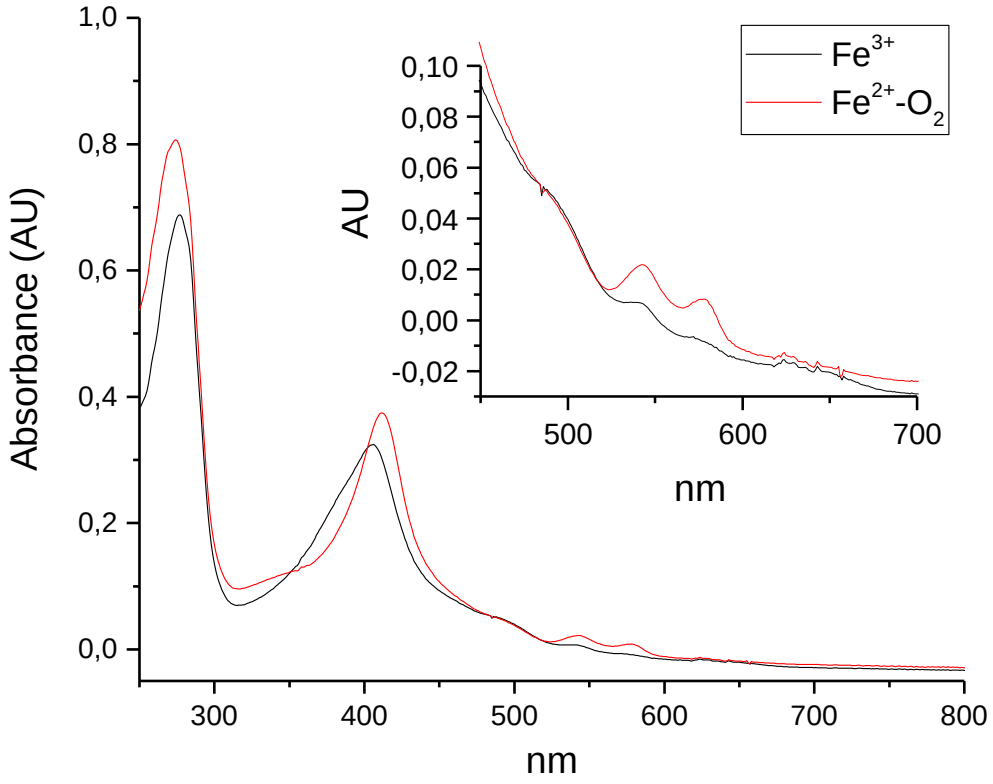


Figure 3.13: UV-vis spectroscopy of FHb: FHb in GF-buffer (pH = 7.5) is reduced by NADPH under aerobic conditions. The Black line is FHb after purification, in the oxidized state. The Red line shows FHb after reduction by NADPH with O_2 binding.

Assays

3.1.7 Protein Concentration

Bradford and Modified Lowry assay was performed, as described in the Methods section, in order to determine protein concentration of the purified FHb samples. In order to calculate protein concentrations of holo protein an extinction coefficient of 88.6 $mM^{-1}cm^{-1}$ for the Soret peak (405-406 nm) was used [34]. To indicate total concentration an extinction coefficient of 44.81 at 280 nm was used. The extinction coefficient was calculated from the protein sequence with Protaparam (Expasy) [38].

The Bradford and Modified Lowry give the total concentration of the FHb. Bovine Serum Albumin (GBiosciences) was used to generate the standard curves shown in figure 3.14. The overall shape of the Modified Lowry Assay curve does not seem to be linear in the

concentration range tested and so a quadratic fit was used. A linear fit was used with the Bradford assay. According to the manufacturer (Bio-Rad) the linear range of the assay is from 125 μg/mL to 1000 μg/mL. The linear fit is calculated for this range and the standard at 100 μg/mL is not taken into consideration since it is outside the linear range. Extrapolation is performed for the Bradford assay, with absorbance of FHb samples ~ 0.63 (absorbance of 1.0 mg/mL BSA = 0.60). The calculated protein concentrations are shown in Table 3.5.

As both protein assays were performed with the same sample concentration there is a rather large discrepancy between the two methods with calculated concentrations of ~1.04 mg/mL and ~0.59 mg/mL for the Bradford and Modified Lowry assay, respectively. Both methods are affected by the standard used, (e.g. BSA or gamma globulin) and the color response varies with the proteins assayed [37, 40, 41]. In addition as extrapolation was performed for the Bradford assay a higher uncertainty is expected.

Table 3.4: Fit for protein assays

Assay	Fit	a	b	c	R^2
Bradford	Linear	0.12358	0.48642	-	0.98164
Lowry	Quadratic	0.09875	2.77648	-1.47847	0.99512

The fit for the standard is expressed in the form: $f(x) = a + bx + cx^2$. Absorbance is expressed by x (concentration in mg/mL) and the constants (a, b, c and d).

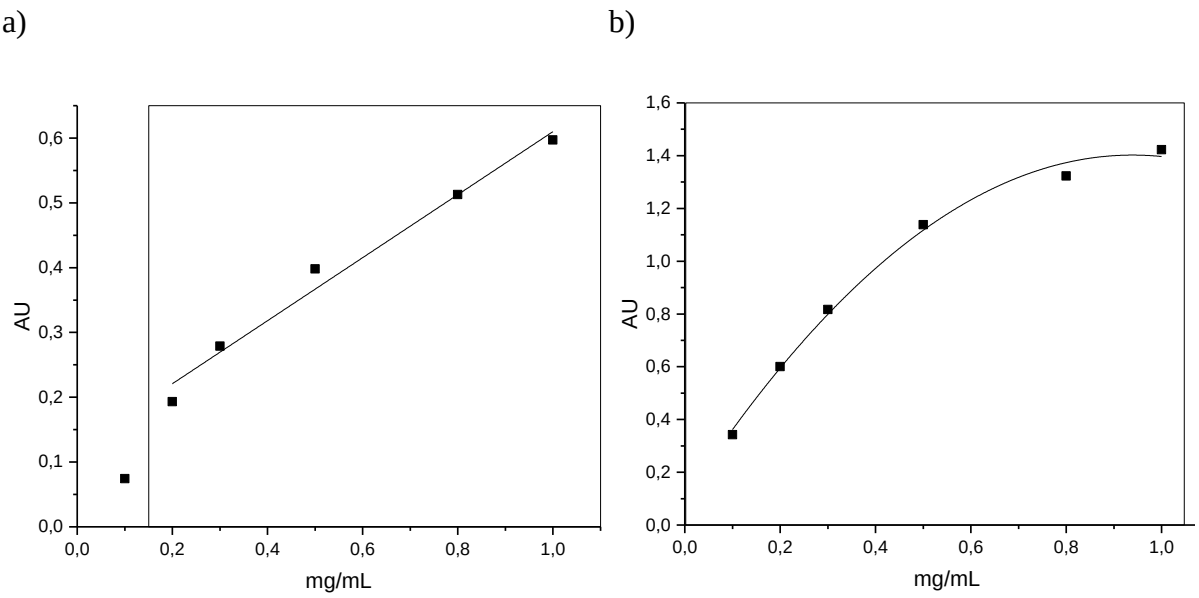


Figure 3.14: Bradford assay and Modified Lowry protein assays. Bradford assay (a) and Modified Lowry assay (b) was performed as described in Methods with a bovine serum albumin (BSA) standard. The squares represent the BSA standard with the fits shown. A linear fit is used for the Bradford Assay and a quadratic fit for the Modified Lowry assay. Only the squares within the box are used for fitting.

Table 3.5: Calculated concentrations for the Bradford assay and Modified Lowry assay

Assay	**Fit / ε**	**Average** mg/mL	**Sample 1** mg/mL	**Sample 2** mg/mL	**Sample 3** mg/mL
Bradford	Linear fit	1.04	1.04	1.02	1.06
Lowry	Quadratic fit	0.59	0.62	0.59	0.56

By comparing the protein concentration measured by the protein assays with absorbance of FHb in solution at 280 nm (measured with a UV-vis spectrophotometer) the following extinction coefficients were calculated, $\varepsilon_{280} = 70.36\ mM^{-1}/cm^{-1}$ and $\varepsilon_{280} = 124.02\ mM^{-1}/cm^{-1}$ for the Bradford assay and Modified Lowry assay, respectively. There is a rather large difference between the two extinction coefficients, also compared to the extinction coefficient calculated by ProtParam ($44.81\ mM^{-1}/cm^{-1}$) [38]. FAD absorbs light at 280 nm which is not taken into account by the ProtParam tool, so the extinction coefficient is expected to be too low. The discrepancy between extinction coefficients calculated by the two protein assays could be due to variations of color response for the assays. The Modified Lowry assay could be more accurate [40], in addition extrapolation was performed to calculate protein concentration for the Bradford assay causing a higher uncertainty.

3.1.8 Heme Assay

The fraction of FHb with heme and FAD varies among organisms and purification runs, although the ratio between concentrations of FHb and heme can be 1:1 [42]. As heme absorbs light at ~405 nm, the ratio between absorbance at 405 nm and 280 nm (OD_{405} / OD_{280}) can indicate heme content. Takaya *et. al.* reported a ratio of 0.85 with 0.5 mole heme and 1 mole FAD per mole FHb [18]. Rafferty *et. al.* reported a ratio of 1.05 with 0.65 mole heme and 1 mole FAD per mole FHb [22].

Heme assays were performed to quantify heme concentrations of FHb samples, and assess the amount of mole heme/mole protein (figure 3.15). For the assay, FHb is reduced with dithionite and reoxidized with ferricyanide. Heme concentrations in the samples are calculated from the absorbance differences of the reduced minus oxidized heme at 556 nm and 539 nm.

