Phytoglobins: Elucidation of Structural, Functional and Practical Features of Non-Symbiotic Plant Hemoglobins

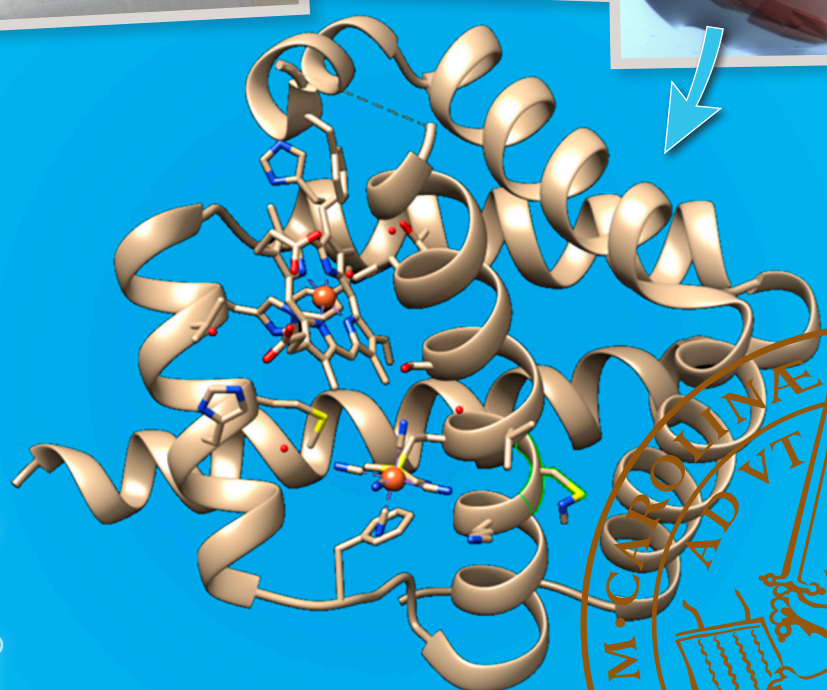
Christensen, Simon

2023

Document Version:
Publisher's PDF, also known as Version of record

Phytoglobins: Elucidation of Structural, Functional and Practical Features of Non-Symbiotic Plant Hemoglobins

SIMON CHRISTENSEN
DEPARTMENT OF CHEMISTRY | FACULTY OF ENGINEERING | LUND UNIVERSITY



Phytoglobins: Elucidation of Structural, Functional and Practical Features of Non-Symbiotic Plant Hemoglobins

Simon Christensen

LUNDS UNIVERSITET

DOCTORAL DISSERTATION

by due permission of the Faculty of Engineering, Lund University, Sweden. To be publicly defended in Lecture Hall B, Kemicentrum, Naturvetarvägen 14, Lund. Wednesday 7th of June 2023 at 9:00.

Faculty opponent
Prof. Hans Bäumler

Institute of Transfusion Medicine, Charité Universitätsmedizin, Berlin

Organization
LUND UNIVERSITY
Division of Pure and Applied Biochemistry
Deparment of Chemistry
P.O Box 124
SE-22100 Lund, Sweden
Author: Simon Christensen

Document name
DOCTORAL DISSERTAION
Date of issue:
7th of June 2023
Sponsoring organization

Title: Phytoglobins; Elucidation of Structural, Functional and Practical Features of Non-Symbiotic Plant Hemoglobins

Abstract

Phytoglobins (Pgb) (plant hemoglobins) are heme-containing proteins and less well known in comparison to the mammalian counterparts hemoglobin (Hb) and myoglobin (Mb). They share the same secondary structure as other globin-related proteins, six to eight α-helices comprising the myoglobin-fold (Mb-fold) with one prosthetic heme moiety per subunit. Pgbs are divided into different classes, where the non-symbiotic Pgbs (nsPgb) are found in vast number of plants. Unlike their mammalian counterparts, nsPgbs usually have an additional coordination site interacting with the iron called the distal histidine, in addition to the proximal histidine below the heme plane. Even though they bind oxygen (O_2), the proposed functions of these proteins are linked to nitric oxide (NO) metabolism, redox balance and energy maintenance, especially in hypoxic conditions, to a higher extent than oxygen transport.

One aim of this thesis was to generate mutant(s) with altered characteristics using site-directed mutagenesis and investigate the role of these residues. One residue in particular was a conserved cysteine residue located at position 86 (C86A) in sugar beet (*Beta vulgaris* ssp. *vulgaris*) nsPgb BvPgb1.2. After crystallization and structure determination, this cysteine-to-alanine substitution had no impact on the tertiary and quaternary structure but major implications regarding functionality were observed. The mutant had faster autoxidation rate and increased thermal stability, while the wild-type (WT) protein showed higher peroxidase activity. Both species appeared as homodimers and did not show any heme loss, unlike human hemoglobin.

To gain more information regarding the intra- and intermolecular interactions in nsPgbs, triple-labeled BvPgb1.2 was expressed, purified and analyzed using 2D Nuclear Magnetic Resonance (NMR). 83% of the expected amide cross-peaks were assigned. The majority of the non-assigned residues were located in G and H α-helices, which are proposed to be an important area of dimer interactions. Hydrogen bonding between specific residues (T53 and E123) and a hydrophobic cluster in opposing monomers were suggested to be important in dimer formation. Structure predictions were conducted to gain knowledge regarding the role of the conserved cysteine in dimerization and mapping the lost residues in the crystal structure. These residues were located at highly flexible regions in the N- and C-termini, as well as in the loops for helices CD/DE. Small angle X-ray scattering (SAXS) confirmed maintained a dimeric quaternary structure at low protein concentrations (~0.15 mg/ml) and physiological salt concentrations for WT and C86A. Different protein-dependent oligomerization tendencies were observed for the proteins, where the effects for C86A were most prominent. High degree of dimerization was observed for both proteins, not affected by the imposed cysteine-substitution to great extent.

To evaluate the biotechnical applicability of these proteins, encapsulation of BvPgb1.2 and Mb were carried out using a lipid-based sponge phase system. High protein concentrations were correlated with increased aggregation tendencies, especially for BvPgb1.2. This effect seemed to be reversible upon agitation and the internal sponge phase structure was maintained. When analyzed with size exclusion chromatography (SEC), no protein leakage was detected for the nsPgb. To study toxicological effects of BvPgb1.2 WT, C86A and Mb, these proteins were labeled with fluorescent markers and studied *in vivo* using a zebrafish model. The globins were injected into brain and lateral muscle tissues to study potential accumulation and toxicity. Granules of C86A and Mb were detected in brain tissue, while all proteins were observed in the muscle tissues. In general, either no/low oxidative stress was observed for the penta- or hexacoordinated globins, except for the most concentrated granules of BvPgb1.2 WT, indicating good tolerance in this model system.

The results highlighted in this thesis provide valuable insights in order to improve/strengthen several areas within biotechnology, including potential generation of resilient crops, iron supplements and oxygen therapeutics.

Key words: Hemoglobin, Phytoglobin, Hexacoordination, Characterization, Protein Engineering, Recombinant Protein Production, Chromatography, Crystalization, Dimerization, Protein Dynamics.

Classification system and/or index terms (if any)

Supplementary bibliographical information — **Language:** English

ISSN and key title — **ISBN** 978-91-7422-936-3

Recipient's notes | Number of pages: 94 | Price

Security classification

Signature Date 2023-04-12

Phytoglobins: Elucidation of Structural, Functional and Practical Features of Non-Symbiotic Plant Hemoglobins

Simon Christensen

LUNDS
UNIVERSITET

Coverphoto by Simon Christensen

Faculty of Engineering
Department of Chemistry
Division of Pure and Applied Biochemistry

ISBN 978-91-7422-936-3 (print)
ISBN 978-91-7422-937-0 (digital)

Printed in Sweden by Media-Tryck, Lund University
Lund 2023

Don´t Hesitate. When The Time Comes, Just Act!

-Star Fox

Table of Contents

All the figures displaying 3D-structural models and protein structure comparisons were performed using Chimera software from UCSF [1, 2]. Note: It has been an effort in the globin community over the last years to present these proteins as "Phytoglobins" instead of symbiotic/non-symbiotic plant hemoglobins [3]. Moreover, "leghemoglobin" or "legHb" are old terms for symbiotic Pgbs from legumes. Since the literature is based on the old description, it should be noted that these terms are equivalent. In addition, the partial pressure of oxygen at which 50% saturation of the hemoglobin is reached (p_{50}), is equivalent to the molar affinity constant K_m for phytoglobins, which is usually used in the literature. Conventional one or three letter abbreviations of amino acids residues are used and not included in the Abbreviations section (except for the cysteine (position 86) and tyrosine (position 115) residues).

Popular Summary

When people hear the word “Hemoglobin”, they probably recognize this term from high school biology as the protein transporting oxygen from the lungs to the tissues in the human body. This property, to carry and deliver oxygen through the blood, is essential for life, as we know it. However, it is less well known that plants also contain hemoglobins, which is strange since this protein is associated with blood and as far as we know, plants do not have blood. So why would plants need these kinds of proteins when they do not use hemoglobin in red blood cells to transport oxygen?

It was in the late 1930s when plant hemoglobins were discovered for the first time. It was found in the root nodules of soybean, a household legume in several cultures. This protein was classified as a symbiotic hemoglobin, because legumes form a symbiosis with soil-living bacteria. In this relationship, the hemoglobin regulates the oxygen levels in the soil while the bacteria supply the plant with nitrogen in return. In other words, a perfect match. It would take another four decades before a new set of hemoglobins were found in other plants than legumes, hence given the name non-symbiotic hemoglobins. Today, we have classified these proteins into three classes; class I, class II and class III non-symbiotic hemoglobins and some form is found in all plants we know of. We call them phytoglobins instead (phyto meaning plant) and the exact role of these proteins is still under investigation. However, substantial progress has been made in this area since phytoglobins were first discovered. The general understanding is that non-symbiotic phytoglobins bind oxygen too tight and release it too slow to function as an efficient oxygen carrier. Instead, these proteins are proposed to function as nitric oxide regulators in plants, an important signaling molecule produced under stressful conditions. Therefore, it is believed that this kind of globin helps the plants in their stress response and combat this by regulating the levels of produced nitric oxide while maintaining a healthy redox and energy balance under these circumstances, instead of oxygen delivery.

Since phytoglobins seem to have a vast array of functions and contain the heme moiety, they have sparked a great biotechnical interest in utilizing them for human purposes. For instance, iron deficiency is one of the leading micronutrient deficiencies in the world and could lead to anemia. This condition affects approximately 42% of children less than five years old and 40% of pregnant women according to World Health Organization estimations. The current treatment for

anemia is primarily supplementation with inorganic iron salts, which commonly results in iron overload and unpleasant side effects, including nausea, constipation and fatigue. Heme iron is generally accepted as a more bioavailable option in terms of iron, not having to ingest huge doses and is not affected by absorption inhibitors as the inorganic form. Would phytoglobins be a suitable replacement to relieve this current problem and increase the level of dietary iron in a sustainable manner instead of meat products?

Over the years, Hemoglobin-based oxygen carriers have received more attention as artificial blood substitutes to counteract the constant shortage of blood donors around the world. However, despite of great efforts, these oxygen therapeutics are still limited. Problems include lack of stability, causing harmful side reactions and insufficient production volumes. Even though phytoglobins are not meant to transport oxygen, would it be possible to engineer these proteins while maintaining their stable structure and deliver oxygen to the most oxygen-deprived tissues?

In a constantly changing world, there is always a great need of stable and sustainable food production, something that many of us take for granted in our generation. In order for us to adapt to climate change and still produce sufficient amounts of food, resilient crops are needed to combat these problems. Since phytoglobins are known to help plants survive during stressful conditions, would overexpression or gene insertion of these proteins directly in plants provide more resilient features?

The biotechnical applications of phytoglobins are many, but more knowledge and research regarding these proteins are needed in order for us to utilize them to our advantage. Work presented in this thesis is primarily based on class I non-symbiotic phytoglobins from sugar beet, including their characteristics and generating mutant(s) with altered properties. By using protein engineering, structural and functional features were investigated as well as intra- and intermolecular protein interactions. In addition, efforts have been made to find suitable practical solutions for these proteins that could facilitate their applicability in biotechnology and biomedicine. This work has been an attempt to further understand the biochemical features of phytoglobins and how to translate them into practical use to relieve some of the issues we are facing in the modern world.

Populärvetenskaplig Sammanfattning

När man hör ordet ”Hemoglobin” känner man förmodligen igen detta från grundskolans biologi och påminns om att det är proteinet som transporterar syre från lungorna till resterande vävnader i kroppen. Egenskapen att transportera syret genom blodet är en förutsättning för livet som vi känner till det idag. Dock är det inte lika känt att växter också innehåller hemoglobiner, vilket är underligt då växter inte innehåller något blod. Så varför behöver växter dessa proteiner då de inte använder hemoglobin i röda blodkroppar för syretransport?

Under 1930-talet upptäcktes växthemoglobin för första gången. Det hittades i rotknölar hos sojabönor, en vanlig baljväxt i flera kulturer. Detta protein kallades för ett symbiotiskt hemoglobin då baljväxter formar en symbios med jordlevande bakterier. I detta förhållande är det hemoglobinet som reglerar syrehalterna i jorden medan bakterierna levererar kväve till växten. Med andra ord så är detta väldigt gynnsamt för båda parter. Det skulle ta ytterligare fyra årtionden innan en ny klass av hemoglobiner hittades i andra växter än baljväxter. Därför fick dessa namnet icke-symbiotiska hemoglobiner. Idag har vi delat in dessa proteiner i tre klasser; klass I, klass II och klass III icke-symbiotiska hemoglobiner och någon form av dessa återfinns i alla växter som undersökts. Vi kallar de för fytoglobiner (fyto för växt) och den exakta rollen som de har undersöks fortfarande. Dock har det gjorts betydande framsteg inom detta område genom åren sedan fytoglobiner först upptäcktes. Den generella uppfattningen är att icke-symbiotiska fytoglobiner binder syre alldeles för hårt och lösgör det för långsamt för att kunna fungera som en effektiv syretransportör. Istället har man föreslagit att denna typ av globiner fungerar som regulator av kväveoxid i växter, vilken är en viktig signalmolekyl som bland annat produceras under stressfulla förhållanden. Därför tror man att dessa proteiner hjälper växten hantera stressresponser genom att reglera mängden kväveoxid och bevara en gynnsam redox- och energibalans under dessa betingelser istället för att leverera syre.

Eftersom fytoglobiner verkar ha många olika funktioner hos växter och det faktum att de innehåller hemgruppen har de väckt ett stort biotekniskt intresse i hur dessa kan användas för mänsklig vinning. Till exempel är järnbrist ett av de ledande bristtillstånden runt om i världen och kan leda till blodbrist (anemi). Denna åkomma drabbar ungefär 42% av världens barn under fem år och 40% av gravida kvinnor enligt Världshälsoorganisationen. Den nuvarande behandlingen för anemi är främst att ta tillskott med oorganiska järnsalter, vilket vanligen resulterar i överdosering av

järn och otrevliga bieffekter, som illamående, förstoppning och utmattning. Hemjärn är generellt sett som ett mer biotillgängligt alternativ, där mängden tillfört järn därför kan minskas samt att upptaget inte förhindras av andra faktorer som för den oorganiska formen. Skulle fytoglobiner vara ett passande alternativ för att lindra de nuvarande problemen och öka mängden järn från kosten på ett hållbart sätt istället för att äta mer köttprodukter?

Genom åren har även forskningen mot artificiellt blod, så kallade Hemoglobin-baserade syrebärare, fått mer uppmärksamhet för att motverka den konstanta bristen på bloddonatorer. Men trots stora satsningar kvarstår fortfarande problem som instabilitet, åstadkommande av skadliga sidoreaktioner och otillräcklig produktionsmängd. Även om fytoglobiner inte är gjorda för att transportera syre, skulle inte en möjlig lösning kunna vara att designa dessa proteiner utifrån detta men samtidigt behålla dess stabilitet och då kunna leverera syre till de mest syrefattiga vävnaderna?

I en konstant föränderlig värld finns det alltid behov av en stabil och hållbar livsmedelsproduktion, vilket är något som ofta tas för givet i vår generation. För att vi skall kunna förhålla oss till de rådande klimatförändringarna och fortfarande producera tillräckliga mängder mat så behövs förmodligen mer motståndskraftiga grödor som kan stå emot detta. Eftersom fytoglobiner tros hjälpa växter i att överleva under stressande förhållanden, vore inte ett alternativ vara att överuttrycka eller introducera motsvarande gener direkt i växterna för ökad tålighet?

De biotekniska applikationerna av fytoglobiner är många men mer grundläggande kunskap och forskning behövs för att kunna använda dem till vår fördel. Arbetet som presenteras i denna avhandling är baserat på klass I icke-symbiotiska fytoglobiner från sockerbeta, där deras funktioner är i fokus samt framtagandet av nya mutanter med nya egenskaper. Strukturella och funktionella egenskaper har undersökts genom användningen av proteindesign, där även intra- och intermolekylära interaktioner har haft ett betydande fokus. Det har även gjorts försök att finna praktiska tillämpningar för att öka den biotekniska och biomedicinska potentialen angående dessa proteiner. Detta arbete har utvidgat den existerande kunskapen gällande egenskaperna hos fytoglobiner och hur man skulle kunna använda dem i praktiken för att lösa de utmaningar som den moderna världen står inför.

Abstract

Phytoglobins (Pgb) (plant hemoglobins) are heme-containing proteins and less well known in comparison to the mammalian counterparts hemoglobin (Hb) and myoglobin (Mb). They share the same secondary structure as other globin-related proteins, six to eight α-helices comprising the myoglobin-fold (Mb-fold) with one prosthetic heme moiety per subunit. Pgbs are divided into different classes, where the non-symbiotic Pgbs (nsPgb) are found in vast number of plants. Unlike their mammalian counterparts, nsPgbs usually have an additional coordination site interacting with the iron called the distal histidine, in addition to the proximal histidine below the heme plane. Even though they bind oxygen (O_2), the proposed functions of these proteins are linked to nitric oxide (NO) metabolism, redox balance and energy maintenance, especially in hypoxic conditions, to a higher extent than oxygen transport.

One aim of this thesis was to generate mutant(s) with altered characteristics using site-directed mutagenesis and investigate the role of these residues. One residue in particular was a conserved cysteine residue located at position 86 (C86A) in sugar beet (*Beta vulgaris* ssp. *vulgaris*) nsPgb BvPgb1.2. After crystallization and structure determination, this cysteine-to-alanine substitution had no impact on the tertiary and quaternary structure but major implications regarding functionality were observed. The mutant had faster autoxidation rate and increased thermal stability, while the wild-type (WT) protein showed higher peroxidase activity. Both species appeared as homodimers and did not show any heme loss, unlike human hemoglobin.

To gain more information regarding the intra- and intermolecular interactions in nsPgbs, triple-labeled BvPgb1.2 was expressed, purified and analyzed using 2D Nuclear Magnetic Resonance (NMR). 83% of the expected amide cross-peaks were assigned. The majority of the non-assigned residues were located in G and H α-helices, which are proposed to be an important area of dimer interactions. Hydrogen bonding between specific residues (T53 and E123) and a hydrophobic cluster in opposing monomers were suggested to be important in dimer formation. Structure predictions were conducted to gain knowledge regarding the role of the conserved cysteine in dimerization and mapping the lost residues in the crystal structure. These residues were located at highly flexible regions in the N- and C-termini, as well as in the loops for helices CD/DE. Small angle X-ray scattering (SAXS) confirmed maintained a dimeric quaternary structure at low protein concentrations (~0.15

mg/ml) and physiological salt concentrations for WT and C86A. Different protein-dependent oligomerization tendencies were observed for the proteins, where the effects for C86A were most prominent. High degree of dimerization was observed for both proteins, not affected by the imposed cysteine-substitution to great extent.

To evaluate the biotechnical applicability of these proteins, encapsulation of BvPgb1.2 and Mb were carried out using a lipid-based sponge phase system. High protein concentrations were correlated with increased aggregation tendencies, especially for BvPgb1.2. This effect seemed to be reversible upon agitation and the internal sponge phase structure was maintained. When analyzed with size exclusion chromatography (SEC), no protein leakage was detected for the nsPgb. To study toxicological effects of BvPgb1.2 WT, C86A and Mb, these proteins were labeled with fluorescent markers and studied *in vivo* using a zebrafish model. The globins were injected into brain and lateral muscle tissues to study potential accumulation and toxicity. Granules of C86A and Mb were detected in brain tissue, while all proteins were observed in the muscle tissues. In general, either no/low oxidative stress was observed for the penta- or hexacoordinated globins, except for the most concentrated granules of BvPgb1.2 WT, indicating good tolerance in this model system.

The results highlighted in this thesis provide valuable insights in order to improve/strengthen several areas within biotechnology, including potential generation of resilient crops, iron supplements and oxygen therapeutics.

List of Papers

The thesis is based on the following publications and manuscripts, numbered from I to V. The publications/manuscripts are appended to the thesis.

I. **Oxidative Implications of Substituting a Conserved Cysteine Residue in Sugar Beet Phytoglobin BvPgb1.2.** Simon Christensen, Leonard Groth, Nélida Leiva-Eriksson, Maria Nyblom and Leif Bülow. 2022. *Antioxidants*. **11**(8): 1615. DOI: 10.3390/antiox11081615.

II. **Conformational Dynamics of Phytoglobin BvPgb1.2 from *Beta vulgaris* ssp. *vulgaris*.** Simon Christensen, Olof Stenström, Mikael Akke and Leif Bülow. 2023. *International Journal of Molecular Sciences.* **24**(4): p. 3973. DOI: 10.3390/ijms24043973.

III. **Evaluation of Oligomer Formation for Sugar Beet Phytoglobin BvPgb1.2 and Cysteine-Substituted Mutant.** Simon Christensen, Henrik Vinther Sørensen, Petra Pernot, Leif Bülow and Marie Skepö. Manuscript.

IV. **Encapsulation of Sugar Beet Phytoglobin BvPgb1.2 and Myoglobin in a Lipid Sponge Phase System.** Jennifer Gilbert, Simon Christensen, Tommy Nylander and Leif Bülow. Submitted.

V. **Toxicological Effects of Penta- and Hexacoordinated Hemeproteins and Their Implications in a Zebrafish Model System.** Simon Christensen, Fredrik Ek, Roger Olsson and Leif Bülow. Manuscript.

My Contribution to Papers:

Paper I: I took a major part in planning, performed the cultivation, expression and purification of the phytoglobins, designed and performed the majority of the experiments and data analysis, as well as co-wrote the manuscript.

Paper II: I took a major part in planning, performed the cultivation, expression and purification of labeled-phytoglobin, assisted in evaluation of dynamic data and co-wrote the manuscript.

Paper III: I took a major part in planning, performed the cultivation, expression and purification of the phytoglobins, conducted simulations using AlphaFold, performed the SAXS measurements and co-wrote the manuscript.

Paper IV: I took a major role in planning, performed the cultivation, expression and purification of the phytoglobin, prepared concentrated protein samples for encapsulation, performed the size exclusion chromatography and co-wrote the manuscript.

Paper V: I took a major part in planning, performed the cultivation, expression and purification of the phytoglobins, labeled the proteins with fluorescent markers, prepared all samples for injections into zebrafish and co-wrote the manuscript.

Documents not Included in Dissertation

I. **Spatial- and Temporal Expression of Novel Phytoglobins in *Avena sativa*.** Leonard Groth, Simon Christensen, Anna Barjuan Grau, Alfredo Zambrano and Leif Bülow. Manuscript.

II. **Engineering and Conjugation of Surface-Exposed Cysteine Residues in Phytoglobin BvPgb1.2 from *Beta vulgaris* ssp. *vulgaris*.** Simon Christensen and Leif Bülow. Manuscript

III. **Importance of Cysteine 86 and Tyrosine 115 in BvPgb1.2 and Cysteine-Substituted Mutant Regarding Redox Maintenance and Electron Transfer.** Simon Christensen, Dimitri A Svistunenko and Leif Bülow. Manuscript.

Patent Application

I. **2151293-4: Generation of Optimized Phytoglobins.** Simon Christensen, Leonard Groth, Leif Bülow and Mats Larsson.

Abbreviations

ABTS	2,2'-Azinobis [3-ethylbenzothiazoline-6-sulfonic acid]-diammonium salt
ALA	δ-Aminolevulinic acid
AF568/647	Fluorescent dyes at 568 and 647 nm, respectively
ApoMb	Apomyoglobin H64Y/V67F (heme scavenger)
ATP	Adenosine Triphosphate
BvPgb	Phytoglobin from Sugar Beet, *Beta vulgaris* ssp. *vulgaris*
cDNA	complementary Deoxyribonucleic Acid
Cgb	Cytoglobin
CN^-	Cyanide ion
CO	Carbon monoxide
CO_2	Carbon dioxide
C86A	Cysteine-to-alanine substitution at position 86 in BvPgb1.2
^{13}C	Carbon isotope
Da	Dalton (g/mole)
Dicot	Dicotyledons
DLS	Dynamic light scattering
E.coli	*Escherichia coli*
Fe^{2+}	Ferrous oxidation state of iron
Fe^{3+}	Ferric oxidation state of iron
Fe^{4+}	Ferryl oxidation state of iron
GMO	Genetically Modified Organisms
Hb	Hemoglobin
HBOC	Hemoglobin-based oxygen carrier
HemD	Uroporphyrinogen III synthase
HemF	Coproporphyrinogen III oxidase
HemH	Ferrochelatase
H_D	Distal histidine
H_P	Proximal histidine
H_2O_2	Hydrogen Peroxide
^2H, D_2O	Deuterium
IPTG	Isopropyl β-D-1-thiogalactopyranoside
K_d	Dimer dissociation constant
K_H	Hexacoordination equilibrium constant

K_m	Molar affinity constant
Mb	Myoglobin
metHb	Hemoglobin in Fe^{3+} oxidation state
Monocot	Monocotyledons
MOPS	3-(N-morpholino) propanesulfonic acid
MQ H_2O	milliQ water
mya	Million years ago
NAD	Sodium dithionite
$NAD(P)^+$	Nicotinamide Adenine Dinucleotide (Phosphate) (Oxidized)
NAD(P)H	Nicotinamide Adenine Dinucleotide (Phosphate) (Reduced)
nanoDSF	Nano-differential scanning fluorimetry
Ngb	Neuroglobin
NH_4^+	Ammonium ion
NMR	Nuclear Magnetic Resonance
NO	Nitric Oxide
NO_2^-	Nitrite
NO_3^-	Nitrate
nsPgb	non-symbiotic Phytoglobin
^{15}N	Nitrogen isotope
OD_{600}	Optical density (cell density) at 600 nm
O_2	Oxygen
$O_2^{\bullet-}$	Superoxide radical
Pgb	Phytoglobin
pI	Isoelectric point
P80	Polysorbate 80
R_g	Radius of gyration
ROS	Reactive oxygen species
SAXS	Small-angle X-ray scattering
SEC	Size Exclusion Chromatography
sPgb	Symbiotic phytoglobin
T_m	Melting temperature (denaturation temperature)
trPgb	Truncated phytoglobin
Tris-HCl	Tris (hydroxymethyl) aminomethane – hydrochloric acid
TROSY	Transverse relaxation optimized spectroscopy
$t_{1/2}$	Half-life: Time required to fall below half its original value
UV-Vis	Ultraviolet-Visible light
Y115, Tyr115	Tyrosine residue at position 115 in BvPgb1.2
α-KG	α-ketoglutarate
λ	Autoxidation rate: Rate to go from ferrous to ferric oxidation state
τ	Scaling time or lifetime: Average time to stay in a particular state

1 Introduction

1.1 Background and Aims

In the modern world, the importance of new and improved methods to relieve the impact of difficult challenges facing humanity has never been greater. For instance, is the world adapted to the increasing need of food to feed a growing population worldwide? Can the crops used today withstand the potential effects from climate change? Perhaps we need additional knowledge of how to create crops that are more resilient in order to fulfil current and future needs. Adaption to constantly changing conditions and finding creative solutions have always been important hallmarks of humankind. This is a vital ability and important to maintain during coming decades, in order to combat modern concerns regarding the climate. Regardless if is sudden draughts, flooding, temperature changes or altered soil conditions, the need for science and technology to solve these issues is always imminent.

Apart from environmental factors, there is always a need to find new ways to promote human health and alleviate conditions for suffering individuals. One example of this is the increasing rate of nutrition deficiencies occurring around the world. For instance, iron deficiency, ultimately leading to anemia if not treated properly, is one of the most common deficiencies of micronutrients and has been on a steady rise during the latest decades [4]. Iron deficiency, with subsequent anemic conditions, is a major public health problem according to the World Health Organization and is defined by hemoglobin (Hb) levels <12 and 13 g/deciliter in women and men, respectively [5]. This affects young children and pregnant women in particular, where approximately 42% and 40% of these groups are estimated to be suffering from this condition, respectively [6]. The occurrence of this and related conditions are most predominant in developing countries but on the rise on a global scale. One explanation could be increasing propensity to eat more vegetarian diets for sustainability and health reasons. Even though plants also contain iron, it is primarily of inorganic forms with a lower bioavailability (<5%) [7] in comparison to meat-based sources (15-35%) [8]. An aspect of this difference is the impact of several plant-derived chemicals (phytochemicals) serving as anti-nutrients. Examples are phytate (phytic acid), polyphenols and oxalic acid but also calcium and other divalent metal ions have a negative effect on iron absorption [9]. A recent study showed high levels of phytate in current plant-based meat substitutes, in addition to low quantities of bioavailable iron [10]. The nutritional limitations of

these products are problematic, especially in providing sufficient amounts of vital micronutrients, in addition to the high anti-nutrient content. These implications adds to the difficulty in maintaining a plant-based diet and still providing appropriate amounts of bioavailable macro- and micronutrients to the body.

Traditionally, meat has been the main source or iron and other micronutrients, where the heme iron is more efficient in terms of absorption and utilization by the human body. The heme iron comes primarily from heme-binding proteins, such as myoglobin (Mb) in muscle tissue and hemoglobin (Hb) from blood. Mentioned inhibitors for inorganic iron do not affect the heme-derived iron. Furthermore, iron in the free ferric (Fe^{3+}) oxidation state is more prone to precipitation, further diminishing absorption [9]. Thus, by changing our eating habits toward a more plant-based diet, there is a risk of lowering the amount of bioavailable iron and other nutrients crucial for maintenance of healthy bodily functions. Could there be an option to have an increased plant-based diet and still provide sufficient amount of iron for the body to thrive?

Another important need is development of Hemoglobin-based oxygen carriers (HBOCs). These are pharmaceutical products designed to mimic the function of Hb and deliver additional oxygen (O_2) when applied. The need for new blood donors is always high and new solutions have to be engineered to keep up with this demand to aid in emergency situations, such as surgeries, accidents or transfusions. Most of current HBOCs are based on engineered human or bovine Hb but often come with several drawbacks, such as severely increasing blood pressure by nitric oxide (NO) scavenging, oversupply of O_2 (low O_2 affinity) and lower cardiac output [11, 12]. In addition, issues regarding storage capacity, biocompatibility and prehospital availability have been focal points for modern HBOCs [13]. Could O_2-carrying proteins from other sources, such as plants, be a future solution to this concern? Would it be possible to ensure O_2 supply to even the most O_2-deprived tissues and provide a universal product simultaneously?

Plant hemoglobins (phytoglobins, Pgb) were first discovered in root nodules of legumes in 1939 [14], where these proteins were believed to aid in nitrogen-fixation by providing proper O_2 conditions for the soil bacteria. Ever since, several different kinds of Pgbs have been discovered with different proposed functions for particular plants and tissues. For instance, non-symbiotic phytoglobins (nsPgb) were discovered for the first time in the 1980s in non-leguminous plants [15, 16]. These proteins have been found in a vast array of plants in some form and have different functions than the symbiotic ones found in legumes. Their functions are linked to NO metabolism, maintaining energy status and redox balance, especially in hypoxic conditions, in addition to aid the plant in response to several abiotic and biotic stresses [17, 18]. Due to the variability of catalyzed reactions and involvement in numerous processes regarding stress responses in plants, it is difficult to pinpoint any specific functions for these proteins. However, this versatility could be a

blessing in disguise, opening up possible biotechnical applications where these could potentially be useful.