μM heme is given by the following formula: $\mu M\ heme = 46.7\ (\Delta A556 - \Delta A539)$ [33, 35]

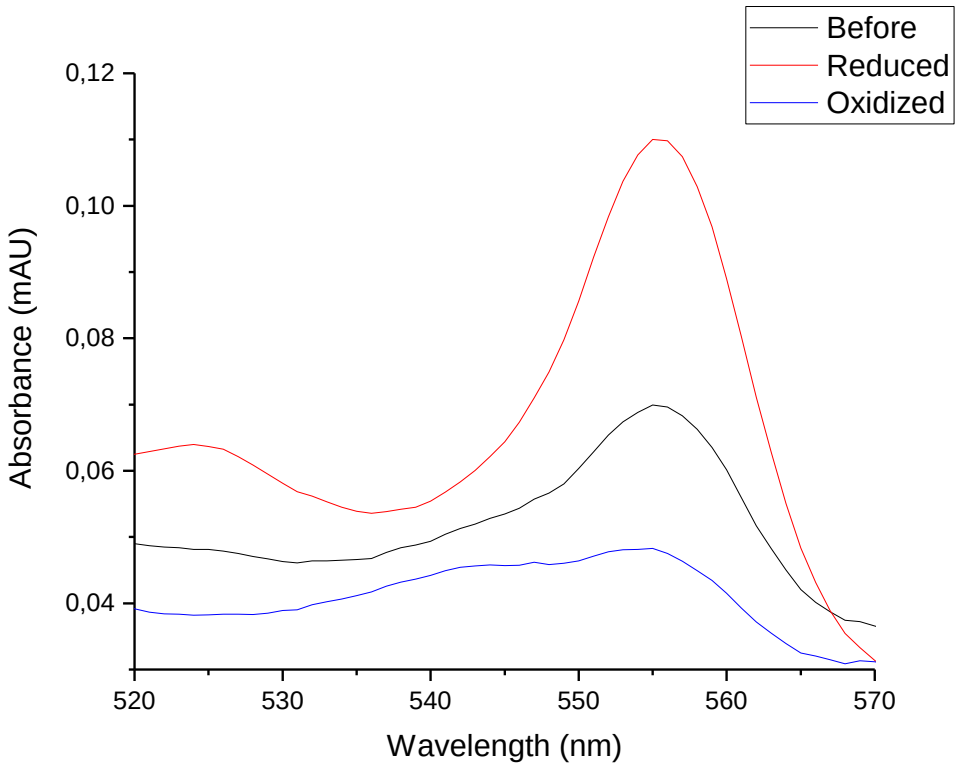


Figure 3.15: A typical UV-vis spectrum for a heme assay. Heme assay performed as described in Methods. Heme concentration is calculated by the formula: μM heme = 46.7 (ΔA556 – ΔA539) The spectra are background corrected at 700 nm (700 nm = 0).

The extinction coefficients for the Soret peak of FHbs from three organisms have been reported, which differ by a factor of 2: ε_{403} = 88.6 $mM^{-1}cm^{-1}$ for *E. coli*, [34] ε_{395} = 123 $mM^{-1}cm^{-1}$ for *R. eutropha* [43] and ε_{395} = 63 $mM^{-1}cm^{-1}$ for *S. cerevisiae* [24]. An extinction coefficient of 88.6 $mM^{-1}cm^{-1}$ has been used to calculate concentration of FHb with heme. The Soret peak of *B. cereus* FHb is at 405 nm, and *E. coli* has the Soret peak at the most similar wavelength (403 nm).

It is worth to note that the FHb assayed contained a quite low 405 nm / 280 nm ratio (Table 3.6). The heme assay gives a low concentration of heme compared to the concentration calculated from absorbance measurements. The heme fraction calculated from absorbance varies from 0.13 to 0.36. If the ratio between OD_{405}/OD_{280} and heme content, is similar to that reported from Takaya *et.al.* [18] and Rafferty *et.al.* [22], then the expected heme fraction is 0.15. This can indicate that the heme assay gives an underestimation of the heme content, and further studies have to be performed to validate the method.

Table 3.6: Results from heme assay

ε_{280}	ε_{280} mM^{-1}/cm^{-1}	[FHb] μM	[FHb-Heme]* μM	[Heme] μM	Heme fraction* (ε_{405})	Heme fraction (assay)	OD_{405}/OD_{280}
ProtParam	44.81	36.45	4.74	1.14	0.13	0.032	0.26
Bradford	70.36	23.02			0,21	0.050	
Lowry	124.02	13.17			0,36	0.087	
$Assays_{avg}$	89.78	18.18			0,26	0.063	

Concentrations of total protein [FHb] and heme fractions have been calculated with different extinction coefficients at 280 nm. $Assays_{avg}$ is calculated from the average values of the protein assays.
*Calculated with an extinction coefficient of 88.6 $mM^{-1}cm^{-1}$ [34].

The FAD-content of FHb can vary, [33] and in order to measure FAD-content FHb was boiled and absorbance measured, but due to high background no conclusions can be drawn. However FAD content can be measured by fluorescence [33].

Reconstitution

Reconstitution attempts were made by incubating the partly apo-FHb with either heme, FAD or both cofactors (figure 3.17-3.19). To separate the excess cofactor from the reconstituted FHb sample, an initial attempt was made by applying the sample to a desalting column as it is a crude size-exclusion column. However, only a single peak was observed, so the column was unable to separate protein from heme and other reactants, results not shown. Therefore, a Superdex 200 gel filtration column was used

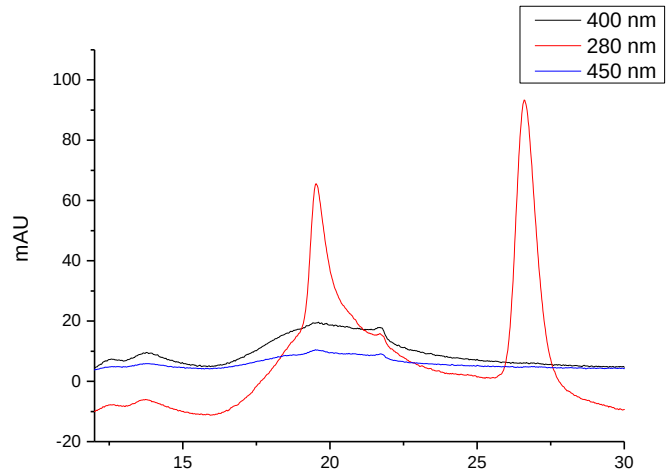
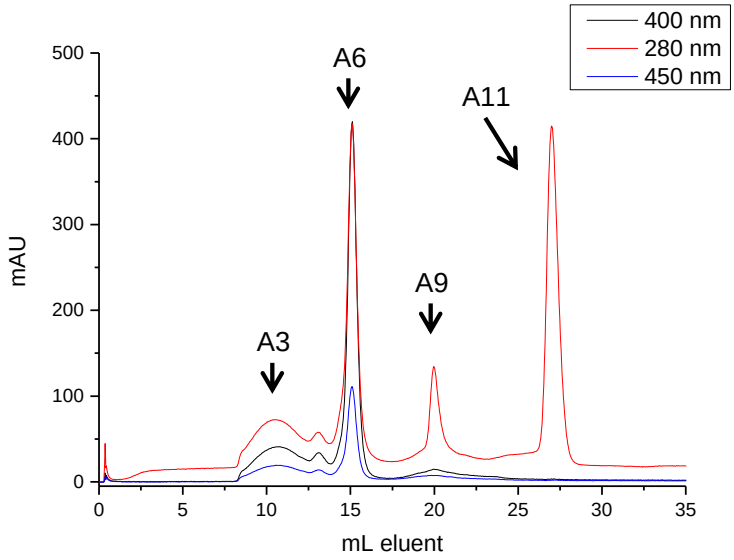


Figure 3.16: Reconstitution mixture applied on Superdex column: a) Solution used for reconstitution applied on a Superdex-200 GL column without FHb present. The black line, red line and blue line correspond to 400 nm, 280 nm and 450 nm, respectively.

a)



b)

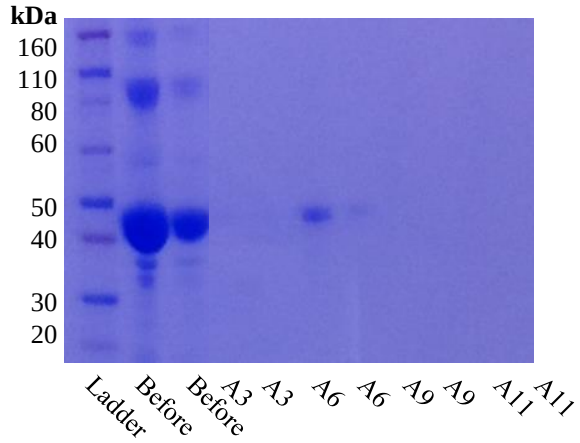


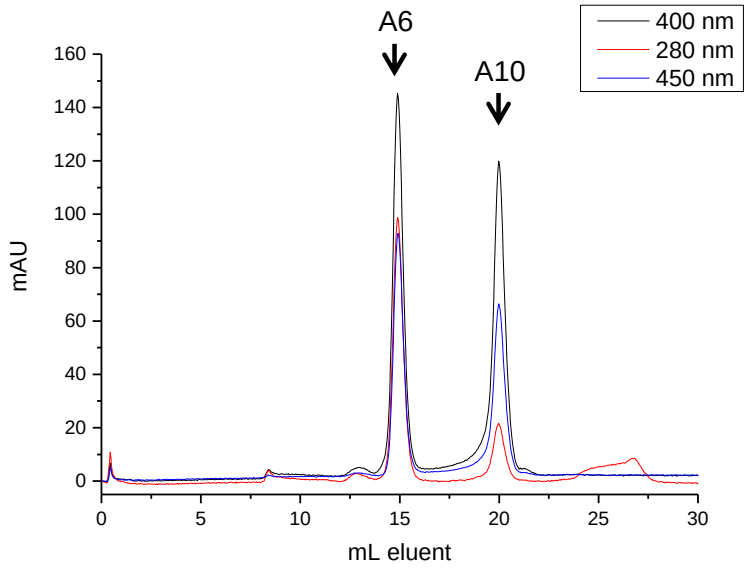
Figure 3.17: Reconstitution with heme: a) Chromatogram of reconstitution with heme. The black line, red line and blue line correspond to 400 nm, 280 nm and 450 nm, respectively. Four peaks are observed at 10 ml, 15 mL, 20 mL and 27 mL. All peaks were collected and SDS-PAGE analysis was performed (b) as A3, A6, A9 and A11 respectively.