The aim of this thesis is to expand the current understanding regarding the characteristics of Pgbs in general and nsPgbs in particular. One important technique in examining the roles of amino acid residues in these proteins is site-directed mutagenesis, where a single or several residues can be substituted and their causal effects studied [19]. In addition, this technique could provide information how to engineer these proteins and make them more suitable for practical applications in the future and help to elucidate their structure and function *in vivo*. This ability, to generate new and improved proteins with desired properties, is an important tool in solving some of the mentioned problems that we are currently facing but also in the future.

This thesis is primarily based on five papers and manuscripts, which have been developed during my doctoral years. The focus has been on class I nsPgbs from sugar beet (*Beta vulgaris* ssp. *vulgaris*). Moreover, identification of novel Pgbs, especially in oat (*Avena sativa*), have also been of interest (not included in the thesis).

The aim of **Paper I** was to examine the role and function of the conserved cysteine residue at position 86 in sugar beet nsPgb BvPgb1.2. We found, by successfully crystallizing and solving the crystal structures for both the wild-type (WT) and cysteine-substituted mutant (C86A), that it had only minor impact on the overall structure of the protein. In addition, both proteins appeared as dimers when analyzed at low protein concentrations. However, the functionality was drastically altered. For instance, the autoxidation rate and thermal stability increased for the mutant while the WT showed 3-fold higher initial peroxidase activity. However, none of the nsPgbs showed any heme loss in contrast to the positive control. The absent heme loss tendency indicated high stability and potentially their less reactive nature in comparison to pentacoordinated Hbs. In this work, we have highlighted the important role of the cysteine regarding the function and activity, even though this residue seemed to be of less structural importance.

In **Paper II**, we expressed triple-labeled (^{2}H, ^{13}C and ^{15}N) BvPgb 1.2 WT in order to analyze the internal protein dynamics and potential dimerization interactions using 2D Nuclear Magnetic Resonance (NMR). The protein was analyzed in its cyanide- and O_2-bound forms. In total, 137 of the 165 expected amide cross-peaks were assigned (83%) for cyano-BvPgb1.2. The non-assigned residues were mostly located in the G and H α-helices, which are one of the proposed regions to be involved in the dimerization, ultimately forming the homodimer. Here, we highlighted the protein dynamics in BvPgb1.2 and identification of flexible regions, features of importance to better understand the functions of nsPgbs *in planta* and for various biotechnological applications.

In **Paper III**, we investigated the role of the conserved cysteine during dimerization in more detail using small angle X-ray scattering (SAXS). Both proteins maintained their dimeric quaternary structure at high/low protein and salt concentrations. Different protein-dependent oligomeric states were observed for BvPgb1.2 WT and C86A, where the mutant seemed to oligomerize to greater extent. Thus, the main function of the conserved cysteine residue is probably associated with activity and redox balance rather than dimeric quaternary structure stabilization. In addition, we performed structure simulations using AlphaFold for both BvPgb1.2 WT and C86A to map the residues missing electron density in the crystal structures, generating a complete structure prediction model. Furthermore, the dimeric crystal structure of C86A was used to investigate important residues in dimer formation in more detail.

In **Paper IV** and **V**, we examined practical applications of these nsPgbs and their potential toxic effects in biological systems, respectively. In **Paper IV,** we examined the heme interactions in a lipid sponge phase system that has been used to encapsulate and potentially protect/delivery proteins to desired locations in the body. This was conducted with one pentacoordinated and one hexacoordinated hemeprotein (Mb and BvPgb1.2 WT, respectively). Major and possibly reversible aggregation tendencies were observed for high BvPgb1.2 concentrations (60 mg/ml), but to lesser extent for Mb. Properties such as isoelectric point (pI), buffer strength, lipid composition and iron coordination might be taken into consideration to optimize the encapsulation in this and similar systems.

In **Paper V**, the heme-containing proteins were also analyzed in a toxicological trial using a zebrafish model. Mb, BvPgb 1.2 WT and C86A were fluorescently labeled using two activated ester-markers. Thus, the accumulation and effect of the proteins could be observed. Labeled C86A and Mb accumulation were detected after injection in brain tissue, but no associated oxidative damage in surrounding tissues were detected. BvPgb1.2 WT was not found in the brain, but all globins were observed in lateral muscle tissues. Most injections showed no or low level of oxidative stress, except for the most concentrated granules of labeled BvPgb1.2 WT. If this effect is consistent or a result of the microinjection procedure itself remains to be proven. Different behavior between protein accumulation in brain and muscle was also found, where fluorescent granules were found close to the site of injection in the muscle tissues. In general, no excessive oxidative stress was observed for either the penta- or hexacoordinated globins, indicating good tolerance in this model system.

The theme of this thesis is class I nsPgbs, whether in regards of expanding the knowledge of previously unknown features and characteristics (**Paper I-III**) or their practical applicability (**Papers IV** and **V**). In addition, finding novel Pgbs to be analyzed and characterized have also been within the scope but not included in the main thesis. In the concluding chapter, the highlights of the presented work are summarized and potential future prospects are discussed. In the end, the goal was to

provide new biotechnical tools and proteins that could be of use for humanity and improve human health in a sustainable manner.

1.2 Hemoglobin and Heme: Structure and Function

An essential basis for life among vertebrates is the transportation of O_2 from the lungs to the different cells and exportation of carbon dioxide (CO_2) and hydrogen ions in the opposite direction. Mentioned ligands are bound, transferred and released to different tissues by the protein Hb in a reversible manner [20]. While Hb is transported around the body in red blood cells by the circulatory system and deliver O_2 where it is needed, Mb is located in muscle tissues. Here, Mb serves as both O_2 supplier/storage and generator of cellular energy when needed by the cells [21].

The transition from an anaerobic to aerobic existence provided the opportunity to utilize more energy than previously. Approximately 15 times more energy from glucose can be harnessed in the presence of O_2 due to the involvement of oxidative phosphorylation in the mitochondria [21]. Because of the importance of these globins, they are among the most heavily studied proteins in vertebrates. Mb and Hb were the first proteins to be crystallized and structure determined using X-ray crystallography [22-24].

The reason why Hb can carry O_2 and other ligands is due to the presence of the prosthetic group *heme.* This group is indicated in Figure 1, in addition to the quaternary structure of human Hb A [25]. The protein exists as a tetramer in its functional form, equipped with one heme moiety per subunit. A central water cavity is also present between the dimers of subunits α_1-β_1 and α_2-β_2. Each subunit consists of six to eight α-helices, depicted A-H, which folds into the motif called the myoglobin fold (Mb-fold) [25, 26]. The fold consists of a 3/3 sandwich structure formed between helices A, B, C and E over helices F,G and H, forming a hydrophobic core containing the heme group [26]. Moreover, the short helix D is usually located in the loop region preceding helix E and in close proximity to helix C.

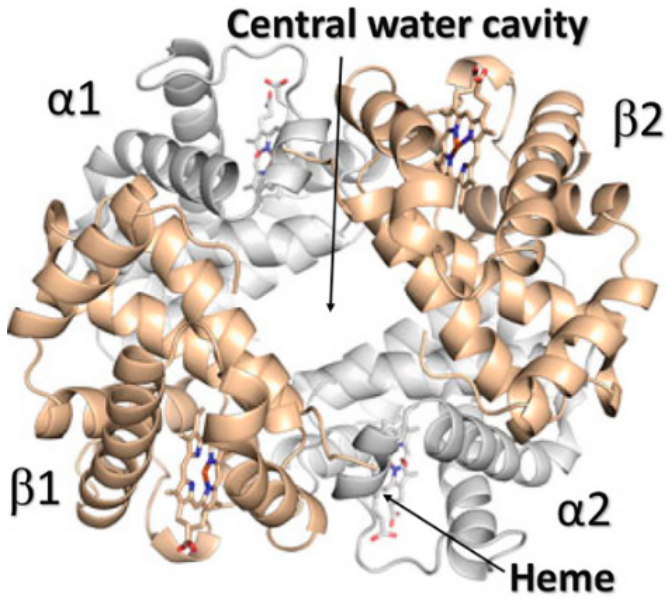


Figure 1. Quaternary structure of human Hb A (PDB-ID: 2HHB [27]). The four subunits forming the tetramer, heme moieties and central water cavity are depicted. Adopted and modified from [25].

The prosthetic heme group, also called protoporphyrin IX, consists of four pyrrole rings linked by methine bridges to form a tetrapyrrole ring with a central iron atom. (Figure 2A) [21, 28]. The bound iron generates the red color of both muscles and blood seen throughout the animal kingdom. Beside the ability to transport and/or store O_2, heme-containing proteins also perform several enzymatic reactions, including peroxidase functions (reacting with hydrogen peroxide (H_2O_2)) and serve as electron transporters in redox maintenance [29]. Furthermore, other ligands than O_2 are capable of binding to the heme, including NO, carbon monoxide (CO) and cyanide (CN^-) among others [25].

Different oxidation states are important in binding and transportation of O_2. For instance, the ferrous oxidation state (Fe^{2+}) is capable of binding O_2 while the ferric form (Fe^{3+}) is not, rendering it inactive [21]. The iron atom can form two additional bonds adding to the four ones from the nitrogen in the tetrapyrrole ring. It is possible for the centered iron to interact with two additional sites located on either side of the heme plane and these binding sites are referred to as the fifth and sixth coordination sites [30]. Thus, Hbs are divided into pentacoordinated (five iron coordination sites) and/or hexacoordinated (six iron coordination sites) depending of the number of interactions with the central iron. In pentacoordinated Hbs, the iron interacts with a histidine located in the F-helix, usually at position F8, termed the proximal histidine [25, 26, 30]. In contrast, all six coordination sites are occupied in hexacoordinated Hbs. In both cases, the heme group is linked to the protein

backbone by the proximal histidine and this residue has been shown to be important in anchoring the prosthetic group to the protein. Hexacoordinated Hbs have an additional histidine residue in the binding pocket, referred to as the distal histidine [21]. The distal histidine is crucial for stabilization of the bound O_2 by donation of a hydrogen bond or interacting with the iron itself [30]. This histidine is located in the E-helix (E7), above the heme plane [25]. Therefore, Hbs can be classified by penta- or hexacoordination depending on the involvement of the distal histidine in iron stabilization. Although, Hbs may have different degrees of distal histidine involvement in ligand stabilization, giving them features from both classes. Figure 2B displays the interactions of the proximal (H_P) and distal histidine (H_D) residues. Here, it is evident that pentacoordinated Hbs also contain the distal histidine residue, but it is not interacting with the central iron.

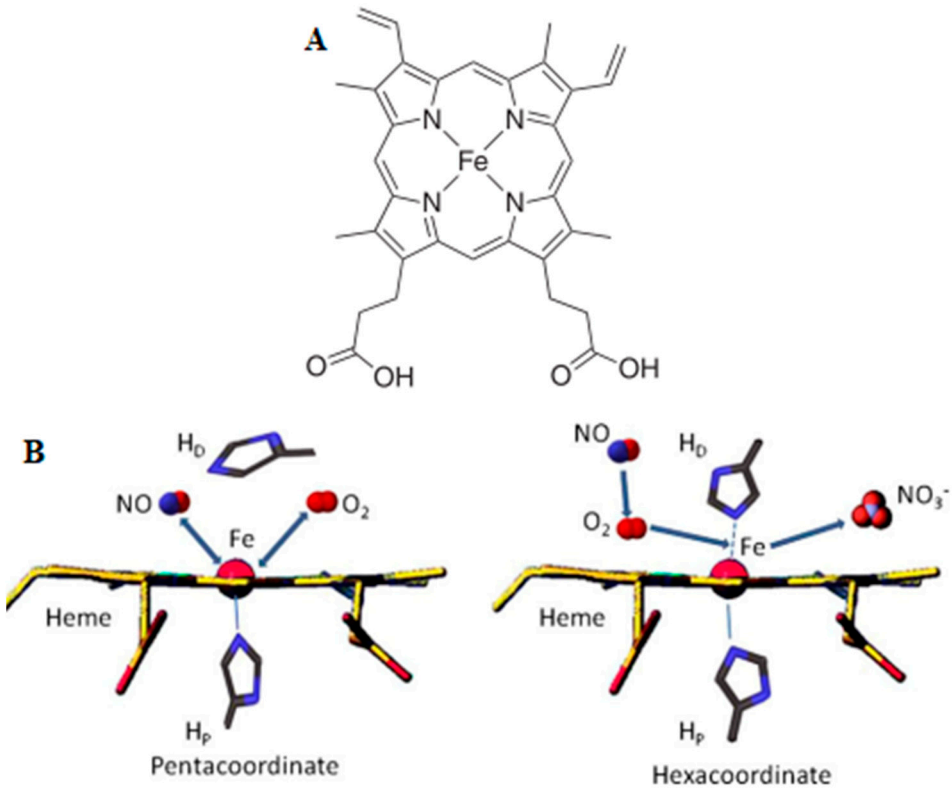


Figure 2. (A): Structure of the prosthetic heme moiety. **(B):** Orientation of the proximal (H_P) and distal histidine (H_D) in penta- and hexacoordinated Hbs. Adopted and modified from [31] and [30], respectively.

Pentacoordinated globins are usually involved in traditional O_2 binding, transport or storage, exemplified by human Hb A in red blood cells and Mb in muscle tissues. On the other hand, examples of hexacoordinated globins are also found in the animal

kingdom, namely neuroglobin (Ngb) and cytoglobin (Cgb), but also the more recently discovered nsPgbs in plants are hexacoordinated [30, 32]. Due to the additional stabilization provided by the distal histidine, these proteins are not associated with O_2 transport because of the tight binding and slow release of O_2. For example, human Ngb is present in brain, retina and several tissues in the nervous system and its exact function is not fully determined [32, 33]. The current proposed model suggests involvement in the protection against hypoxic and oxidative stress, probably through a NO scavenging mechanism. This model indicates the neuroprotective properties of Ngb, especially during O_2-deprived conditions [32]. In similar fashion, Cgb is proposed to scavenge NO during fibroblast proliferation and providing cellular protection in oxidizing conditions [34, 35]. This protein is found in many different tissues in humans and upregulated under hypoxia. Furthermore, cells expressing nsPgbs have been found to prevent cell death in hypoxic conditions, in addition to increased level of ATP and low NO accumulation in these cells [36]. In particular, class I nsPgbs are important for maintenance of the redox balance and energy conversion in O_2-deprived environments [37, 38].

Although the precise functions of penta- and hexacoordinated globins are not fully understood, it is evident that the two coordination states of the heme iron provide different characteristics and functionalities. The rest of the chapter will investigate the origin of Pgbs in particular, how/why they evolved, what is known in terms of how they affect plant tissues and their current proposed functions *in vivo*.

1.3 Phytoglobins: From Origin to Characterization

1.3.1 Evolution of Globins

Since the first crystal structures of Hb and Mb were published in the 1960s, a vast number of different globins and related proteins have been discovered, ranging from simpler bacteria to more advanced eukaryotic organisms. The development in recombinant DNA technology, sequencing output and computer power over the last 50 years have significantly facilitated the ability to sequence entire genomes and compare these to elucidate ancestral traits.

In the globin case, there are three lineages in the bacterial super family that either belongs to two structural classes; 3/3- and 2/2-fold globins [26]. The two lineages belonging to 3/3-fold are the F family, where flavohemoglobins and similar single-domain globins are found and the S family, where so called globin-coupled sensors and related proteins are included [39, 40]. The final lineage, the T family, possesses the unconventional 2/2-fold, mainly due to conversion of helix F into a loop and shortening of helix A. The T family is divided into three different subfamilies [26], but will not be examined in further detail. It should be noted that this kind of

phylogenetic analysis is a continuous process since more globin-like proteins are discovered due to easier access of genomic data. For instance, the discovery of hexacoordinated nsPgbs, Ngbs and Cgbs further added to the great variability of these proteins. One of the more recent additions of hexacoordinated globins, called androglobin, contains both protease/globin domains. These globins have been found in testis tissue and may be part of a more novel metazoan globin lineage [41].

It is evident that all globins from vertebrates and non-vertebrates, as well as symbiotic and non-symbiotic Pgbs and several microbial eukaryotes, possess the conventional 3/3-sandwich fold and have homologous gene sequences to the F family of bacterial proteins [26]. Globins from the T family have been found in plants and a number of microbial eukaryotes [42]. In contrast, fungi seem to be the only group with flavohemoglobins and single domain globins from the S lineage and archaea only express globins from the S family [43, 44].

A phylogenetic tree model of globin evolution can be seen in Figure 3A. The analysis was conducted using genomic sequencing originating from different species, including globins from plants, algae (*Chlorophyte* and *Stramenopile*), vertebrates and members of the bacterial F and T families [45]. A close relationship between vertebrate Ngb and nsPgbs (depicted plant nsHb) was found, originating from bacterial flavohemoglobins. This evolution of eukaryotic globins might have been a result of horizontal gene transfer (transfer of genetic material between bacterial cells) from respective bacterial lineages. The origin of this transfer involved one or both of the crucial events of endocytosis, the basis for life as we know it; creation of mitochondria involving a α-proteobacterium and chloroplasts involving cyanobacterium [40, 42, 46]. Therefore, horizontal gene transfer form an ancestor of the proteobacteria to the last eukaryote common ancestor (~1200 million years ago (mya)) could be a representation of current vertebrate Ngb, algal and plant 3/3-folded globins [45]. In addition, the relationship between penta- and hexacoordinated globins is also evident in Figure 3A. This suggests that ancestors of Ngb first developed into Cgb and later into the pentacoordinated vertebrate Mb and α/β subunits of Hb. Moreover, both class II nsPgbs and symbiotic Pgbs (sPgb, Lbs) are both derived from the ancestral class I nsPgb [45]. Thus, it is reasonable that pentacoordinated Pgbs originated from hexacoordinated ancestors [32, 47, 48].

A representation of the evolution from bacterial to vertebrate and other eukaryotic globins and the relationship between them can be seen in Figure 3B. Here, globins from the F family, containing the conventional 3/3 Mb-fold, developed from the chimeric form of flavohemoglobins to the current form of single-domain Hbs, losing its original reductase domain. These events were the foundation and further developed into the current globins, mainly: globins in vertebrates (Hb, Mb, Ngb and Cgb), Pgbs, fungi and other microbial eukaryotes.

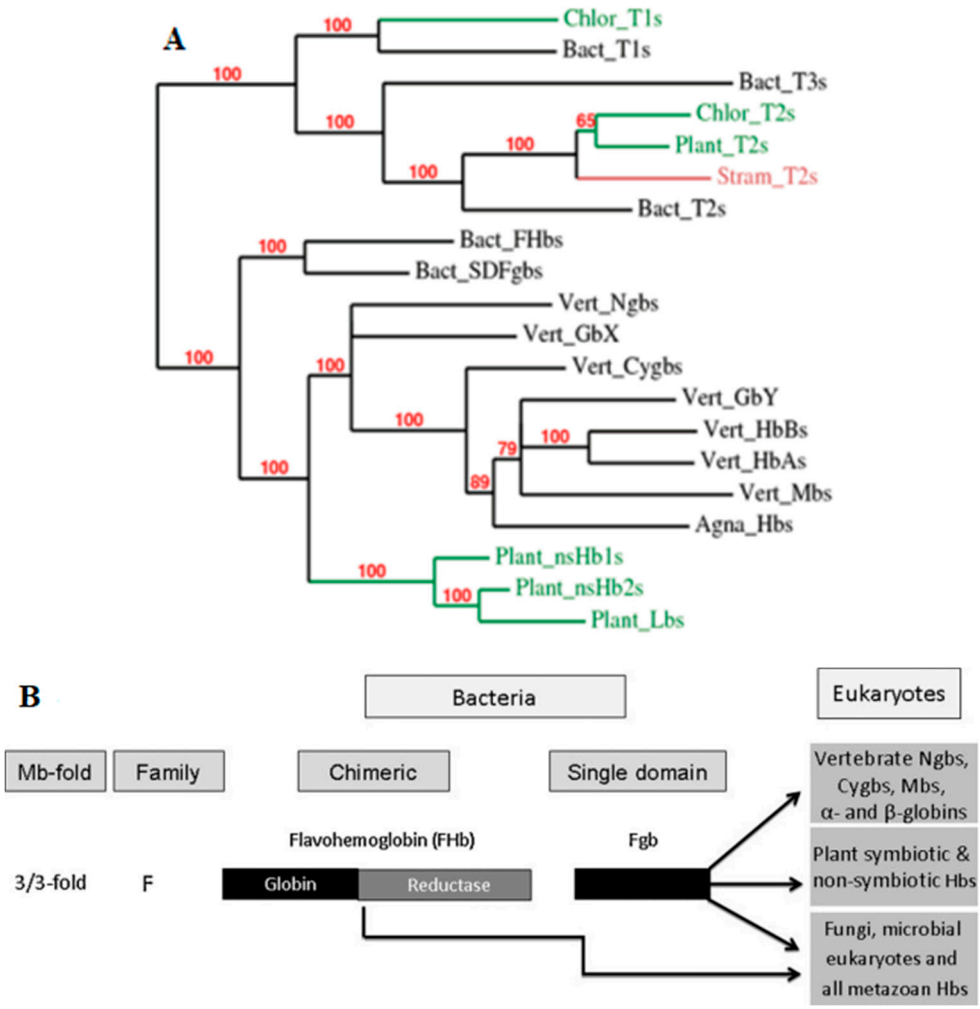


Figure 3. (A): Phylogenetic tree of globin evolution. Present vertebrate Ngb, Cgb and Pgbs originated from a bacterial flavohemoglobin ancestor. Majority of bootstrap values (red) were 100% (high percentage of support). Longer distance between the branches indicates more sequence diversity. More separated branches represent more genetic distance/timed evolved. **(B):** Schematic representation of modern globins originating from the bacterial flavohemoglobin F family. Abbreviations: Agna: Agnatha (fish), Bact: Bacterial, Chlor: Chlorophyte, Fgb: single-domain flavohemoglobin, FHb: chimeric flavohemoglobin, GbX: Globin in fish, amphibians and reptiles, GlbY: Globin in amphibians, nsHb: nsPgb, Plant Lbs: symbiotic Pgbs, Stram: Stramenopile, SDFgbs: single domain flavoglobins and Vert: Vertebrate. Adopted and modified from [45] and [26], respectively.

The historic timeline of globin development took place over millions of years, rendering it difficult to predict exactly when important stages in the evolution occurred. Globins are probably more than 3000 million years old and prokaryotic ancestors contained early members of the F, S and T families [42]. Several genomic transitions occurred during these years to ultimately form the current globins. The first transition is believed to occurred between 450 to 800 mya and developed

members of bryophytes, such as mosses [49]. This was the period when plants started to colonize the land, instead of being limited to aquatic life. The second major event, approximately 320 mya, yielded the first spermatophytes (seed plants), which were split into ancestors of gymnosperms (woody plants) and angiosperms (flowering plants). The third and fourth transitions diversified the early angiosperms and formed monocotyledons (monocots) and dicotyledons (dicots) (both between 140 to 180 mya) [26, 50]. It is also postulated that genome duplication events, in addition to horizontal gene transfers, took place prior to the second and third events, which were important stages for genome diversification and generation of new genomic varieties [51].

The first observed ancestors of non-symbiotic/symbiotic Pgbs and Pgbs from legumes are approximately 150 and 60 million years old, respectively. Therefore, it is likely that symbiotic Hbs emerged from early versions of non-symbiotic Hbs and spread among flowering plants [26]. However, it is not entirely clear if globins from legumes stems from the same origin as all other sPgb [52]. Although, it is likely that leguminous Hbs arose from class II nsPgbs, since these hexacoordinated globins are found in early dicots, whereas the more recent development of legumes suggests transition to the pentacoordinated sPgb found in leguminous root nodules [53].

In summary, 3/3 land plant nsPgbs emerged from an ancient ancestor shared with algae and developed into class I and II through the diversification of angiosperms ~140 to 180 mya (later split into monocots and dicots). Whether or not symbiotic globins in legumes or non-legumes coevolved, these proteins emerged at a later stage and probably originate from early versions of class I nsPgbs [26]. During this event, the iron coordination of the heme transitioned from its hexacoordinated precursors to the pentacoordinated sPgbs. Examples of these alterations were decreasing length of the loop between helices C/D and N- and C-terminals, leading to a less mobile distal histidine [54]. Thus, the ability for the distal histidine to interact with the heme iron and adopt both penta- and hexacoordinating configurations was lost in this transition. Finally, it should be mentioned that 2/2 truncated Pgbs (trPgb, class III nsPgb) emerged from a bacterial ancestor through a horizontal gene transfer to early ancestors of all eukaryotes or at least the ones shared by algae and land plants [26].

1.3.2 Classification of Phytoglobins

Ever since Pgbs were first discovered in the 1930s [14], the elucidation of their origin and characteristics have been a major focal point. In the 1980s, nsPgbs where observed for the first time in non-nodulating plants [15], adding to the diversity of these proteins. Today, Pgbs are divided into different classes depending on their evolution and features. Here, the main classes are presented and described in terms of suggested functions and characteristics.

1.3.2.1 Class I Non-Symbiotic Phytoglobins

Class I nsPgbs are usually expressed in low μM concentrations (5-20 μM) when induced *in vivo* [55, 56]. These proteins have a very high O_2 affinity, (K_m~2 nM), making them unlikely to function as O_2 transporters in plants [47, 57]. Instead, these proteins are proposed to have NO scavenging functions due to the increased levels of this compound and upregulations of class I nsPgb in abiotic and biotic stressful environments [58, 59]. Elevated levels of NO has been linked to control several developmental stages in plants, for instance seed dormancy and root development [60], in addition to environmental impact from cold and hypoxia [61]. These levels are probably modulated by class I nsPgbs, aiding the plant in adaptive response mechanisms and these proteins are present in all angiosperms [44].

Class I nsPgbs appears as dimers in its active configuration, comprised of monomers ~17-19 kDa in size. Dimeric interactions are proposed to occur in the interface of helices G and BC loop region with a dissociation constant in the μM-level [56, 62]. The dimer interactions and flexible regions for class I nsPgbs are discussed in Chapter 3 (**Paper II** and **III**). These proteins usually contain one or two conserved cysteine residues, located in the E helix [56]. When this residue in barley nsPgb (*Hordeum vulgare*) (position 79) was mutated to serine, an intermolecular disulfide bond was detected in the WT protein using mass spectrometry and suspected to stabilize the quaternary structure and maintaining proper redox balance [63]. Moreover, the distal histidine tends to interact with the bound ligand through hydrogen bonding. This interaction could be disrupted when the cysteine is absent due to solvent entry into the heme pocket. The features of cysteine substitution in class I nsPgbs will be further examined in Chapter 2 (**Paper I**).

In regards of hexacoordination, the hexacoordination equilibrium constant (K_H), binding constant for the distal histidine residue, is comparably low (K_H~2) in class I nsPgbs [30, 64]. This results in an equilibrium between penta- and hexacoordination states, allowing ligands to bind when the distal histidine is directed away from the heme iron. This causes very high O_2 affinity observed for these proteins, especially regarding slow O_2 release (0.038 s^{-1}) [47, 48] compared to O_2 transporting Hbs (> 5 s^{-1}) [44]. The tight binding and slow release of O_2 further implicate their involvement in NO scavenging in hypoxic conditions, preferably in association with a reducing agent (reductase activity and NAD(P)H) to regain the active Fe^{2+} state [65, 66]. The reaction between O_2 and NO in nsPgbs yields nitrate (NO_3^-), which the plant can more easily metabolize. Nitrogen in NO_3^- and ammonium ions (NH_4^+) can be incorporated into new amino acids and stored/transported to support growth and overall metabolism [67]. The elevated expression levels of class I nsPgbs under anaerobic and other stressful environments where NO is produced imply their importance in plant survival and resiliency under these conditions [68]. A representation of this proposed mechanism is illustrated in Section 1.3.3.

1.3.2.2 Class II Non-Symbiotic Phytoglobins

Class II nsPgbs are found in non-leguminous dicots/early angiosperms and less information exits regarding this class in comparison to class I [44, 69]. Unlike class I nsPgbs, class II proteins have a lower oxygen affinity (K_m~300 nM) [70]. Hence, their function is more linked oxygen supply in plants, especially in developing tissues and are not normally expressed under hypoxic conditions [17, 30]. In addition, this class have tighter hexacoordination (K_H~100) [47], indicating heavy involvement of the distal histidine occupying the sixth coordination site instead of ligand. Therefore, a possible function for these proteins might be O_2 sensing, storage and/or diffusion and share characteristics with sPgbs [32]. However, overexpression of a class II nsPgb in *Arabidopsis* showed earlier bolting, hence involvement in NO metabolism might be possible [55]. In addition, this overexpression produced significantly increased amounts of total lipid content in seed and increased the energy status up to 5 times the normal level [70]. Nevertheless, no precise functions have been attributed for this class and they might be involved in a diverse range of processes in plants. Expression of this class have been observed in roots, stems, flowers, leaves and immature fruits in several plants, including tomato (*Solanum lycopersicum*) [71]. Further characteristics will be discussed in the section regarding class II nsPgb from sugar beet in Section 2.1.

1.3.2.3 Class III Non-Symbiotic Phytoglobins (Truncated Phytoglobins)

As previously mentioned, trPgbs possess the unconventional 2/2 globin fold due to shortening or absence of helices A and F. These proteins have lower O_2 affinity (K_m~1500 nM) [72] in comparison to the 3/3 folded counterparts. TrPgbs are usually pentacoordinated in their oxygenated state, but have also shown transient hexacoordination tendencies [72]. The functions related to these proteins have not been elucidated but they could be involved in O_2 delivery in oxygenated environments. For instance, downregulation of truncated globins in bacteria made them more sensitive in environmental changes, such as presence of H_2O_2 and high O_2 solution pressures [73]. A trPgb from *Arabidopsis* showed expression in roots and shoot tissue but was downregulated under hypoxic conditions, opposite to class I nsPgbs [74]. This observation in addition to the lower O_2 affinity indicates different functions than NO scavenging. These proteins are present in both monocots and dicots [44].

1.3.2.4 Symbiotic Phytoglobins

SPgbs were first detected in nitrogen-fixating root nodules of soybean (*Glycine max*). Hence, the old terminology for sPgbs are “legHb” due to the first observation in legumes [30]. SPgbs are usually monomeric and have moderate O_2 affinity (K_m~50 nM) [47]. These proteins have rapid O_2 binding (230 $\mu M^{-1}s^{-1}$), especially compared to class I and class II nsPgbs (24 and 1.3 $\mu M^{-1}s^{-1}$, respectively), but do not show any hexacoordinated configuration (K_H=0) [47, 75].

The root nodules expressing the proteins are usually red due to their high concentration *in vivo* (up to 0.7 mM) and these proteins are believed to be important in maintaining proper O_2 levels for the sensitive nitrogenase process in nitrogen fixation [30]. This process is an interplay between the plant and *Rhizobium* bacteria, where the sPgb buffers the O_2 levels and facilitate fixation of atmospheric nitrogen into useable NO_3^- for the plant. The symbiosis is crucial for the nitrogen-fixating ability in legumes and help to promote/maintain soil fertility by increasing the level of fertilizing nitrogen [76]. This effect was demonstrated by gene silencing of three sPgbs in the legume *Lotus japonicas*, where growth was affected and maintained only when nitrogen fertilizer was added. The mutants showed increased levels of O_2 in the nodules, indicating the importance of these proteins in regulating O_2 in nitrogen-fixating symbiosis, especially when external nitrogen is not added [77, 78]. Furthermore, increase of oxide radicals and peroxide, in addition to disrupted energy balance in these mutants highlighted there importance in nitrogen fixation [79]. Due to the more recent development of leguminous plants and the sequence similarity between sPgbs and class II nsPgbs, this underlines the transition from hexacoordinated ancestors to the pentacoordinated symbiotic globins, as previously mentioned [53].

1.3.3 Function of Class I Non-Symbiotic Phytoglobins

Some of the characteristics of class I nsPgbs were described in previous sections. Here, one of the proposed models of why these proteins are expressed and what functions they might have in plants is described.