Reconstitution with heme (figure 3.17) resulted in 4 peaks in the chromatogram and the peak at 15 mL is expected to be FHb as it shows a band at the expected size on the SDS-PAGE. The peaks at 20 mL and 27 mL are expected to be heme, as they are present when a heme solution is applied to the Superdex column, figure 3.16. The peak at 10 mL could be aggregates of FHb as filtration was not performed before application to the column, although no band is present in the SDS-gel. A UV-vis spectrum of the fraction however shows similar characteristics as FHb and the peak at 15 mL, but with a lower absorbance at 405 nm. A small peak at 13 mL might be catalase. The reconstitution clearly increased the heme content of the FHb as can be seen by the increase in 405 nm / 280 nm ratio from 0.72 to 0.99, from 0.38 to 0.70 and 0.54 to 0.70 (Table 3.7). Therefore the reconstitution procedure of heme seemed to work. If compared to 405 nm / 280 nm ratios from other FHbs [18, 22], the heme content is expected to be 0.59 and 0.42 for the 405 nm / 280 nm ratios of 0.99 and 0.70, respectively.

The Chromatogram from reconstitution with FAD (figure 3.18) also shows separation of protein from FAD with peaks at 15 mL and 20 mL, respectively. This is confirmed by SDS-PAGE analysis as only the first peak contains a protein band. The tiny peak at 13 mL might be catalase. Table 3.7 shows an increase in absorbance at 450 nm, although lower than for heme which is expected as FAD has a lower extinction coefficient than heme at this wavelength ($\varepsilon_{450} \sim 11.2\ mM^{-1}cm^{-1}$) [44]. FAD absorbs some light at 405 nm as well.

As reconstitution with both heme and FAD separately indicate successful reconstitution a trial of reconstitution with heme and FAD by adding both cofactors to the same tube was tested. The chromatogram (figure 3.19) is a mixture of the former reconstitutions experiments with the expected peaks, in addition to a shoulder at 13 mL. The change in 405 nm / 280 nm ratio is similar as for reconstitution with heme alone with an additional increase in 450 nm / 280 nm ratio, especially for the A7 peak. It seems like reconstitution of heme and FAD can be performed by adding both cofactors at the same time. An SDS-PAGE analysis should be performed, in addition to further testing to possibly increase the amount of holo-FHb further.

a)



b)

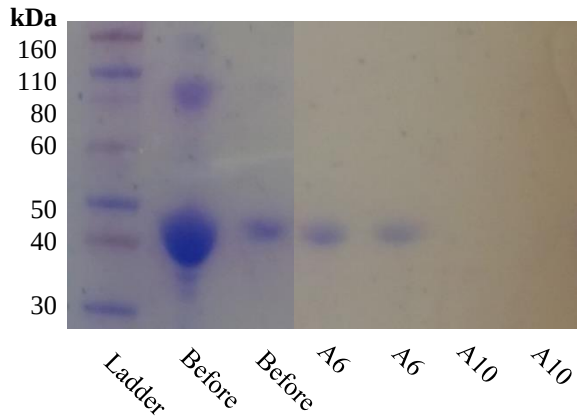


Figure 3.18: Reconstitution with FAD: a) Chromatogram of reconstitution with FAD, The black line, red line and blue line correspond to 400 nm, 280 nm and 450 nm, respectively. Two peaks are observed at 15 mL, 20 mL and 27 mL. Both peaks were collected and SDS-PAGE analysis was performed (b) as A6 and A10, respectively.

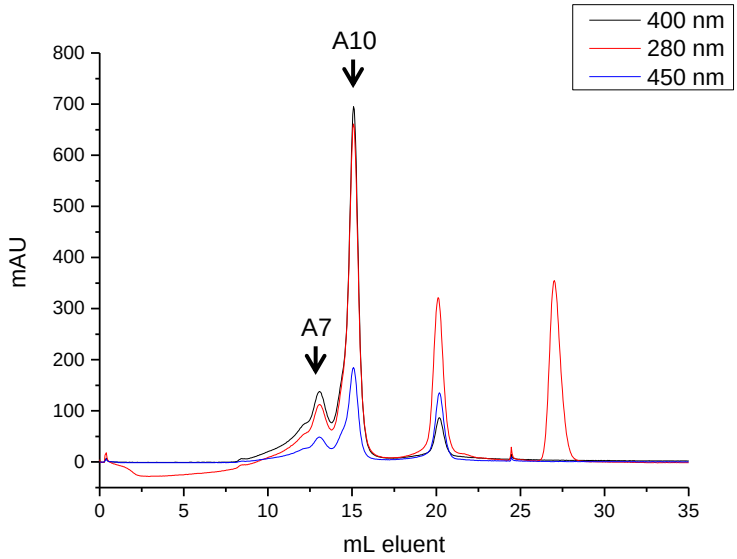


Figure 3.19: Reconstitution of heme and FAD at the same time: Chromatogram of reconstitution with heme and FAD. The black line, red line and blue line correspond to 400 nm, 280 nm and 450 nm, respectively. Four peaks are observed at 13 mL, 15 mL, 20 mL and 27 mL. The first two are collected as indicated on the figure.

Table 3.7: 400 nm/280 nm

Reconstitution	Peak	Before		After	
		OD_{405}/OD_{280}	OD_{450}/OD_{280}	OD_{405}/OD_{280}	OD_{450}/OD_{280}
Heme	A6	0.72	0.15	0.99	0.25
Heme	A6	0.38	0.09	0.70	0.21
Heme*,**	A6	0.54	0.17	0.70	0.31
FAD**	A6	0.69	0.15	0.75	0.20
Heme+FAD**	A7	0.72	0.16	1.05	0.39
	A10			0.99	0.28

*Performed at 4°C, *Performed with 100 μM FHb instead of 20 mM

Sequence alignment

A sequence alignment of *B. cereus* FHb (figure 3.31) was performed with flavohemoglobins from *E. coli*, *R. Eutropha* and *S. cerevisiae* with sequence identities of 46.72, 46.63 and 44.53, respectively. The organisms listed all have crystal structure in the Protein data bank (PDB) [17, 24, 25]. The conserved residues mentioned in the introduction are all present in *B. Cereus* FHb. The FHb sequence alignment in addition to spectroscopy and SDS-PAGE analysis confirm that the purified protein is a member of the FHb family.