NO has been linked to several functions in plants. This signaling molecule is overproduced in response to abiotic stress factors, such as low nutrient availability, high salinity, drought, high/low temperatures and radiation [80]. Here, NO influences hormonal signals, interacting with reactive oxygen species (ROS) and modulating both gene expression and protein activity to adapt to stressful environments. It also mediates signals in biotic stress, such as pathogenic infections [81]. In response to plant hormones, NO can induce seed germination, control root elongation, facilitate different developmental stages and promote iron uptake [82]. However, this molecule is also a free radical that can induce oxidative damage and toxic effects if overproduced and not neutralized [83]. Thus, it is crucial for the plant to regulate this signaling molecule and mitigate its toxic implications in response to mentioned factors. One important aspect here might be elevated expression levels of class I nsPgbs.

As mentioned before, the expression of nsPgbs occurs in a variety of tissues and due to a vast range of stimuli. For example, they are expressed in calli (wound healing), seeds, roots and stems in members of monocot and dicot families [84]. Metabolic stimuli shown to upregulate levels of class I nsPgbs include hypoxia and sucrose [85, 86], low energy levels in cells (low ATP status) [38] and in rapid growing

tissues, such as root meristems in germinating seeds [56]. Hypoxic conditions have also been linked to decreased ATP levels and class I nsPgb expression is further exacerbated by calcium ions [55]. Moreover, elevated expression levels have also been observed when plants have been exposed to osmotic stress and high salinity, heat and cold exposure, infections, treatment with various compounds (NO, nitrite (NO_2^-), NO_3^-, salicylic acid and H_2O_2 [18, 82]. Furthermore, hormone regulation also stimulate class I nsPgb expression, exemplified by elevated expression levels seen in roots when treated with polyamines and jasmonic acid [87].

In general, these proteins are expressed and upregulated when plants respond to stressful conditions in order to combat whatever stimuli it may be, abiotic or biotic. One attempt of summarizing proposed events in NO metabolism and related responses under hypoxic conditions is shown in Figure 4. Here, these proteins may be used to regenerate $NAD(P)^+$ from NAD(P)H to maintain glycolysis and ATP production, thus regulating energy levels. The reduction in ATP during hypoxia stimulates the production of NO, which is toxic in high concentrations. The class I nsPgb oxidizes NO into NO_3^- by incorporating O_2 and is reduced back to its active Fe^{2+} state through the reductive capacity of NAD(P)H and reductase activity. This allows the protein to be regenerated and react with additional NO. In the end, both NAD(P)H and NO are oxidized, contributing to maintenance of the redox balance and energetic status of the plant cells, ultimately promoting survival [17, 18]. In addition, these proteins may also be involved in signal transduction pathways for several plant hormones, including auxins, ethylene, abscisic acid and cytokinins, adding to the ones previously mentioned.

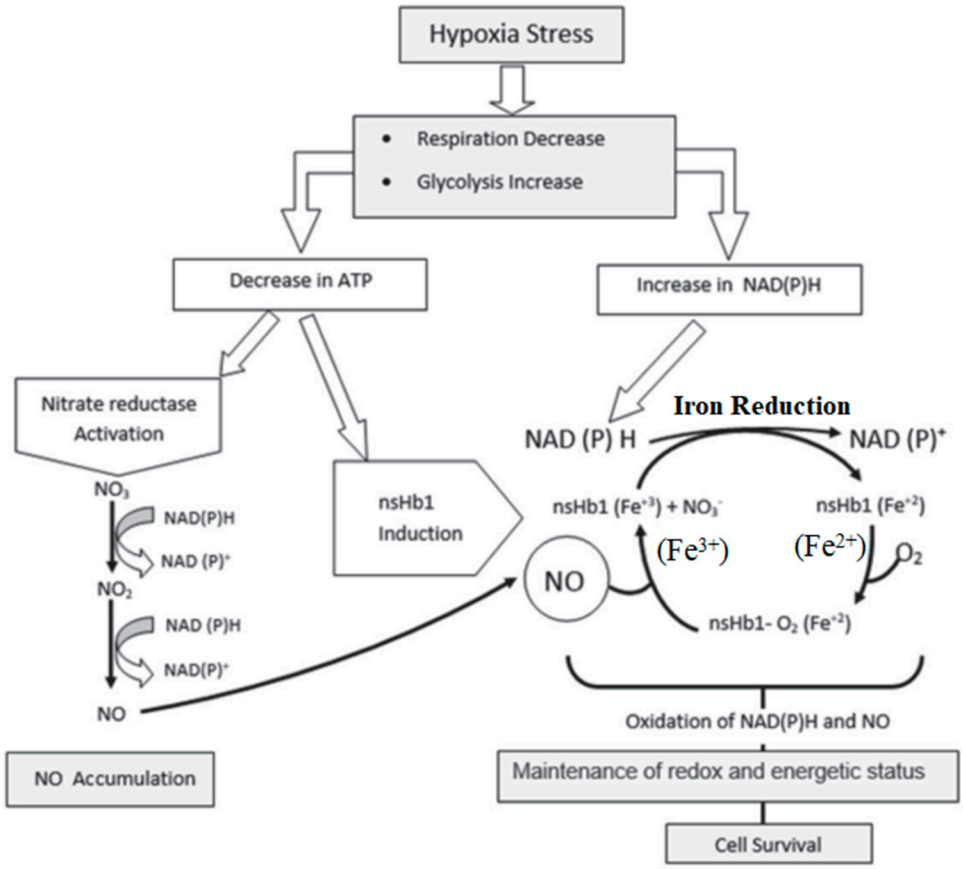


Figure 4. Proposed function of class I nsPgbs (nsHb) in NO metabolism and energy maintenance under hypoxic conditions. Class I nsPgbs (nsHb1) are induced by low energy levels and NO accumulation. By incorporating O_2 to form NO_3^-, the toxic effects of NO are circumvented, promoting redox and energetic status maintenance. Adopted and modified from [17] and [18].

Due to the ability for class I nsPgbs to aid in adaption and survival in hypoxic-related conditions, these might be useful for metabolic engineering applications. Many mentioned studies have shown to improve the resiliency of plants under different stresses, an important property in the modern world due to changing climate and other factors leading to spoiled harvests. For example, overexpression of nsPgb in barley showed promoted plant development and efficient utilization of atmospheric NO upon NO fumigation, in contrast to plants with nsPgb knockdown and WT protein [88]. Studies like this, in addition to overexpression and induction upon stressful growing conditions of other nsPgbs mentioned in this thesis, highlight the potential applicability of these proteins and similar ones in plant resistance programs, producing crops that are more resilient.

1.3.4 Recombinant Expression of Phytoglobins

The majority of mentioned work regarding functionality of nsPgbs have been conducted *in vivo*, by either gene expression analysis, overexpression or suppression of these proteins. Although these procedures are crucial to investigate the impact and functionalities of these proteins in plants, it is difficult to study specific protein interactions and characteristics. The focus of this section is previous work regarding the recombinant expression of nsPgbs and other globins.

One of the first globins from a non-leguminous plant was isolated from *Parasponia* [16]. From this point, questions concerning evolution of plant hemoglobin genes were raised and how this new, isolated type of protein could provide new information in terms of origin and function in plants. A few years later, another protein was solubilized, purified and characterized from the non-leguminous *Casuarina glauca,* but the expression of the Pgb was still inside the root nodules [89].

The first nsPgb in monocots was discovered in barley, especially in seed and root tissues under anaerobic conditions [90]. Furthermore, cDNA of this gene was cloned and expressed in *Escherichia coli* (*E. coli*), the first recombinantly expressed nsPgb [91]. The barley nsPgb consisted of an 18.5 kDa protein and a dimeric native quaternary structure was observed. It had a very slow O_2 dissociation rate, as well as hexacoordinated spin state [91]. Ever since, several other nsPgbs have been expressed using recombinant techniques. Soon after, class I and class II nsPgbs from *Arabidopsis thaliana* were expressed for the first time [92]. It was concluded that these two classes probably have different functionalities in the plant. In addition, two nsPgbs from rice (*Oryza sativa*) once again showed the extremely high O_2 affinity of class I nsPgbs and removal of the distal histidine revealed the hexacoordinated implications [64]. Examples of recombinant expression and crystallization of globins using *E. coli* as the expression system can be seen in Table 1. Today, these are important characteristics of class I nsPgbs and other globins, laying the foundation of future globin-related research.

Table 1. Examples of recombinant globin expression and protein crystallization using *E. coli* as expression system.

Type of Protein (Origin)	PDB-ID	References
nsPgb (Hordeum vulgare)	2OIF	[53]
nsPgb (Arabidopsis thaliana)	3ZHW	[93]
nsPgb (Oryza sativa)	1D8U	[62]
nsPgb (Trema tomentosa)	3QQQ	[94]
nsPgb (Parasponia andersonii)	3QQR	[94]
nsPgb (Beta vulgaris ssp. *vulgaris)*	7ZOS	[95]
Hb (Homo sapiens)	2HHE	[96]
Mb (Homo sapiens)	3RGK	[97]
Ngb (Homo sapiens)	1OJ6	[98]
Cgb (Homo sapiens)	1URV	[99]

One potential limiting factor in recombinant expression of Pgbs and other heme-containing proteins is the heme synthesis and incorporation into the final holoprotein. The heme synthesis pathway of *E. coli* can be seen in Figure 5. Here, α-ketoglutarate (α-KG) from the tricarboxylic acid cycle is converted into δ-aminolevulinic acid (ALA) by a set of different enzymatic reactions [100]. Further catalysis leads to the formation of protoporphyrin IX-structure before the iron incorporation by the enzyme ferrochelatase (HemH). Even though the bacteria are capable of heme production intrinsically, low ALA accumulation/production is severely diminishing the amount of functioning hemeprotein. However, this problem is often mitigated by addition of ALA when protein expression is induced [91]. Moreover, overexpression of uroporphyrinogen III synthase (HemD) and coproporphyrinogen III oxidase (HemF) enzymes in the bacterial heme pathway have shown to increase ALA accumulation [100].

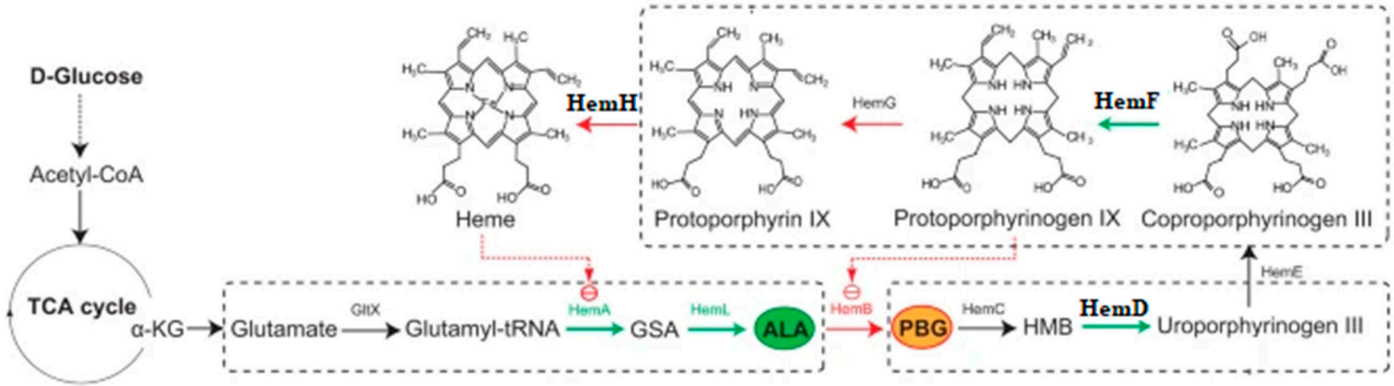


Figure 5. Heme pathway in *E. coli*. Important enzymes for increased heme production are HemD (uroporphyrinogen III synthase), HemF (coproporphyrinogen III oxidase) and HemH (ferrochelatase). Addition of ALA during protein expression has also been shown to increase globin levels. Adopted and modified from [100].

Human-derived Hb genes have also been expressed in different expression hosts. When human Ngb was discovered, this protein was expressed in *E. coli*, purified and characterized [101]. It was found that humans, like plants, have hexacoordinated Hbs and the similarities between Ngb and nsPgbs were observed. Production of human Hb has also been conducted successfully in yeast [102] and vertebrate Mb has been overexpressed in tobacco plants (*Nicotiana benthamiana*) to increase the iron content in these plants [103]. Moreover, efficient production and excretion of soy bean sPgb in the yeast *Pichia pastoris* by genetic engineering reduced the need of cell lysis and purification [104]. The recombinant expression of Hbs have opened up vast opportunities for biomedical use, both for pharmaceutical and nutritional applications.

This thesis is primarily based on a class I nsPgb from sugar beet (BvPgb1.2). Here, 0.3 mM of ALA is used to boost production of protein [105] in a similar manner as for nsPgb from barley. Even though ALA production and accumulation might be a bottleneck, it is not a linear increase in Pgb expression. Instead, when higher and lower concentrations of ALA were added for induction, no significant increase or decrease of nsPgb expression was observed, respectively (unpublished results). Thus, the precise concentration of ALA might not be of great importance as long as it is supplemented. However, high optical density at 600 nm (OD_{600}) and addition of isopropyl β-D-1-thiogalactopyranoside (IPTG) are needed for proper Hb expression, at least using the T7 expression system. An OD_{600} >2 is preferred for human fetal Hb and several nsPgbs [105, 106], while lower cell densities might be used for other recombinant proteins in *E. coli*, such as lactate dehydrogenase [107]. Therefore, expression protocols are usually protein specific and have to be customized for the particular proteins.

Efficient purification of proteins is essential for maximizing protein yields and ensure sufficient purity. For BvPgb1.2, a two-step chromatography protocol has been developed to produce high-purity protein fractions used for further characterization and usage. The first step involves anionic exchange chromatography, where buffer pH is normally set to approximately one pH-unit above the pI of the protein to promote ionic interactions between the colon resins and protein. The protein fractions are later eluted using an increasing salt gradient [105, 108]. The elevated salt concentrations disrupt the protein-resin interactions, promoting elution.

The second step relies on hydrophobic interaction chromatography, where proteins are separated based on surface hydrophobicity. The obtained fractions after anionic exchange are saturated with ammonium sulfate to allow precipitation of contaminating proteins before applied to the column [95, 105]. The hydrophobic interactions are weakened through a water gradient, decreasing the conductivity of the medium. Therefore, the elution process in hydrophobic chromatography stands in contrast to ionic exchange procedures, which have been effective in BvPgb1.2 purification. Although, buffer concentrations and elution conditions are highly

protein-dependent. Thus, careful optimization of expression and purification protocols are of great importance to achieve sufficient amounts and purity of protein.

The functionality of purified, recombinant BvPgb1.2 can be seen in Figure 6. The figure displays the characteristic Soret peak for oxygenated nsPgb (412 nm, blue), reduced with sodium dithionite (NAD) (424 nm, orange) and CO-saturated (416 nm, gray). The shift in Soret peak and heme-specific peaks (between 500-600 nm) indicate an active and functioning recombinant globin. Advances in recombinant expression and chromatography procedures have made it possible to produce high-quality protein samples, which greatly facilitated further protein characterization.

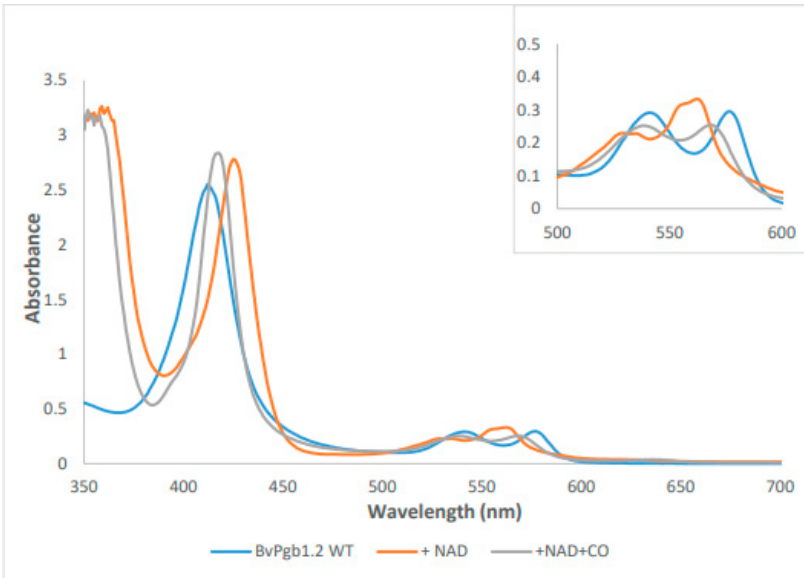


Figure 6. UV-Vis spectrum for oxygenated (blue), reduced (orange) and CO-saturated (gray) recombinant BvPgb1.2. The protein was expressed and purified as described above. The shift in Soret and heme peaks for the different states/ligands indicate a functional globin.