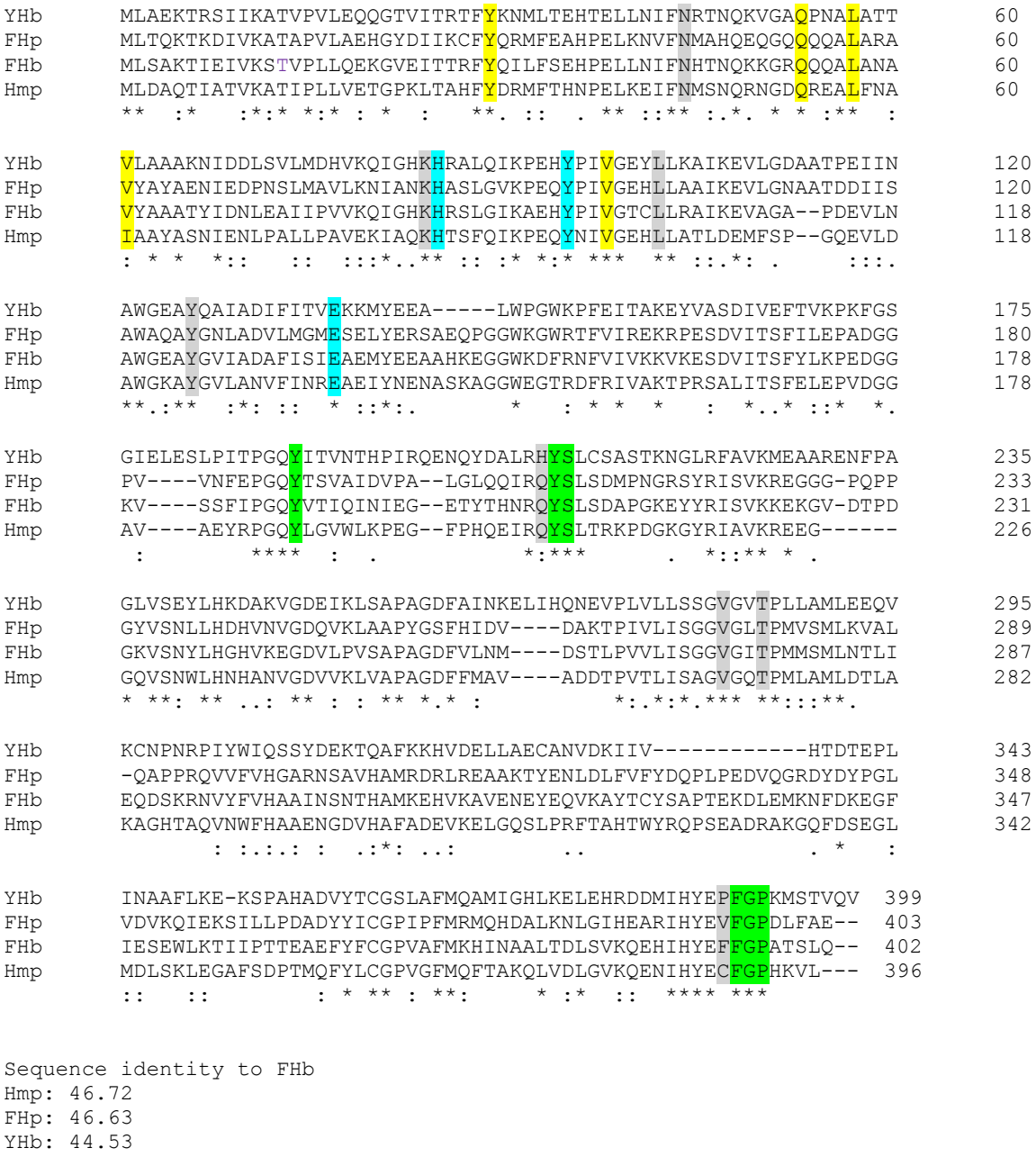


Figure 3.20: Sequence alignment of FHb. Sequence alignment of flavohemoglobin from *B. cereus* (FHb), *S. cerevisiae* (YHb), *E. coli* (Hmp), and *R. Eutropha* (FHp) performed by Clustal Omega [45]. An "*" (asterisk) indicates positions which have a single, fully conserved residue. A ":" (colon) and a "." (period) indicate conservation between groups of strongly similar properties and weakly similar properties, respectively.

Amino acids either surrounding the active site and/or are known to play a role in catalysis are marked in yellow. Amino acids in cyan are believed to constitute the push-and-pull mechanism. Side chains responsible for FAD binding are marked in green [13]. In addition amino acids otherwise mentioned in the introduction are marked in gray.

Protein crystallization

For crystallization screening a range of different commercial crystallization kits were used (Table 3.8). Very small red or orange needles grew from amorphous precipitate under various conditions (figure 3.21) (Table 3.8). Optimization of crystallization conditions were set out manually under a light microscope. None of the initial crystals/needles were reproduced during optimization. FHb precipitated during mixing of the drops for crystallization which might be due to light sensitivity of the flavin as crystallization was performed under a light microscope. Flavins exhibit some light-sensitivity and are susceptible to photo-degradation, although FAD is less sensitive than the other flavins (by an order of magnitude). Regardless Nobre *et al.* and Zhou *et al.* performed purification of FHb protected from light and estimated mole heme per mole protein to be ~1 and 0.89, respectively [20, 21].

In addition the ratio between OD_{400} and OD_{280} varied between each purification. A high purity and low apo-content (homogenous solution) is important for crystal growth. The relatively high apo-content could be one of the reasons for the limited success of the crystallization screening. To obtain a sample with higher holo-content, a successful reconstitution is one possibility, as well as being able to separate the apo- and holo-FHb during purification.

A lack of diffracting crystals resulted in no X-ray data and consequently no crystal structure.

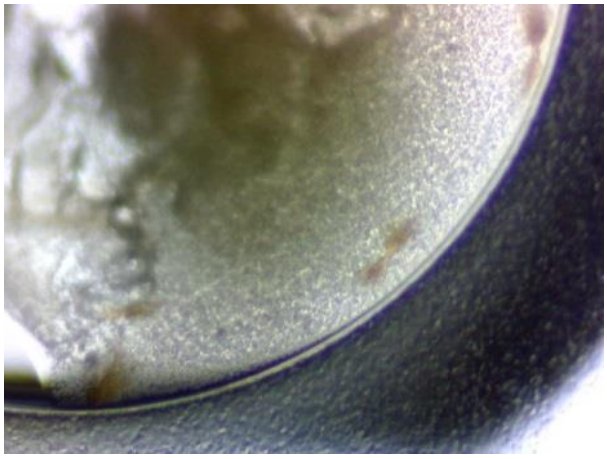


Figure 3.21: Small needles grown from precipitate. The picture of needles grown from precipitate is taken with a light microscope.

Table 3.8: Commercial crystallization screens used for initial screening.

Screen name	Manufacturer	Room Temperature	4°C
Crystal Screen 1+2	Hampton Research	C2, F10	-
Hampton Index I+II	Hampton Research	X	-
JCSG+	Molecular Dimensions	A9*****, B9*****, D7, H4, H6, H7*****	D7
MIDAS	Molecular Dimension	A5, A6	X
Morpheus	Molecular Dimensions	X	-
Natrix	Hampton Research	X	-
Wizard I + II	Jena Bioscience	E3	-

Wells with small needles are named while screens with no crystals are marked with an "X".

Table 3.9: Conditions with small needles grown from precipitate

Screen name	Well	Contents
Crystal Screen 1 + 2	C2	0.2 M Ammonium acetate, 0.1 M sodium citrate tribasic dehydrate pH 5.6 30% v/v MPD
	F10	0.1 M MES monohydrate pH 6.5 12% (w\v) PEG 20 000
JCSG+	A9	0.2 Ammonium chloride pH 6.3 20% (w\v) PEG 3350
	B9	0.1 M Citric acid pH 5, 20% (w\v) PEG 6000
	D7	0.2 M Lithium sulfate, 0.1 M Tris pH 8.5, 40% (w\v) PEG 400
	H4	0.2 M Calcium chloride, 0.1 M Bis-Tris pH 5,5, 45% (w\v) MPD
	H6	0.1 M Ammonium acetate, 0.1 M Bis-Tris pH 5.5, 17% PEG (w\v) 10 000
	H7	0.2 mM Ammonium sulfate, 0.1 M Bis-Tris pH 5.5, 25% (w\v) PEG 3350
MIDAS	A5	0.5 M ammonium phosphate monobasic, 12.5% (w\v) poly(acrylic acid sodium salt) 2100
	A6	19% (v/v) poly(acrylic acid-co-maleic acid), 0.1 M Tris pH 8.5
Wizard I + II	E3	20% (w/v) PEG 8000, 0.1 M Tris pH 8.5; 0.2 M $MgCl_2$

Conclusion and future perspectives

The focus of this master project was to purify, characterize and obtain a crystal structure of flavohemoglobin (FHb) from *B. Cereus*. FHb is an NO-dioxygenase which is proposed to protect against nitric oxide by dioxygenation of NO to nitrate (NO_3^-). As well as being a potential therapeutic target [13] it can be a useful tool for probing mammalian nitric oxide biology [16].

The protein was successfully cloned and expressed and a purification protocol has been developed, with ammonium sulfate precipitation, anion exchange chromatography and gel filtration. A high yield with high purity was obtained, but with varying degrees of heme and FAD content. Crystallization screening resulted in only some small needles grown from precipitate, and as reproduction or optimization of crystallization failed, no crystal structure was obtained.

With varying degrees of apo-protein, heme assay was performed in order to identify heme to protein ratio. Bradford assay and Modified Lowry assay was used to obtain the protein concentration and calculate extinction coefficients. Testing of reconstitution of heme and FAD was performed with promising results. Both heme and FAD content was clearly increased, in addition reconstitution of both cofactors can be performed at the same time.

Future work includes reconstitution, more crystallization screening and further characterization. Purification protected from light should be tested as well.

In summary, this thesis presents an expression and purification protocol and the first results of characterization and reconstitution of the flavohemoglobin protein from *Bacillus Cereus*.

4 Appendices

Appendix 1 – Terms and abbreviations

(v/v)	(volume/volume)
(w/v)	(weight/volume)
Amp	Ampicillin
B. cereus	*Bacillus cereus*
BSA	Bovine serum albumin
DMSO	Dimethyl sulfoxide
DNA	Deoxyribonucleic acid
DTT	Dithiothreitol
E. coli	*Escherichia coli*
EDTA	Ethylene diamine tetraacetic acid
EtOH	Ethanol
FAD	Oxidized Flavin Adenine Dinucleotide
$FADH_2$	Reduced Flavin Adenine Dinucleotide
FAD_{SQ}	Semiquinone state of Flavin Adenine Dinucleotide
FHb	Flavohemoglobin
hAC	Acetic Acid
HEPES	N-2-hydroxyethylpiperazine-N'-2-ethanosulfonic acid
H_2O	water

IPTG	Isopropyl β-D-1-Thiogalactopyranoside
LB	Lysogeny broth
MOPS	3-(*N*-morpholino)propanosulfonic acid
MPD	2-methyl-2,4-pentanediol
NAD+	Oxidized Nicotinamide adenine dinucleotide
NADH	Reduced Nicotinamide adenine dinucleotide
NADP+	Oxidized Nicotinamide adenine dinucleotide phosphate
NADPH	Reduced Nicotinamide adenine dinucleotide phosphate
N_2	Nitrogen gas
PAGE	Polyacrylamide gel electrophoresis
PDB	Protein Data Bank
PEG	Polyethylene glycol
R. eutorpha	*Ralstonia eutropha*, formerly *Alcaligenes eutrophus*
S. cerevisiae	*Saccharomyces cerevisiae*
SDS	Sodium dodecyl sulfate
TB	Terrific broth
Tris	Tris(hydroxymethyl) aminomethane
UV	Ultraviolet
UV-vis	Ultraviolet-visible light

Appendix 2 – Materials

Chemicals

Chemical	Manufacturer
Acetic acid	Merck
δ-aminoluvelinic acid (ALA)	Sigma
Ammonia	Kebo Lab
Ammonium chloride	Sigma / Merck
Ammonium sulfate	Merck
Ampicillin sodium salt	Sigma
Bacto agar	Becton, Dickinson & Co.
Bovine Serum Albumin standard, prediluted	GBiosciences
Bradford Assay Dye Reagent	Bio-Rad laboratories
Brilliant Blue R250	Sigma
Catalase	Sigma
Citric acid	Fluka
Deconex (Detergent)	VWR
DTT	V'WR
Ethylenediaminetetraacetic acid (EDTA)	Sigma
Ex-Cell Antifoam	Sigma
Dimethyl sulfoxide (DMSO)	Sigma

Dithiothreitol (DTT)	Sigma
Ethanol	Arcus
FAD	Sigma
Glycerol	VWR
Hydrochloric acid ≥ 37%	Sigma
Hemin	Fluka
HEPES	Sigma
MOPS SDS-Running buffer	Noves®
Mucosit®-P Disinfecting Powder Cleaner	Sigma
NADPH	Applichem
NuPAGE ® Bis-Tris Mini Gels (10 & 15 wells)	Novex®
InstantBlue coomasie staining	CBS
Isopropyl β-D-1-thiolgalactopyranoside (IPTG)	Thermo
PEG (3350, 6000 and 8000)	Sigma
Potassium chloride	Merck
Potassium ferricyanide	Aldrich
Dipotassium hydrogen phosphate (K_2HPO_4)	Merck
Potassium dihydrogenphosphate (KH_2PO_4)	Merck
Protease pill (cOMplete ULTRA tablets, mini Easypack)	Roche
Pyridine	Merck
Riboflavin (Vitamin B_2)	Sigma

Sodium dithionite	Fluka
Sodium citrate	Merck
Streptomycin sulfate salt	Sigma
Tris(hydroxymethyl)aminomethane (Tris)	VWR
Tryptone	Sigma
Yeast Extract	Merck

Chromatographic column materials

Material	**Manufacturer**
Hitrap Desalting	GE Healthcare
Sephadex 25	GE Healthcare
HPQ Sepharose	GE Healthcare
Superdex 200 10/30 GL	GE Healthcare
Superdex 200	GE Healthcare

Hardware

Material	Function	Manufacturer
Äkta purifier system	Purifier stystem	GE Healthcare
JA-10 rotor	Rotor	Beckman Coulter
JA-25.50 rotor	Rotor	Beckman Coulter
pH meter (PHM240)	pH-meter	Radiometer Analytical
MililQ plus	MilliQ-water	Millipore
Agilent 8453 Spectrophotmeter	UV-vis spectrophotometer	Agilent Technologies
Series 25 Incubator shaker	Incubator	New Brunswich Sc.
Tomy SS-325	Autoclave	Tomy
Emmer autoklav	Autoclave	Emmer
Dri-Block DB-2D	Heat block	Techne
WB4MS	Water Bath	Biosan
REAX 2000	Vortex	Heidolph
Mosquito LCP	Crystallization robot	TTP Labtech
Oryx6 Robot	Crystallization robot	Douglas Inst. Ltd.
Avanti J-25 Centrifuge	Large centrifuge	Beckman
Avanti J-20XP Centriguge	Large centrifuge	Beckman Coulter
BioCen20	Table centrifuge	Orto Alresa

Software

Material	Source
845 x UV visible Chem Station	Agilent Technologies
Unicorn 531 (build 743) 2011 Äkta	GE Healthcare
Unicorn 407 (build 743) 2006 Äkta	GE Healthcare

Equipment

Material	Manufacturer
Amicon Ultra (30 K)	Millipore
Amicon Ultra 4 Centrifugal filters	Millipore
Amicon Plus-70 Centrifugal filters (30 K)	Millipore
Centrifugation tubes	Nalgene
Crystallization plates	Qiagen
Cuvettes (Quartz)	Hellmma
Disposable cuvettes	Plastibrand
Microlance 3	BD
PCR-Tubes	Axygen
Pipette tips	VWR
Precision syringes	Hamilton
Sterile 0.22 μm filter	Sarstedt

Crystallization Kits

Screen name	Manufacturer
Crystal Screen	Hampton Research
Hampton Index I+II	Hampton Research
JCSG+	Qiagen / Molecular Dimensions
MIDAS	Molecular Dimensions
Morpheus	Molecular Dimensions
Natrix	Hampton Research
Proplex	Molecular Dimensions
Wizard I + II	Emerald Biosystems / Rigaku Reagents

Appendix 3 – Media and solutions

4.1.1 Media

LB-medium (1 L)

10 g NaCl

10 g tryptone

5 g yeast extract

1 L MilliQ-water

Adjusted to pH 7.5 and autoclaved

Phosphate Buffer

0.17 M KH_2-phosphate

0.72 M K_2H-phosphate

pH 7.5

autoclaved

TB-medium (1 L)

12 g tryptone

25 g yeast extract

0.9 L MilliQ-water

4 mL glycerol *dissolved in water*

0.1 L phosphate buffer, pH 7.5 *autoclaved separately*

Autoclaved without glycerol and phosphate buffer

4.1.2 Buffers

Buffer A	**Buffer B**
50 mM Tris-Cl, pH 7.5	1 M KCl
50 mM Tris-Cl, pH 7.5	2 mM DTT
2 mM DTT	

Buffer C	**Buffer D**
100 mM HEPES, pH 7.5	50 mM Tris-Cl, pH 8
100 mM KCl	1 mM EDTA
2 mM DTT	10 mM DTT
	3000 U catalase

Appendix 4 – Sequences

Amino acid sequence of flavohemoglobin from *B. cereus* (402 aa)

```
MLSAKTIEIV KSTVPLLQEK GVEITTRFYQ ILFSEHPELL NIFNHTNQKK   50
GRQQQALANA VYAAATYIDN LEAIIPVVKQ IGHKHRSLGI KAEHYPIVGT   100
CLLRAIKEVA GAPDEVLNAW GEAYGVIADA FISIEAEMYE EAAHKEGGWK   150
DFRNFVIVKK VKESDVITSF YLKPEDGGKV SSFIPGQYVT IQINIEGETY   200
THNRQYSLSD APGKEYYRIS VKKEKGVDTP DGKVSNYLHG HVKEGDVLPV   250
SAPAGDFVLN MDSTLPVVLI SGGVGITPMM SMLNTLIEQD SKRNVYFVHA   300
AINSNTHAMK EHVKAVENEY EQVKAYTCYS APTEKDLEMK NFDKEGFIES   350
EWLKTIIPTT EAEFYFCGPV AFMKHINAAL TDLSVKQEHI HYEFFGPATS   400
LQ                                                       402
```

DNA sequence of flavohemoglobin from *B. cereus* (1209 bp for gene and 1257 bp total)

TCTAGAAATAATTTTGTTTAACTTTAAGAAGGAGATATACAT**ATG**TTAAGTGCAAAAACAAT
TGAAATCGTAAAGTCAACAGTACCATTATTACAAGAAAAAGGTGTTGAAATTACAACGAGAT
TCTATCAAATTTTATTTTCGGAACATCCGGAATTGTTGAATATTTTCAACCATACGAATCAG
AAAAAGGGAAGACAACAACAAGCGTTAGCGAATGCTGTTTATGCAGCTGCAACGTACATTGA
TAATTTAGAAGCAATTATTCCAGTTGTAAAACAAATTGGTCATAAGCATAGAAGTTTAGGGA
TTAAAGCTGAGCATTATCCGATTGTAGGTACATGTTTACTACGTGCCATTAAAGAGGTCGCA
GGTGCACCTGATGAAGTTTTAAATGCATGGGGAGAAGCATATGGTGTAATTGCTGATGCATT
CATTAGCATTGAAGCAGAGATGTATGAAGAAGCTGCACATAAAGAAGGTGGATGGAAAGACT
TCCGCAACTTTGTGATTGTAAAAAAAGTGAAGGAAAGCGATGTTATTACGTCATTTTATTTG
AAACCTGAAGATGGAGGGAAAGTTTCTTCATTCATCCCAGGTCAATATGTAACAATTCAAAT
CAATATTGAAGGTGAGACATATACACATAATCGTCAATACAGTTTATCCGATGCTCCTGGGA
AAGAATATTATCGTATTAGTGTAAAAAAAGAAAAAGGTGTAGATACACCAGACGGTAAAGTG
TCTAATTACTTACATGGACATGTAAAAGAAGGAGATGTTTTACCAGTAAGTGCACCAGCGGG
AGATTTCGTGTTAAATATGGATTCAACATTACCAGTTGTACTAATTAGTGGTGGAGTGGGGA
TTACACCGATGATGAGTATGTTAAATACGTTAATTGAACAAGATTCAAAACGTAATGTATAT
TTTGTTCATGCAGCAATAAATAGTAATACACATGCAATGAAAGAACACGTTAAGGCAGTAGA
AAATGAATATGAACAAGTTAAAGCATATACTTGTTATTCTGCACCGACTGAAAAAGATTTAG
AAATGAAGAACTTTGATAAAGAAGGTTTCATTGAAAGTGAATGGTTAAAAACTATTATTCCG
ACAACTGAAGCAGAGTTCTATTTCTGTGGTCCAGTAGCATTTATGAAGCATATAAATGCTGC
ACTAACTGATTTAAGTGTGAAACAAGAGCATATTCATTATGAATTTTTCGGCCCAGCAACAA
GCTTACAA**TAA**GGATCC