2 Phytoglobins: Structural and Biochemical Features

As previously stated, some forms of Pgbs are found in essentially every plant, where the expression of different classes may be different depending on the origin of the plant. This thesis is primarily based on class I nsPgbs from biennial sugar beet, which will be highlighted in the following chapters.

2.1 Sugar Beet Phytoglobins

Even though several Pgbs from different plants exist and have been studied to varying degrees, there is a need to search for additional proteins, adding information and open up possibilities for utilization. Biennial sugar beet is a valuable crop, both for economical and sustainable reasons [109]. This plant is a eudicot (part of the dicots), having two seed leaves upon germination [110]. Since other members of this clade of plants have shown expression of class I, II and III nsPgbs, similar findings were suspected for sugar beet. When genomic information from this eudicot was investigated, four novel Pgbs were discovered, namely two class I nsPgbs (BvPgb1.1 and BvPgb1.2), one class II (BvPgb2) and one class III, truncated globin (BvPgb3) [110]. BvPgb1.1 and 1.2 showed 67% sequence identity and both displayed ~57% identity compared to BvPgb2. Furthermore, BvPgb1.1 had a longer N-terminal, which is probably cleaved when translocated to the chloroplast. BvPgb1.2 and 2 were found in the cytoplasm until they are transported to their final locations. Similar proteins have been found in the cytosol, nucleus and plastids in plant cells [110].

Phylogenetically, these nsPgbs are related to similar proteins from both monocots and dicots (Figure 7) [44]. Here, the globin sequence from a trPgb in *Arabidopsis thaliana* served as base (outgroup) for the rest of the phylogenetic tree, showing class I for monocots, class I and class II for dicots and sPgbs (Lbs). The nsPgbs from sugar beet, BvPgb1.1 (Bvul1-1), BvPgb1.2 (Bvul1-2) and BvPgb2 (Bvul2), are also highlighted in the figure. The relationship between the class I, class II and symbiotic globins should be noted. This analysis underlined the previous mentioned aspect regarding the evolution of globins, where class I nsPgbs from monocots and dicots

served as the foundation for further development into class II nsPgbs and later into symbiotic ones [26, 45]. The tendency to go from hexacoordination to more pentacoordinated features is evident here.

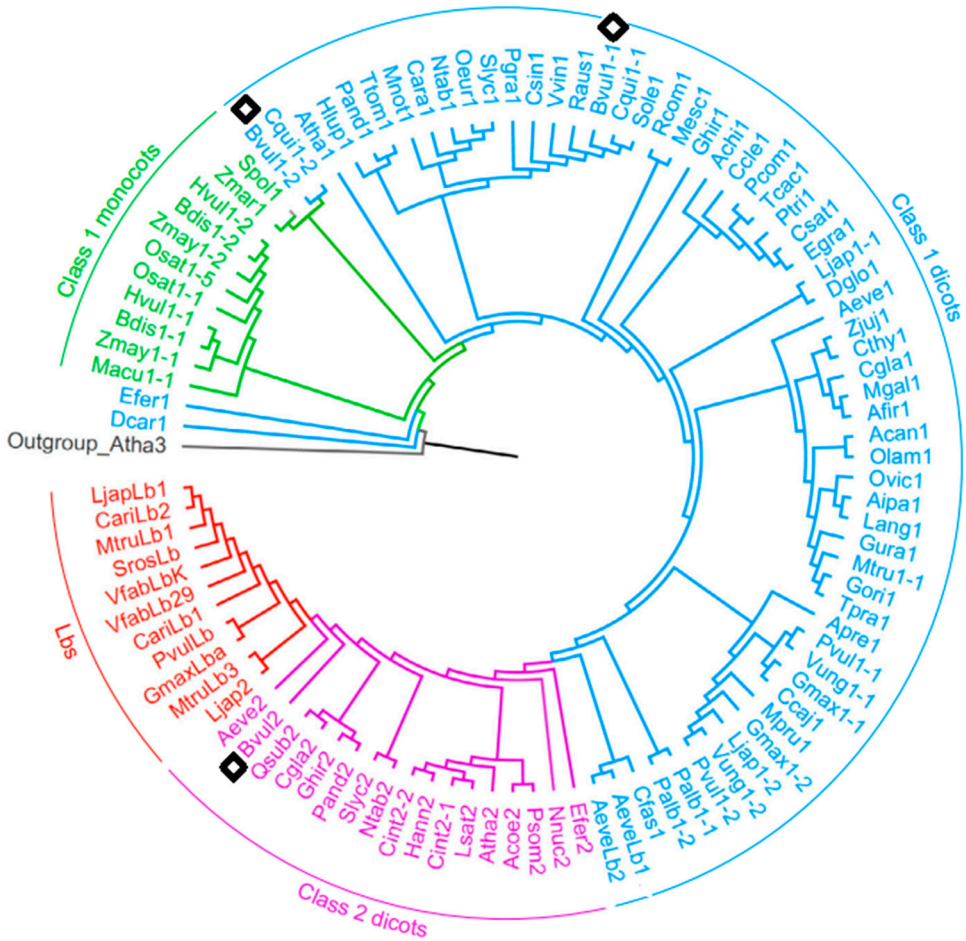


Figure 7. Phylogenetic tree of globin origin for class I nsPgbs in monocots (green), class I and II in dicots (blue and purple, respectively) and sPgbs (Lbs) in legumes (red). The transitions from class I to class II nsPgbs and later to sPgbs (Lbs) were observed. The absence of class II nsPgbs in monocots is also evident. Atha3 was used as outgroup and the majority of bootstrap values were close to 100% except Efer1/Dcar1 branches (<70%). Adopted and modified from [44].

In terms of expression, BvPgb2 was the most widely expressed protein in many tissues of the sugar beet while BvPgb1.2 were mostly found in seeds and early plant development, together with the second class I globin [110]. Class I nsPgbs have previously been shown to be heavily expressed in germinating seeds, probably to maintain energy and redox balance in the early developmental stages [55].

Moreover, BvPgb1.1 and 2 were expressed throughout plant development, where BvPgb2 was the most prominent one. Differential expression during vernalization (promoted flowering) and different photoperiods (long and short days) was observed for both proteins, where BvPgb2 showed the highest overall expression. This might indicate the important role of this protein throughout plant development. Even though BvPgb1.2 was expressed to a lower degree, its importance in seed germination was highlighted.

Spectroscopic measurements and other characterizations of the same proteins were also conducted. It was found that the class I nsPgbs appeared as dimers, while the class II protein consisted in its monomeric form. The degree of hexacoordination showed tighter and less flexible tendencies for BvPgb2, which is in line with the previous knowledge of difference between the classes [105]. In addition, O_2 binding constants and affinities were in accordance with previously reported results (Section 1.3.2) [47]. Furthermore, when reactions with NO were followed, BvPgb1.2 showed the highest NO dioxygenase activity, the ability to convert NO into NO_3^-. This further strengthened the observation of class I nsPgbs being mostly expressed in germinating seeds, where nitrate levels are high. Nitrite reductase activity was also observed for all proteins and was most predominant under low O_2 and nitrogen levels. From this and several other studies regarding characteristics of nsPgbs, it is evident that the function of the proteins will ultimately depend on the current state of the plant and environmental factors, changing the amounts of available ligand concentrations. Even though NO scavenging under hypoxic condition may be one of primary roles for these kinds of proteins (mechanism in Section 1.3.3), the great versatility of possible reactions they could participate in makes it difficult to pinpoint the "correct" function [111].

Due to the versatility of these proteins, new and interesting approaches regarding the applicability of nsPgbs from sugar beet have also been investigated. For instance, BvPgb2 was immobilized and evaluated for its direct electron transfer properties between the protein and carbon electrodes. Its intrinsic electrochemical and catalytic potential displayed high stability and could be suitable to use in this kind of electrode, especially in reaction with H_2O_2 [112]. The same protein was overexpressed both in *Arabidopsis thaliana* and tomato, resulting in increased resistance against drought and osmotic pressure, in addition to increased intrinsic iron levels, especially in the tomato plants [113]. Moreover, both BvPgb1.2 and 2 were deployed in electrical sensors to detect and signal for reactive oxygen and nitrogen species, utilizing their electrochemical and catalytic features in this kind of application [114]. Here, the heme reduction potentials were determined to -349 mV for BvPgb1.2 and ranging between -345 and -457 mV for BvPgb2, depending on the hexacoordinated state.

The structural features of BvPgb1.2 in particular will be described in the next section. This protein was crystallized and characterized together with a cysteine-substituted mutant [95]. Since there are several similarities and common ancestry

between human Ngb and class I nsPgbs [45], a comparison of the crystal structures is presented in Figure 8. The characteristic 3/3 globin fold is present in both proteins with only small differences in positioning of helices and loops. However, the loop between helices D and E is not present in the BvPgb1.2 structure due to lack of electron density information after X-ray diffraction. This loop is depicted for the Ngb. In addition, the distal and proximal histidine residues in helices E and F, respectively, are marked in Figure 8. Even though the heme plane has shifted slightly in relation to each other, both proteins showed the characteristic hexacoordination where both histidine residues interact directly with the heme iron. A sequence comparison between these proteins is shown below, where (*) indicates an identical residue match, (:) denotes residues with similar properties and (.) with different properties [115].

```
Human_Ngb_(1OJ6)      ----------------MERPEPELIRQSWRAVSRSPLEHGTVLFARLFALEPDLLPLFQY      44
BvPgb1.2_(7ZOS)       MSFTNVNYPASDGTVIFTEEQEALVVQSWNVMKKNSAELGLKLFLKIFEIAPTAKKMFSF      60
                                    : . :  *: ***..:.:.  * *  ** ::* : *    :*.:

Human_Ngb_(1OJ6)      NGRQFSSPEDSLSSPEFLDHIRKVMLVIDA---AVTNVEDLSSLEEYLASLGRKHRAVGV      101
BvPgb1.2_(7ZOS)       VRDSDVPLEQN---QKLKGHAMSVFVMTCKSAAQLRKAGKVTFGESSLKHMGSVHLKYGV      117
                         .    *:.    :: .*  .*:::       : :. .::  *. *  :*  *   **

Human_Ngb_(1OJ6)      KLSSFSTVGESLLYMLEKSLGPAFTPATRAAWSQLYGAVVQAMSRGWDGE----    151
BvPgb1.2_(7ZOS)       VDEHFEVTRFALLETIKEAVPEMWSPEMKNAWAEAFNHLVAAIKAEMQRLSTQP    171
                        . *...  :**  :::::   ::*  : **:: :. :* *:.   :
```

Although, the sequence identity is relatively low (18.9%), the iron coordination and tertiary structure are conserved, adding to the versatility of globins. As mentioned earlier, these types of proteins were probably the ancestors of the O_2 transporting/storing Hb and Mb in vertebrates, as well as class II and symbiotic versions in plants, transitioning from hexa- to pentacoordinated characteristics.

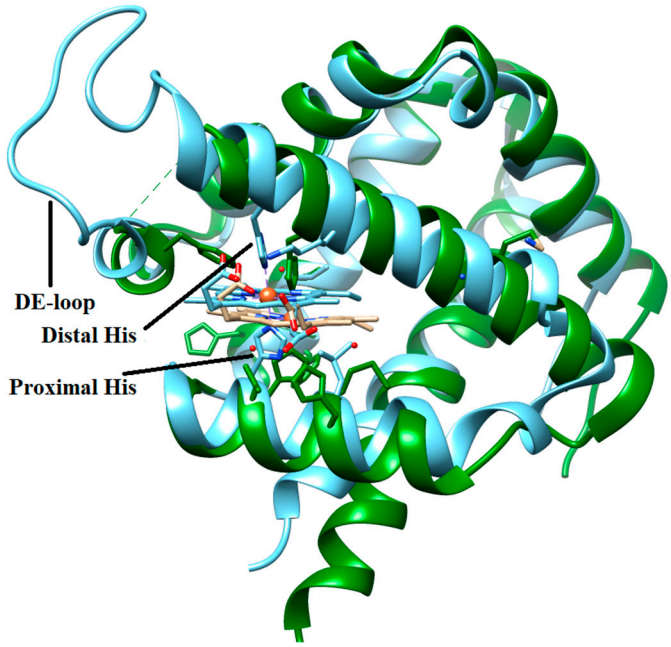


Figure 8. Structure comparison of BvPgb1.2 (PDB-ID: 7ZOS) (green) and human Neuroglobin (PDB-ID:1OJ6 [98]) (light blue). The iron coordinating distal and proximal histidine residues are highlighted, as well as the D/E loop.

2.2 Biochemical Implications of Conserved Cysteine in Phytoglobins

2.2.1 Autoxidation and Structural Impact of Cysteine-Substituted Mutant C86A.

Several interesting amino acid residues have been conserved throughout the evolution of Hbs and Pgbs, which might have important properties in terms of structural stability and/or functionality. In addition to the proximal and distal histidine residues, other conserved residues found in all Pgbs, except trPgbs, are phenylalanine (loop CD1 and B10) and proline (helix C2). [72, 116]. Another example of such conserved residues is the cysteine located at position 86 in BvPgb1.2. Implications of cysteine residues in the globin family have been studied in the past. In human Hb, Cys93 in the β-subunits has been proposed to be important for the subunits to combine into the functional quaternary structure in a correct manner, in addition to affect the redox potential and O_2 affinity [117, 118]. Moreover, cysteine residues α-Cys104 and β-Cys112 are located in the interior of

the globin and suggested to participate in salt bridge formation between subunits, stabilizing the final folded state [119].

Class I nsPgbs have one or several cysteine residues, usually located in helix E [62, 63]. This is where the cysteine in BvPgb1.2 and the corresponding residue in barley nsPgb are found. The residue in barley, in position 79, was substituted to a serine residue in previously published work. This residue seemed to be indirectly involved in reduction of metHb (ferric form) formation and NO turnover [66]. Furthermore, the mutant appeared as a monomer to a higher degree instead a homodimer. The authors hypothesized that the removal of the cysteine led to destabilization of quaternary structure and decreased protection of the Fe^{2+} state inside the heme pocket, probably due to disrupting disulfide bridge formation between monomers [63]. To further characterize the importance of the conserved cysteine, we constructed a cysteine-substituted mutant at position 86 in BvPgb1.2 (C86A) and analyzed its implications regarding structure, stability and functionality (**Paper I**) [95].

Autoxidation is one feature used to examine the proneness of Hbs to transition from its reduced Fe^{2+} state to the oxidized and inactive Fe^{3+}, in addition to superoxide ($O_2^{\bullet -}$) formation [120]. The complete autoxidation of BvPgb1.2 WT and C86A can be seen in Figure 9 (WT: A and B, C86A: C and D, at 25 and 37 ˚C, respectively). The oxidation process was most evident at 542 and 576 nm, indicated in Figure 9 (arrows). The mutant was substantially more prone to reach the inactive Fe^{3+} state in comparison to the WT, estimated at 576 nm.