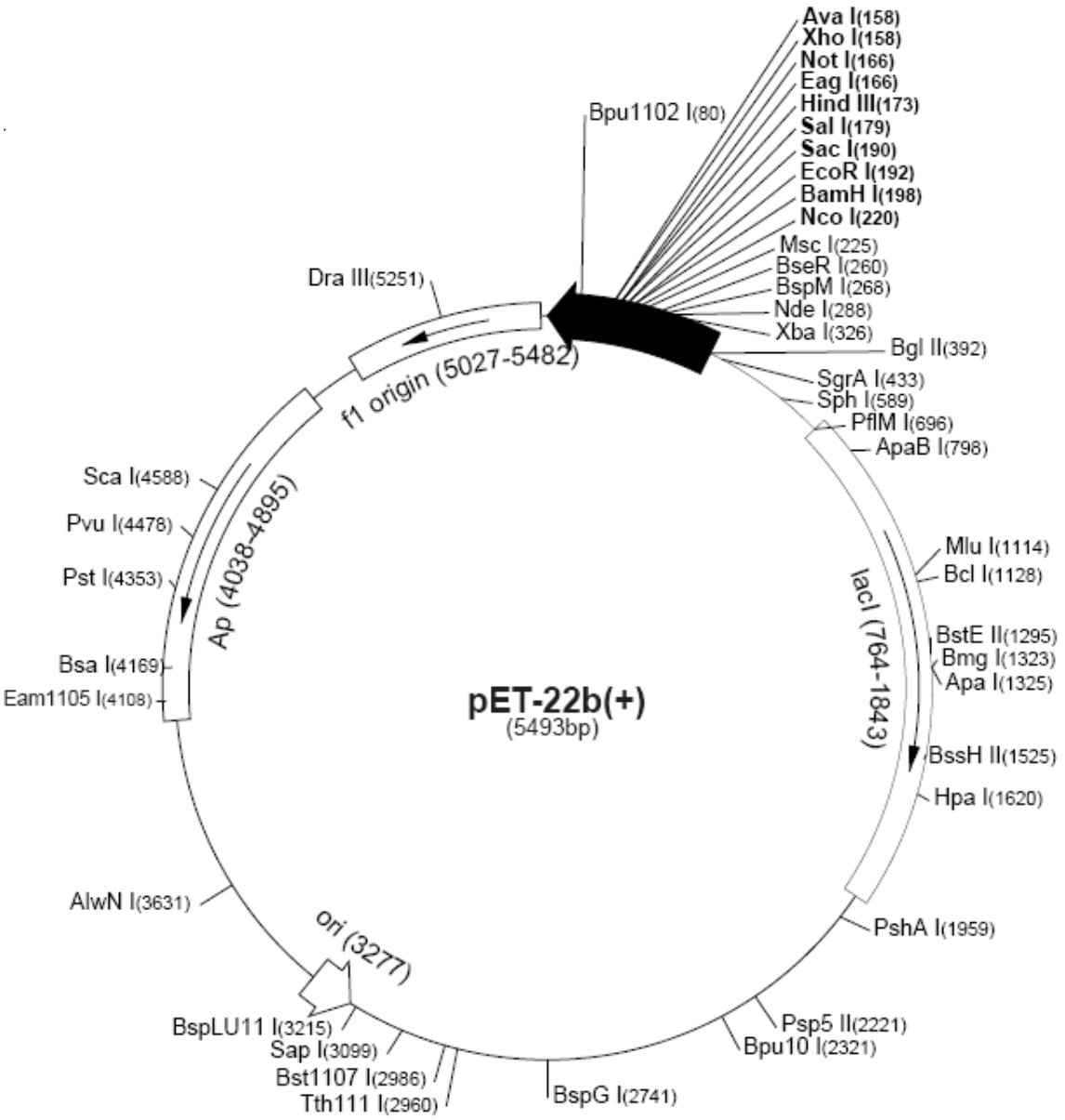
Start and stop codon are marked in bold, Restriction sites for cutting are underlined and a part of the original vector sequence is shown in blue.

The restriction enzymes have the following restriction sites:

XbaI:	TCTAGA	used for cutting
BamHI:	GGATCC	used for cutting
NdeI:	CATATG	often used by the group

NdeI was not used as a restriction enzyme as a restriction site is present within the protein sequence. In order to make minimal changes to the vector, the additional bases of the vector between the XbaI and NdeI were added.

Appendix 5 – Vector map



Appendix 6 – Gel filtration chromatograms

An overview of the chromatograms in this section is found in the table below. Gel filtration has been performed as described in the Methods section.

The figures names are given in this format:

Figure “Letter”, “ammonium sulfate fraction”, “anion exchange program used”

For the figures, the black line, red line and blue line correspond to 400 nm, 280 nm and 450 nm, respectively. Protein with absorption at 400 nm is collected as indicated on the figures.

Modified Table 3.1 Overview of anion-exchange experiments performed

Fraction	Gradient type	Gradient	IEX-chromatogram	GF
0.43 g/mL	Linear	0-100%	Fig. 3.4a	-
	Linear	0-30%	Fig 3.6a	Figure A
	Step	Step1	Fig 3.6b	Figure B
0.30 g/mL	Linear	0-100%	Fig 3.4b	-
	Linear	0-50%	Fig 3.7b	Figure C
	Step	Step1	Fig 3.7a	Figure D
Mix	Linear	0-30%	Fig. 3.9	-
	Step	Step1	Fig. 3.8a	Figure E
	Step	Step2	Fig. 3.8b	Figure F
Apo*	Step	Step2	Fig. 3.8b	Figure G

*Collected from second half of the peak in figure 3.8b with very low absorbance at 400 nm.