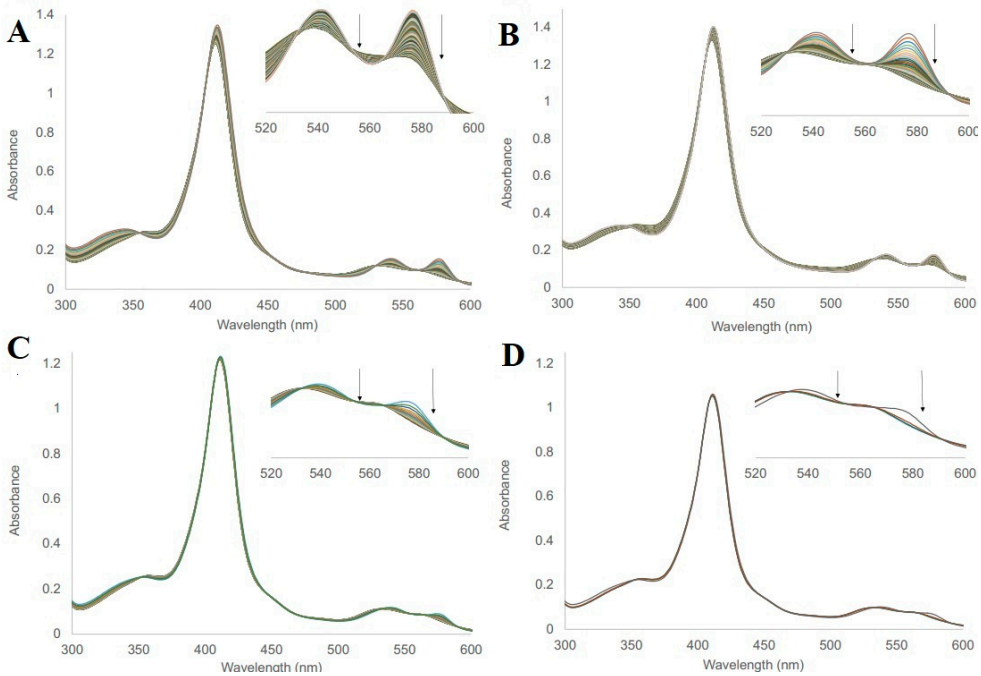


Figure 9. (A) and (B): Complete autoxidation spectra for BvPgb1.2 WT. **(C) and (D):** Complete autoxidation spectra for C86A mutant. The spectra were collected at 25 (A and C) and 37 ˚C (B and D), respectively. Adopted and modified from [95] (**Paper I**).

The autoxidation rate (λ) was 16.7 and 14.4-fold faster for the mutant at 25 ˚C and 37 C˚, respectively (Table 2). In accordance, the scaling time (τ) and half-life ($t_{1/2}$) were higher for the WT but similar values for WT at 25 ˚C and C86A at 37 ˚C were observed. Even though the cysteine substitution affected the tendency to oxidize substantially in BvPgb1.2, the difference was less prominent in comparison with the corresponding results for the barley nsPgb, which reported a 1000-fold increase in autoxidation for the mutated protein [63]. In addition, no differences in dimerization between the proteins were observed, unlike the more predominant monomeric form in barley nsPgb (described in Chapter 3). Corresponding autoxidation rates for human Hb were determined to 0.13 and 0.44 h^{-1} at 25 and 37 ˚C, respectively [121]. This pinpoints the slower tendency to reach the oxidized Fe^{3+} state for hexacoordinated globins, implicating their increased oxidative stability.

Table 2. Autoxidation rate (λ), scaling time (τ) and half-life ($t_{1/2}$) for BvPgb1.2 WT and C86A at 25 and 37 ˚C.

BvPgb	*λ (h^{-1})*	*τ (h)*	*$t_{1/2}$ (h)*
WT 25 ˚C	0.014 ± 0.002	72.2 ± 0.180	50.1 ± 0.125
WT 37 ˚C	0.214 ± 0.003	4.68 ± 0.069	3.24 ± 0.005
C86A 25 ˚C	0.234 ± 0.012	4.28 ± 0.226	2.97 ± 0.157
C86A 37 ˚C	3.08 ± 0.035	0.342 ± 0.004	0.225 ± 0.003

Both BvPgb1.2 and C86A were successfully crystallized and structure determined (Figure 10). The comparison showed only minor differences regarding the structure (A and B), where the characteristic Mb-fold and heme orientation were similar. Moreover, the orientation of the cysteine side chain in the WT (green) did not change compared to the alanine (red) in C86A, indicating that the substitution did not affect tertiary structure in this case. Most class I nsPgbs orient the E-helix towards the center of the porphyrin ring to coordinate the distal histidine in proximity to the iron of the heme moiety [56, 62]. No apparent difference was observed for the C86A mutant.

Similarities between the barley nsPgb and C86A were also observed (C and D). The cysteine in barley nsPgb (green) seemed to have similar orientation as the alanine (red) in C86A. In both cases, the cysteine side chain are oriented toward the surface interface instead of the neighboring monomer forming the homodimer. Even though the possibility remains, it is unlikely that the disulfide formation would occur between the monomers, at least for BvPgb1.2. In addition, the WT seemed to oligomerize to a lesser degree than C86A, strengthening the unlikelihood of disulfide bonding to stabilize quaternary structure in BvPgb1.2 (**Paper III**), as opposed to the barley protein.

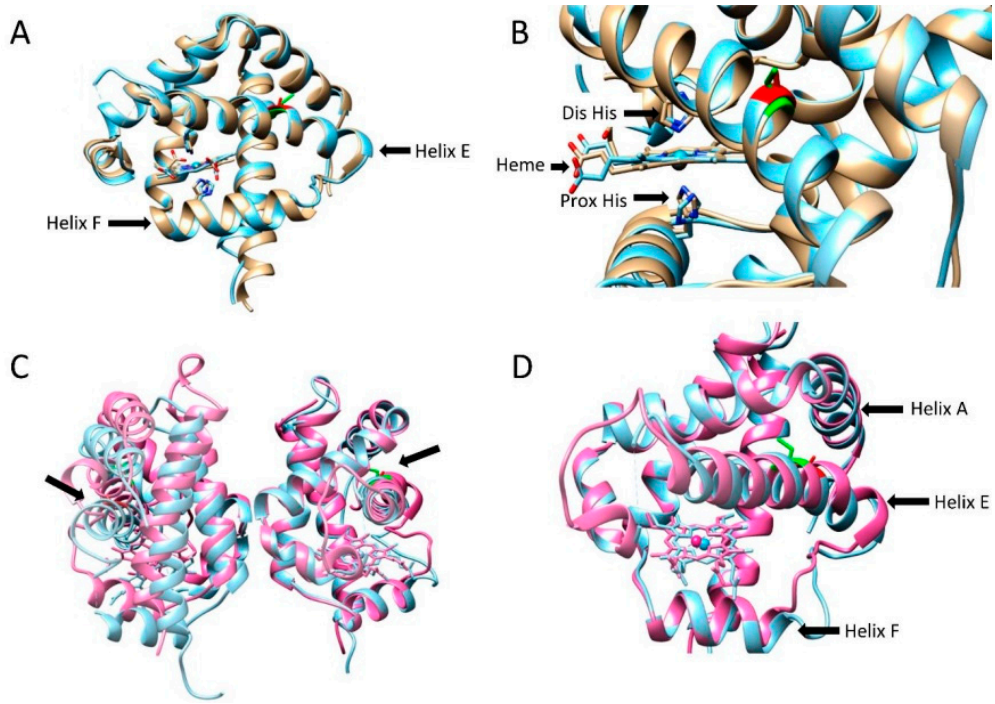


Figure 10. Crystal structure comparison between BvPgb1.2 WT, C86A and Barley nsPgb. **(A) and (B):** Comparison between BvPgb1.2 WT (PDB-ID: 7ZOS) (gray) and C86A (PDB-ID: 7Z1U) (blue), highlighting helices E and F, as well as the orientation of distal and proximal histidine. The cysteine (green) and alanine (red) are also indicated. **(C) and (D):** Comparison between C86A (blue) and barley nsPgb (PDB-ID: 2OIF [53]) (pink), highlighting cysteine and helix orientation. The same colors were used to depict the cysteine and alanine as previous panels. Adopted and modified from [95] (**Paper I**).

Since both side chains are pointing away from the heme group, a direct interaction with heme iron is less likely. However, electrons are often transferred from the heme iron to cysteine residues in moderate proximity [121]. The sulfhydryl side chains in cysteine residues are capable of acting like electron transporters. The cysteine can regenerate the reduced state if antioxidants are present in the environment, which is possible if the residue is facing the surface instead of the core of the protein. This is the case in BvPgb1.2, where the sulfhydryl group is facing the surroundings, in addition of being located 17.3 Å away from the heme iron [95] (Figure 11B and C). Even though it is not in close proximity, it is within the 19 Å distance limit needed to enable electron transfer on a millisecond timescale seen in other Hbs [122, 123]. This might indicate a flexibility in the E-helix in order for the cysteine to move closer to the heme iron if required. Hence, a mechanism where disruption of the hydrogen bond between the distal histidine and ligand could be facilitated by removal of the cysteine. This change in redox balance might cause solvent entry into the heme pocket and increasing the possibility of Fe^{3+} state formation. The flexibility of helix E is also discussed in Chapter 3 (**Paper II**).

Apart from cysteine residues, tyrosine residues also have important roles in electron transfer and they are often involved in redox activity of Hbs [124, 125]. In BvPgb1.2, a tyrosine is located at position 115, in close proximity to the heme iron (9.2 Å) (Figure 11A and C). Thus, we suggested a possible mechanism where these two residues (C86 and Y115) are important in maintaining a stable and regenerative redox environment in the heme pocket of this class I nsPgb. This could explain why the overall structure of the protein was not affected, but as described in the following section, the activity was majorly impacted.

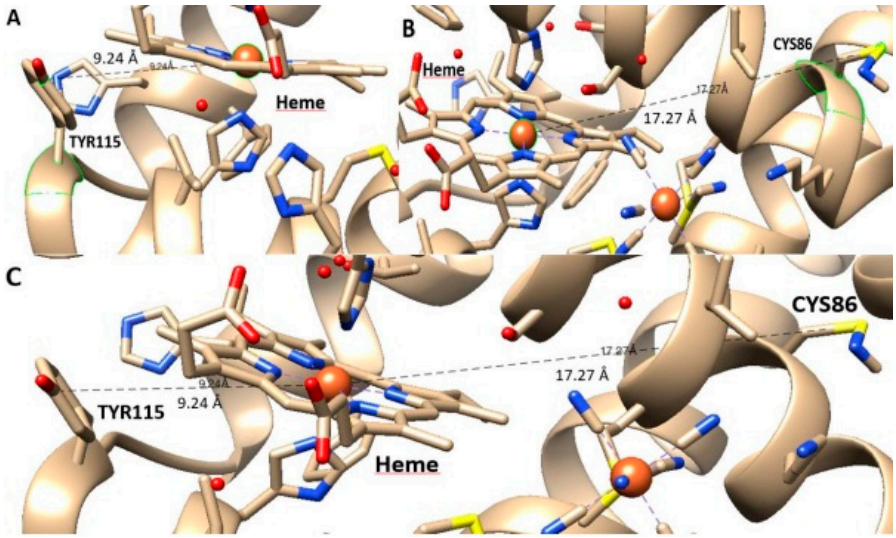


Figure 11. (A) and (C): Localization of Tyrosine 115 (Tyr115) (9.24 Å). **(B) and (C):** Cysteine 86 (Cys86) (17.27 Å). Both distances are in relation to the heme iron. Both residues are within observed distance for possible electron transfer in globins (<19 Å). Adopted and modified from [95] (**Paper I**).

2.2.2 Stability and Activity Implications of Cysteine-Substituted Mutant C86A

We also investigated if the inferred mutation had an impact in the heme stability of the protein. An effective way to assess the heme loss tendency is to use a heme scavenger. In this case, apomyoglobin H64Y/V67F (ApoMb) was used [121, 126]. The heme transfer to the ApoMb was detected at 600 nm, where the holoenzyme has its peak absorbance. To convey the difference regarding heme stability between penta- and hexacoordinated Hbs, human fetal Hb was used as a positive control [106] (Figure 12). The significant increase of the absorbance at 600 nm showed a functioning ApoMb and the tendency for the pentacoordinated Hb to release its heme to the scavenger, indicated by the arrows in Figure 12B.

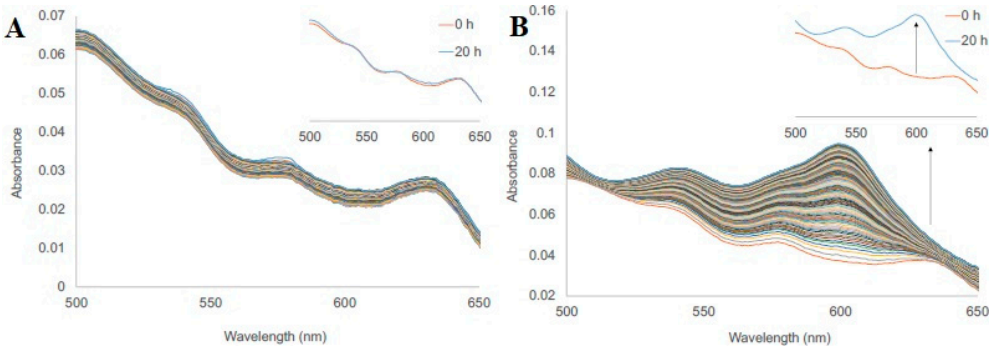


Figure 12. Heme loss assay for human fetal Hb using ApoMb as heme scavenger. **(A):** No ApoMb. **(B):** Added ApoMb. The increased absorbance at 600 nm indicates a functioning ApoMb. Both reactions where conducted at 25 ˚C for 20 h. Adopted and modified from [95] (**Paper I**).

BvPgb1.2 WT and C86A were examined under identical conditions (Figure 13), where (A) and (C) did not contain ApoMb, while (B) and (D) did. No increase of the characteristic absorption at 600 nm was observed for either of the proteins, corresponding to no heme transfer to the scavenger protein. To my knowledge, this is the first results showing this kind of heme stability for hexacoordinated Hbs in general and for nsPgbs specifically. This property is desirable in terms of biotechnical applications for these proteins since the loss of heme would lead to increased side reactions and oxidative damage.

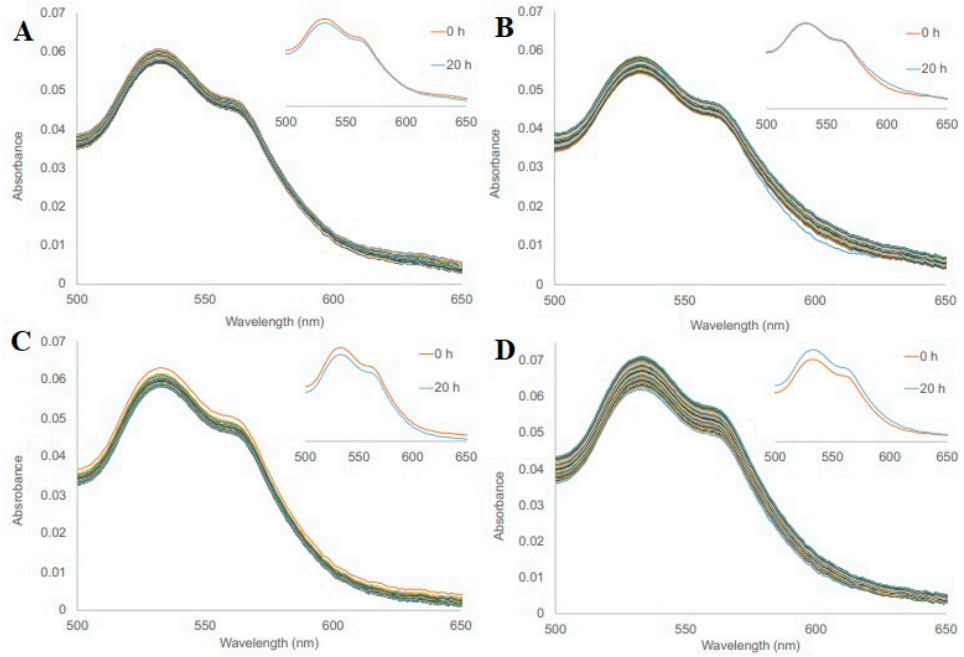


Figure 13. Heme loss assay for BvPgb1.2 WT and C86A using ApoMb as heme scavenger. **(A) and (B):** BvPgb1.2 WT without and with ApoMb, respectively. **(C) and (D):** C86A without and with ApoMb, respectively. No change in absorbance at 600 nm was observed, indcating intact nsPgb holoprotein. Adopted and modified from [95] (**Paper I**).