Figure A, 0.43 g/mL, 0-30%

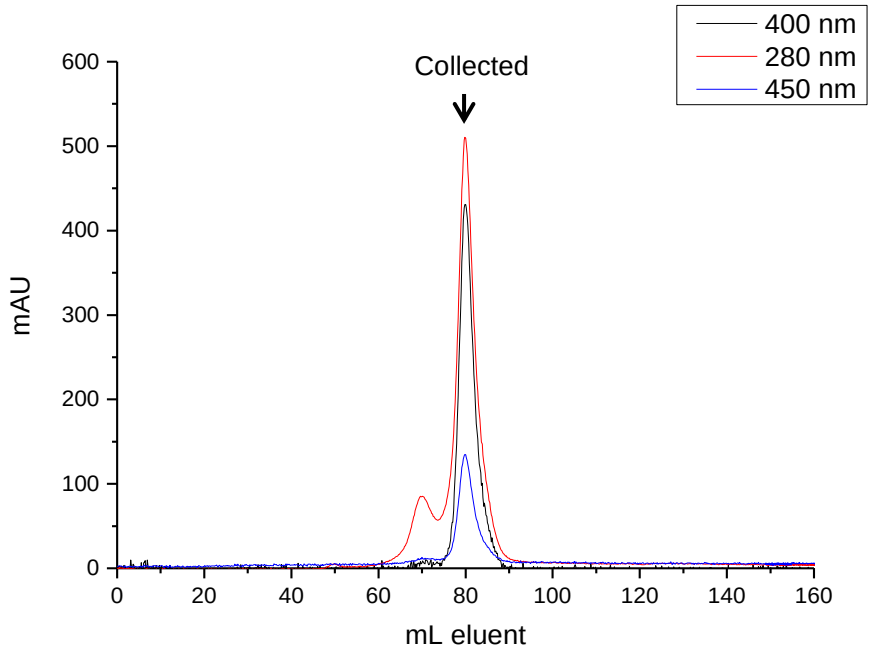


Figure B, 0.43 g/mL, Step1

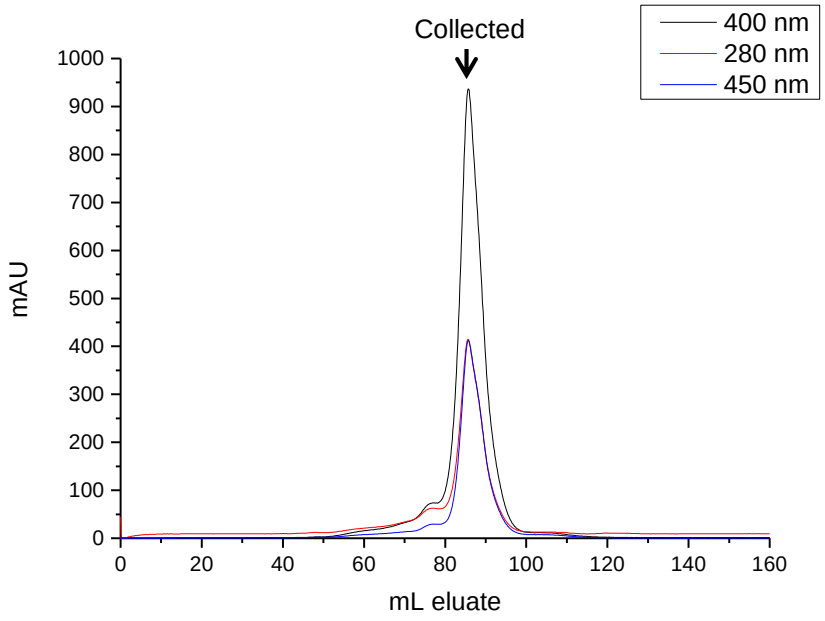


Figure C, 0.30 g/mL, 0-50%

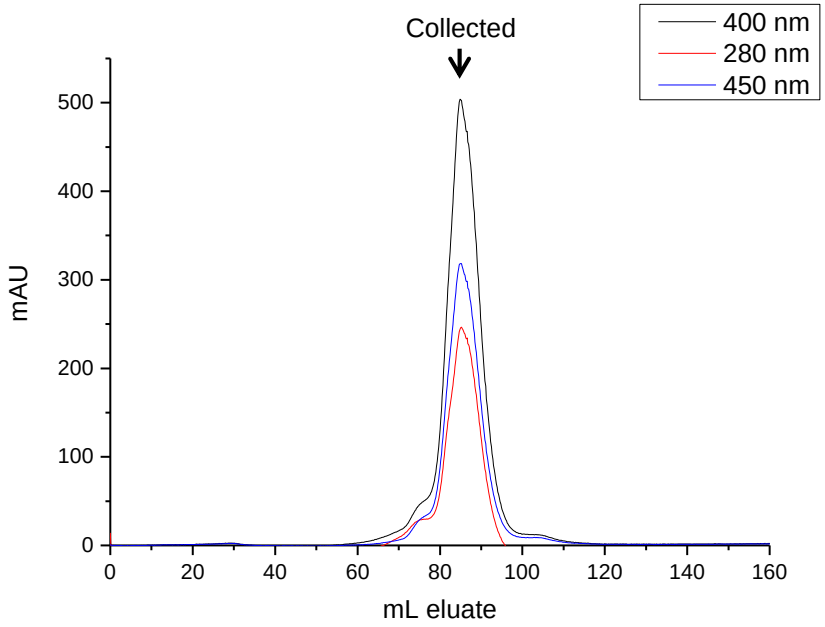


Figure D, 0.30 g/mL, Step1

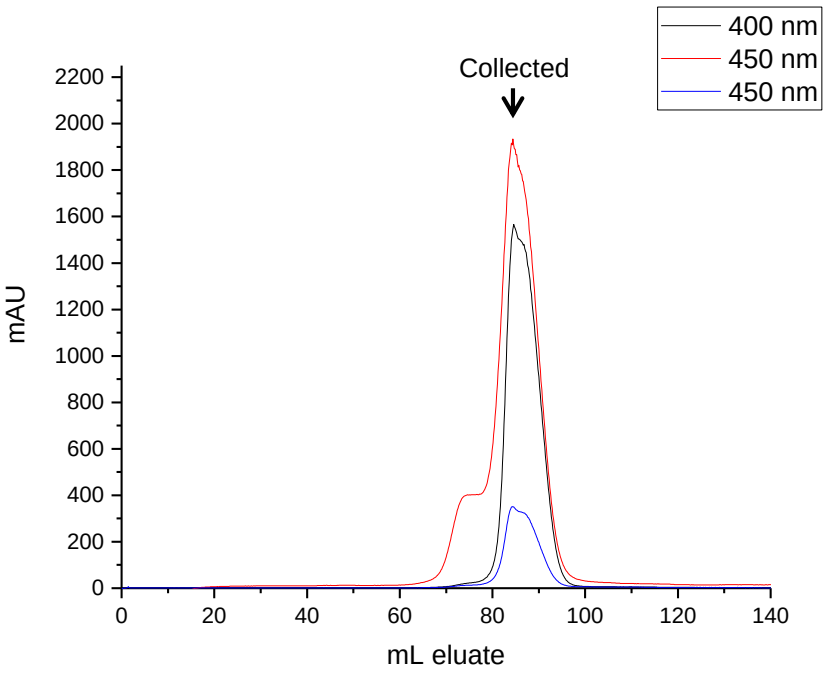


Figure E, Mix, Step1

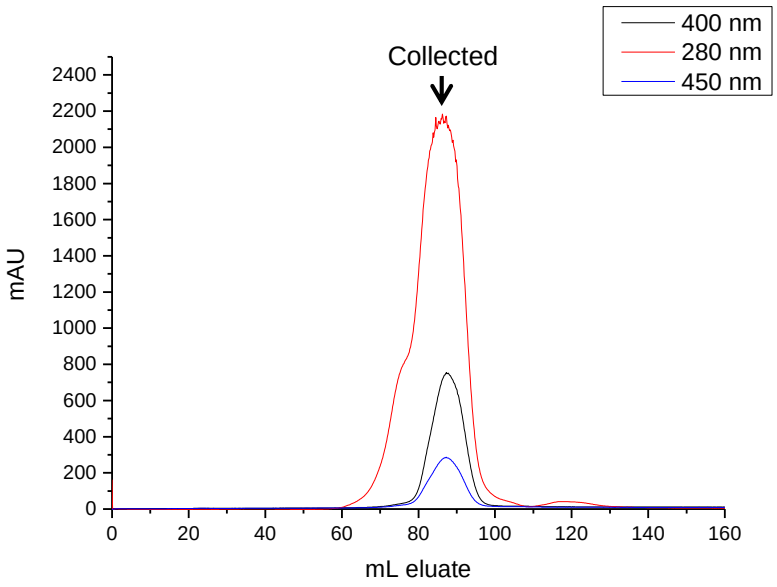


Figure F, Mix, Step2

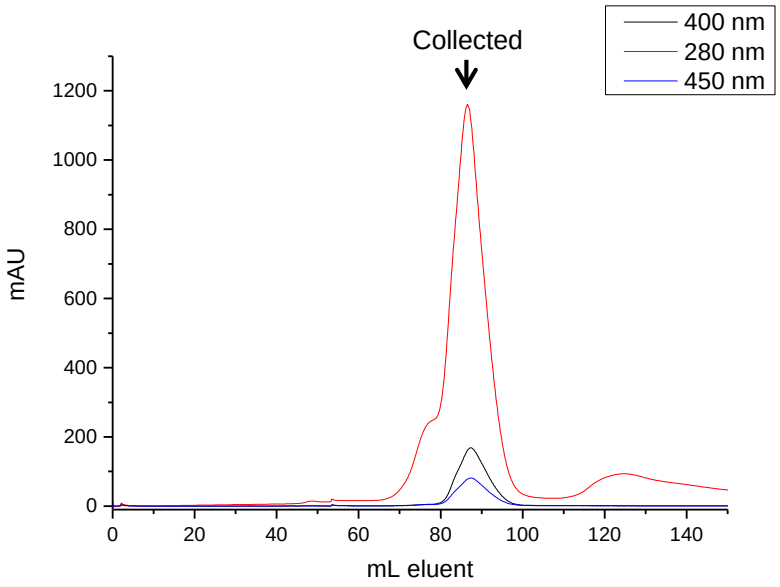
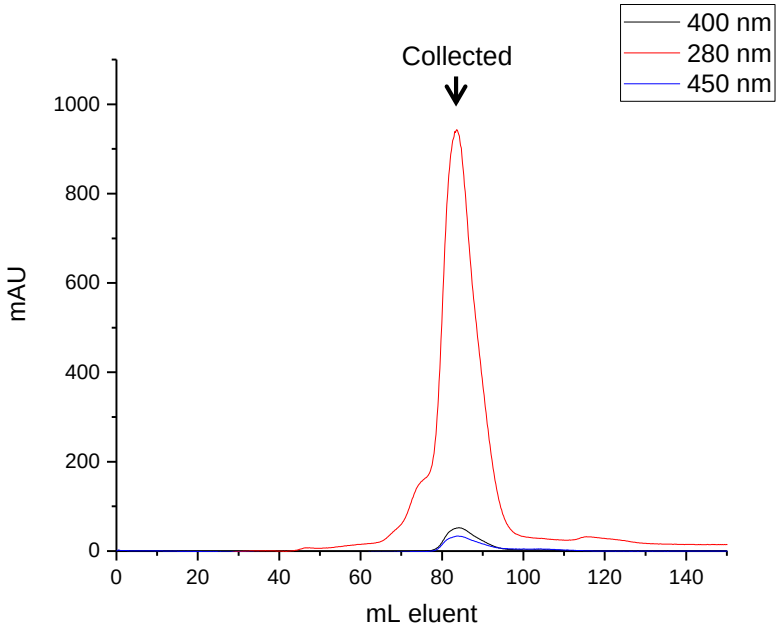


Figure G, Mix (Apo*), Step2



5 References

1. Zhang, L., *Introduction*, in *Heme Biology: The Secret Life Of Heme In Regulating Diverse Biological Processes*, L. Zhang, Editor. 2011, World Scientific Publishing CO PTE LTD.
2. Bali, S., Palmer, D.J., Schroeder, S., Ferguson, S.J., and Warren, M.J., *Recent advances in the biosynthesis of modified tetrapyrroles: the discovery of an alternative pathway for the formation of heme and heme d 1.* Cellular and Molecular Life Sciences, 2014. **71**(15): p. 2837-63.
3. Hersleth, H.-P., Ryde, U., Rydberg, P., Görbitz, C.H., and Andersson, K.K., *Structures of the high-valent metal-ion haem–oxygen intermediates in peroxidases, oxygenases and catalases.* Journal of Inorganic Biochemistry, 2006. **100**(4): p. 460-476.
4. Macheroux, P., Kappes, B., and Ealick, S.E., *Flavogenomics - a genomic and structural view of flavin-dependent proteins.* FEBS Journal, 2011. **278**(15): p. 2625-2634.
5. Edwards, M.A., ed. *Structure and General Properties of Flavins*. Flavins and Flavoproteins: Methods and Protocols, ed. S. Weber and E. Schleicher. Vol. 1146. 2014, Humana Press Inc, 999 Riverview Dr, Ste 208, Totowa, Nj 07512-1165 USA. 3-13.
6. Cronk, J.D. *Flavin adenine dinucleotide (FAD)*. [cited 2014 27.10]; Available from: http://guweb2.gonzaga.edu/faculty/cronk/biochem/F-index.cfm?definition=FAD.
7. Nelson, D.L. and Cox, M.M., *Lehninger Principles of Biochemistry*. 5 ed. 2008.
8. Stern, A.M. and Zhu, J., *An Introduction to Nitric Oxide Sensing and Response in Bacteria*, in *Advances in Applied Microbiology, Vol 87*, S. Sariaslani and G.M. Gadd, Editors. 2014, Elsevier Academic Press Inc: San Diego. p. 187-220.
9. Hill, B.G., Dranka, B.P., Bailey, S.M., Lancaster, J.R., Jr., and Darley-Usmar, V.M., *What part of NO don't you understand? Some answers to the cardinal questions in nitric oxide biology.* Journal of Biological Chemistry, 2010. **285**(26): p. 19699-704.
10. Foundation, T.N. 1998; Available from: http://www.nobelprize.org/nobel_prizes/medicine/laureates/1998/.
11. Campbell, M.G., Smith, B.C., Potter, C.S., Carragher, B., and Marletta, M.A., *Molecular architecture of mammalian nitric oxide synthases.* Proc Natl Acad Sci U S A, 2014. **111**(35): p. E3614-23.
12. Cheng, H., Wang, L., Mollica, M., Re, A.T., Wu, S., and Zuo, L., *Nitric oxide in cancer metastasis.* Cancer Letters, 2014. **353**(1): p. 1-7.
13. Forrester, M.T. and Foster, M.W., *Protection from nitrosative stress: A central role for microbial flavohemoglobin.* Free Radical Biology and Medicine, 2012. **52**(9): p. 1620-1633.
14. Zeier, J., Delledonne, M., Mishina, T., Severi, E., Sonoda, M., and Lamb, C., *Genetic elucidation of nitric oxide signaling in incompatible plant-pathogen interactions.* Plant Physiology, 2004. **136**(1): p. 2875-86.
15. Mishina, T.E., Lamb, C., and Zeier, J., *Expression of a nitric oxide degrading enzyme induces a senescence programme in Arabidopsis.* Plant, cell & environment, 2007. **30**(1): p. 39-52.
16. Forrester, M.T., Eyler, C.E., and Rich, J.N., *Bacterial flavohemoglobin: a molecular tool to probe mammalian nitric oxide biology.* Biotechniques, 2011. **50**(1): p. 41-5.

17. Ermler, U., Siddiqui, R.A., Cramm, R., and Friedrich, B., *Crystal structure of the flavohemoglobin from Alcaligenes eutrophus at 1.75 A resolution.* The EMBO journal., 1995. **14**(24): p. 6067-77.
18. Takaya, N., Suzuki, S., Matsuo, M., and Shoun, H., *Purification and characterization of a flavohemoglobin from the denitrifying fungus Fusarium oxysporum.* FEBS Letters, 1997. **414**(3): p. 545-548.
19. Gardner, P.R., Gardner, A.M., Martin, L.A., and Salzman, A.L., *Nitric oxide dioxygenase: An enzymic function for flavohemoglobin.* Proceedings of the National Academy of Sciences, 1998. **95**(18): p. 10378-10383.
20. Nobre, L.S., Gonçalves, V.L., and Saraiva, L.M., *Chapter Eleven - Flavohemoglobin of Staphylococcus aureus*, in *Methods in enzymology*, K.P. Robert, Editor. 2008, Academic Press. p. 203-216.
21. Zhou, S., Fushinobu, S., Nakanishi, Y., Kim, S.-W., Wakagi, T., and Shoun, H., *Cloning and characterization of two flavohemoglobins from Aspergillus oryzae.* Biochemical and Biophysical Research Communications, 2009. **381**(1): p. 7-11.
22. Rafferty, S., Luu, B., March, R.E., and Yee, J., *Giardia lamblia encodes a functional flavohemoglobin.* Biochemical and Biophysical Research Communications, 2010. **399**(3): p. 347-351.
23. Gupta, S., Pawaria, S., Lu, C., Hade, M.D., Singh, C., Yeh, S.-R., and Dikshit, K.L., *An Unconventional Hexacoordinated Flavohemoglobin from Mycobacterium tuberculosis.* Journal of Biological Chemistry, 2012. **287**(20): p. 16435-16446.
24. El Hammi, E., Warkentin, E., Demmer, U., Marzouki, N.M., Ermler, U., and Baciou, L., *Active site analysis of yeast flavohemoglobin based on its structure with a small ligand or econazole.* FEBS Journal, 2012. **279**(24): p. 4565-4575.
25. Ilari, A., Bonamore, A., Farina, A., Johnson, K.A., and Boffi, A., *The X-ray Structure of Ferric Escherichia coliFlavohemoglobin Reveals an Unexpected Geometry of the Distal Heme Pocket.* Journal of Biological Chemistry, 2002. **277**(26): p. 23725-23732.
26. Grisham, C.M. and Garrett, R.H., *Biochemistry*. 3 ed. 2004: Brooks / Cole Publishing Co.
27. Bonamore, A. and Boffi, A., *Flavohemoglobin: Structure and reactivity.* IUBMB Life, 2008. **60**(1): p. 19-28.
28. Mowat, C.G., Gazur, B., Campbell, L.P., and Chapman, S.K., *Flavin-containing heme enzymes.* Archives of Biochemistry and Biophysics, 2010. **493**(1): p. 37-52.
29. Ferreiro, D.N., Boechi, L., Estrin, D.A., and Martí, M.A., *The key role of water in the dioxygenase function of Escherichia coli flavohemoglobin.* Journal of Inorganic Biochemistry, 2013. **119**(0): p. 75-84.
30. Muhlig, A., Kabisch, J., Pichner, R., Scherer, S., and Muller-Herbst, S., *Contribution of the NO-detoxifying enzymes HmpA, NorV and NrfA to nitrosative stress protection of Salmonella Typhimurium in raw sausages.* Food Microbiology, 2014. **42**: p. 26-33.
31. Nakano, M.M., *Essential role of flavohemoglobin in long-term anaerobic survival of Bacillus subtilis.* Journal of Bacteriology, 2006. **188**(17): p. 6415-8.
32. Edebo, L., *A new press for the disruption of micro-organisms and other cells.* Journal of Biochemical and Microbiological Technology and Engineering, 1960. **2**(4): p. 453-479.
33. Gardner, P.R., *Chapter Twelve - Assay and Characterization of the NO Dioxygenase Activity of Flavohemoglobins*, in *Methods in enzymology*, K.P. Robert, Editor. 2008, Academic Press. p. 217-237.
34. Helmick, R.A., Fletcher, A.E., Gardner, A.M., Gessner, C.R., Hvitved, A.N., Gustin, M.C., and Gardner, P.R., *Imidazole Antibiotics Inhibit the Nitric Oxide Dioxygenase*

Function of Microbial Flavohemoglobin. Antimicrobial Agents and Chemotherapy, 2005. **49**(5): p. 1837-1843.
35. Appleby, C.A., *Purification of Rhizobium cytochromes P-450.* Methods in enzymology, 1978. **52**: p. 157-66.
36. Berry, E.A. and Trumpower, B.L., *Simultaneous determination of hemes a, b, and c from pyridine hemochrome spectra.* Analytical Biochemistry, 1987. **161**(1): p. 1-15.
37. Bradford, M.M., *A rapid and sensitive method for the quantitation of microgram quantities of protein utilizing the principle of protein-dye binding.* Analytical Biochemistry, 1976. **72**: p. 248-54.
38. Gasteiger, E., Hoogland, C., Gattiker, A., Duvaud, S.e., Wilkins, M., Appel, R., and Bairoch, A., *Protein Identification and Analysis Tools on the ExPASy Server*, in *The Proteomics Protocols Handbook*, J. Walker, Editor. 2005, Humana Press. p. 571-607.
39. Thermo. *Instructuions, Modified Lowry Protein Assay Kit*. [Instruction manual] 2011; Available from: https://www.funakoshi.co.jp/data/datasheet/PCC/23240.pdf.
40. Thermo, *Thermo Scientific pierce Protein Assay Technical Handbook*. 2010: Thermo Scientific. p. 44.
41. Lowry, O.H., Rosebrough, N.J., Farr, A.L., and Randall, R.J., *Protein measurement with the Folin phenol reagent.* Journal of Biological Chemistry, 1951. **193**(1): p. 265-75.
42. Gardner, A.M., Martin, L.A., Gardner, P.R., Dou, Y., and Olson, J.S., *Steady-state and Transient Kinetics of Escherichia coli Nitric-oxide Dioxygenase (Flavohemoglobin): The B10 Tyrosine Hydroxyl is essential for dioxygen binding and catalysis.* Journal of Biological Chemistry, 2000. **275**(17): p. 12581-12589.
43. El Hammi, E., Warkentin, E., Demmer, U., Limam, F.r., Marzouki, N.M., Ermler, U., and Baciou, L., *Structure of Ralstonia eutropha Flavohemoglobin in Complex with Three Antibiotic Azole Compounds.* Biochemistry, 2011. **50**(7): p. 1255-1264.
44. Ghisla, S., Massey, V., Lhoste, J.M., and Mayhew, S.G., *Fluorescence and Optical Characteristics of Reduced Flavines and Flavoproteins.* Biochemistry, 1974. **13**(3): p. 589-597.
45. Sievers, F., Wilm, A., Dineen, D., Gibson, T.J., Karplus, K., Li, W., Lopez, R., McWilliam, H., Remmert, M., Söding, J., Thompson, J.D., and Higgins, D.G., *Fast, scalable generation of high-quality protein multiple sequence alignments using Clustal Omega.* Molecular systems biology, 2011. **7**: p. 539.