Another important aspect in terms of stability is the melting temperature (T_m). This temperature indicates when proteins denature due to the increase in thermal energy and this property has not been studied extensively for nsPgbs. We used nano-differential scanning fluorimetry (nanoDSF) to assess the T_m for BvPgb1.2 WT and C86A. The intrinsic fluorescence of the proteins at 330 and 350 nm were used to estimate the unfolding event during a gradual temperature increase (20-95 ˚C). In Figure 14, the results for these proteins incubated at 37˚C for seven days are shown. Both proteins showed high thermal stability, where the mutant displayed ~2.5 ˚C increased T_m compared to WT (88.5 vs 86˚C). Similar results were obtained with cyanide and O_2-bound forms. Proteins with T_m> 65˚C are considered to be thermally stable [127], which these proteins certainly were. Similar proteins with measured T_m:s are horse heart Mb (T_m= 76˚C) [128] and the thermally robust human Ngb (T_m= 100˚C) [127, 129]. These observations further strengthens the claim regarding the increased stability for hexacoordinated globins in relation to pentacoordinated ones.

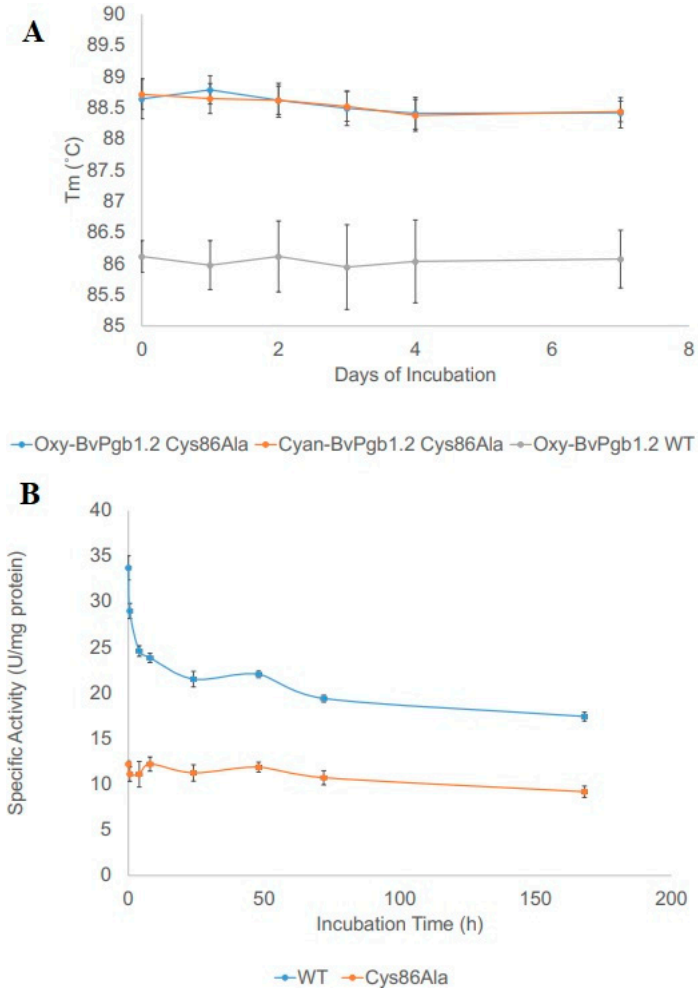


Figure 14. (A): Thermal stability (Oxy and cyano-forms) and **(B):** Peroxidase activity comparison between BvPgb1.2 WT and C86A at 37˚C. The proteins were incubated at 37 ˚C during seven days and stability/activity were measured at least once per day. Adopted and modified from [95] (**Paper I**).

Although the thermal stability was not affected to a great extent, the functionalities of the nsPgbs certainly were. By designing an assay based on previously published work [130], the peroxidase activity was estimated for both proteins. By adding H_2O_2, the proteins were oxidized from ferrous to ferric and ultimately to ferryl form ($Fe^{2+} \rightarrow Fe^{3+} \rightarrow Fe^{4+}$). The reactive ferryl oxidation state will react with the substrate ABTS, yielding the radical product A•, from the protonated form AH, which can be detected at 415 nm [131]. The reaction scheme is illustrated below, not showing the initial $Fe^{2+} \rightarrow Fe^{3+}$ where additional Fe^{4+} is generated:

$$\text{Heme-Fe}^{3+} + H_2O_2 \rightarrow \text{Heme}^{\bullet+}\text{-Fe}^{4+}\text{=O} + H_2O \quad (1)$$

$$\text{Heme}^{\bullet+}\text{-Fe}^{4+}\text{=O} + AH \rightarrow \text{Heme-Fe}^{4+}\text{=O} + A\bullet + H^+ \quad (2)$$

$$\text{Heme-Fe}^{4+}\text{=O} + AH + H^+ \rightarrow \text{Heme-Fe}^{3+} + A\bullet + H_2O \quad (3)$$

The specific activity for BvPgb1.2 WT decreased from 33.7 to 17.4 U/mg protein, while the mutant had a constant activity 9.2 U/mg protein during the seven-day incubation period (Figure 14B). Previous studies regarding the peroxidase activity for nsPgb-1 in rice (*Oryza sativa*) showed 86x lower affinity for H_2O_2 compared to horseradish peroxidase [132]. However, nsPgbs are capable to function as pseudo peroxidases, although it is not their primary function *in vivo*. These results indicated the implications of removing the conserved cysteine residue in terms of functionality, in accordance with the difference in autoxidation rate.

In summary, the substitution of the conserved cysteine residue in BvPgb1.2 resulted in dramatic effects of the functionally and activity of the protein. However, the structural integrity, thermal stability, dimerization and heme loss tendencies were less affected. The mutation altered the redox environment of the heme substantially and this work provided important knowledge regarding intrinsic properties for globins in general and for nsPgbs specifically. By using protein engineering in this way, new possibilities have opened up where these proteins can be utilized for biotechnical purposes. In Chapter 3, intra- and intermolecular interactions in BvPgb1.2 were investigated in more detail (**Paper II** and **III**).

3 Intra- and Intermolecular Interactions in Phytoglobins

3.1 Nuclear Magnetic Resonance Studies of Phytoglobins

3.1.1 Bacterial Adaption and Expression of Isotope-Labeled Phytoglobin

Investigations of protein dynamics and internal interactions during small timescales could be important to understand how proteins interact and fold into their final quaternary structures. This would provide crucial information in order to design proteins with desired properties but also insight into protein behavior and interactions *in vivo.* Even though the crystal structures for both BvPgb1.2 WT and C86A were solved (**Paper I**), this usually provides information of the structure as a snapshot in time. Thus, it is difficult to make claims regarding both dynamic intra- and intermolecular interactions. One option would be to label proteins with detectable isotopes and deploy 2D Nuclear Magnetic Resonance (NMR) to characterize these dynamic interactions.

NMR is a useful tool to investigate structure, dynamics and interactions, complementing the structural information gained from X-ray crystallography [133]. These internal motions are often essential for protein function and it is possible to study these in more detail using NMR relaxation methods [134, 135]. However, one major obstacle is the molecular size of the protein since interference increases with size, such as spectral overlap and line broadening. These problems can be mitigated by labeling proteins with stable isotopes, such as ^{13}C and ^{15}N in combination with deuterium (^{2}H, D_2O), reducing the interference and making it possible to study proteins in the 40-50 kDa range [136, 137].

The molecular weight of the BvPgb1.2 monomer is 19.2 kDa. Since this protein forms a homodimer in its active form and low concentrations (**Paper I** and **III**), the total molecular weight is below the normal size-limiting range mentioned above. This would make it possible to study the internal motions of the dimer if it could be expressed and produced with stable isotopes.

Due to the different properties of these isotopes, a specific protocol had to be developed for the *E. coli* to express the labeled and functioning protein in sufficient amounts [138]. This was achieved by slowly adapting the bacteria to grow in ^{2}H-containing minimal media by gradually increasing the concentration of the isotope to the final concentration of 99.8% D_2O [139]. This was complemented by adaption of the ^{13}C and ^{15}N isotopes. The induction of protein expression was initiated after the adaptation and adjustments of cultivation and expression times were made in accordance to the slow growth of the bacteria. The incubation time to reach optimal OD_{600} was ~4.5-fold longer for the labeled medium and the final yield was 2.5 mg protein/g cell weight; a 75% reduction compared to expression in the non-labeled medium. Although, neither the purification nor functionality of BvPgb1.2 were altered due to the isotopes. The amount of BvPgb1.2 was sufficient to obtain both cyano- and O_2-bound samples.

In order to perform high-quality NMR measurements, buffer conditions must be optimized for the protein. Like in **Paper I**, nanoDSF was used to evaluate protein stability in a range of different buffer salts, ionic strengths and pH-values. In the case of BvPgb1.2, 3-(N-morpholino) propanesulfonic acid (MOPS) showed maintained protein stability in low molar concentrations and deemed to be a suitable buffer. The final buffer conditions for the NMR experiments were 50 mM MOPS pH 7.2 and were used for both forms of BvPgb1.2 [138].

3.1.2 Assignment of Residues in BvPgb1.2

By obtaining ^{1}H-^{15}N TROSY spectra (Figure 15) from the cyano-bound BvPgb1.2, 137 backbone amides were sequence-specific cross-peak assigned of the 165 expected ones (83% coverage). Due to the special properties of proline residues, including lack of amide protons and low dispersion of chemical shifts [140], these were not included in the 165 expected ones. Furthermore, an additional 13 residues were partially assigned, where one or more of carbonyl chemical shifts could be seen, while the rest of the residues remained unassigned. The results are summarized in Table 3.

Table 3. Residues with partial/missing assignments and multiple conformations.

Type of Assignment	*Residues*
Partial Assignments	N5 P9 I16 I47 F48 T53 P67 T85 E102 K114 P138 P143 V157
Missing Assignments	M1 S2 F3 T4 P52 H112 L113 L129 L130 E131 T132 I133 A150 E151 A152 F153 N154 H155 L156 I160 P171
More than one confirmation	V6 N7 Y8 A10 S11 D12 G13 T14 V15 F17 V98 E144 M145 K146 N147 A148 W149 R166 L167 S168

In comparison, 48 of the expected 165 cross-peaks in the O_2-bound BvPgb1.2 could be assigned. This was probably due to the paramagnetic character of the Fe^{2+} oxidation state of the heme iron. This phenomenon is associated with unpaired electron spins, for example in transition metals and radicals [141]. This high-spin state when associated with O_2 could cause problems due to nuclear spin relaxation and unpaired electrons attracted to the external magnetic field, making it challenging to distinguish residue peaks from each other [142, 143]. Instead, the low-spin state for cyanide-bound BvPgb1.2 showed slower relaxation, yielding more separated peaks and better NMR spectra [144].

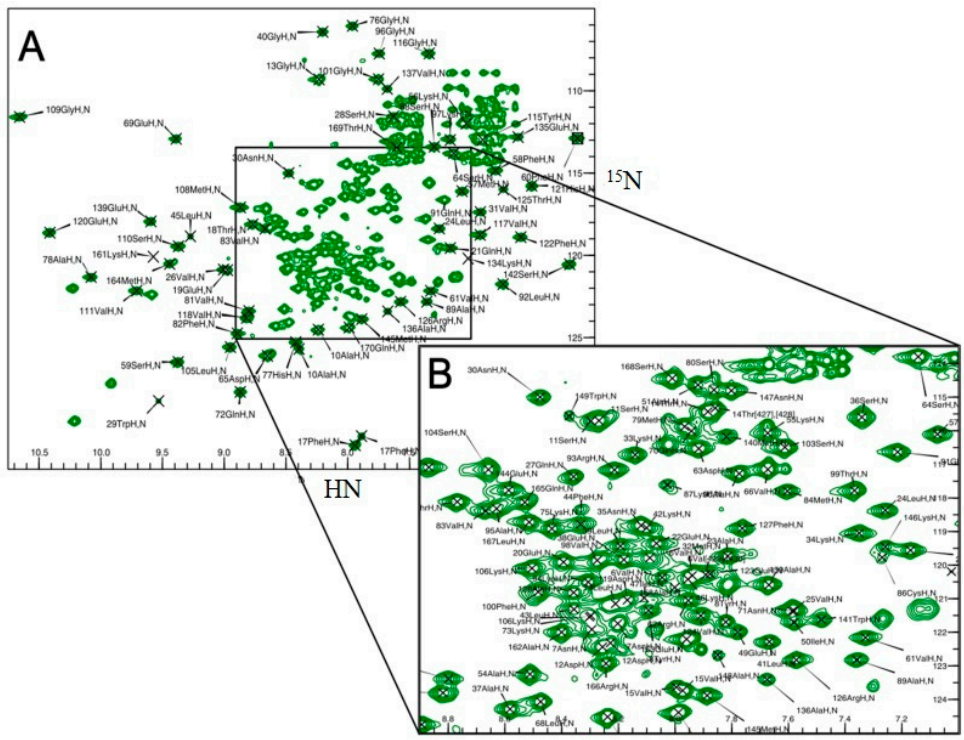


Figure 15. (A): ^{1}H-^{15}N TROSY spectrum for cyanide-bound BvPgb1.2. **(B):** Zoom-in depiction of residue-rich section of the spectrum. These spectra are a complement to the information in Table 3. The spectrum regarding the paramagnetic character of O_2-bound BvPgb1.2 was not included here. Adopted and modified from [138] (**Paper II**).

3.1.3 Dimerization and Flexible Regions in BvPgb1.2

The dimer dissociation constants (K_d) for similar class I nsPgbs from *Parasponia andersonii*, *Trema tomentosa*, and *Arabidopsis thaliana* (72.2, 74.1 and 74.4% sequence identity, respectively) were determined to K_d ~1 μM [48, 94, 145]. Using

the same constant for our model and samples, we predicted to have >97% dimeric BvPgb1.2 in the analyzed samples. This high degree of dimerization was consistent with the findings in **Paper III**, even at low protein concentrations (0.15 mg/ml). Furthermore, several of the residues missing assignments seem to be located in the G and H helices, which have been proposed to be involved in dimer formation [62]. Possible explanations could be a sliding motion of these protomers, which has been seen in the same region of different Hbs or ´tunnel´ formation for ligand diffusion [146, 147]. Based on our observations, we suggest that there might be two distinct processes taking place in this region: transient dimerization between the subunits and local conformational exchange between or within the helices at different timescales.

The dimerization and flexible regions were further characterized using ^{15}N relaxation to study conformational dynamics in BvPgb1.2 Here, the longitude (R_1) and transverse (R_2) relaxation rate constants were determined, which probe motions from picosecond to millisecond timescales. Nuclear excited states will tend to display R_1 and R_2 relaxation towards equilibrium, where R_1 and R_2 describe the recovery of net magnetization and loss of coherent magnetization among the collection of spins, respectively [148]. In other words, the net magnetization after the frequency pulse will be tilted to the XY-plane instead of the Z-plane and the succeeding decrease in the XY-magnetization vector is defined as T_2 (transverse relaxation time). Meanwhile, the Z-plane magnetization vector will increase again, defined as the T_1 (longitude relaxation time). By measuring T_1 and T_2 at different delay intervals, it is possible to determine the corresponding relaxation rates as described above, where R_1 and R_2 relate to respective time constants of relaxation in an inverse relationship [148].

The relaxation quota R_2/R_1 indicate increased flexibility (low values) for 25 residues, namely 28, 34–35, 63–66, 99, 101, 104, 119 and 124 other than the highly flexible N- and C termini (Figure 16). The high flexibility may have several implications, including being positioned in surface-exposed areas and/or end of helices but also where electron density information was missing from the structure determination. Moreover, the localization of V124 (helix G) indicated its involvement in close-contact hydrophobic interactions with the opposing subunit (**Paper III**).

Residues with higher R_2/R_1 included V81, M84, K87, S88, Q91 and R93. Most of these residues are located on the same side of helix E, separated on turn of helix away from each other. Moreover, V81 is located one turn of helix away from the heme-coordinating H77. A similar pattern was observed for residues M108, S110 and V111 in helix F, which also harbors H112, the other heme-coordinating residue. Furthermore, residues V81, M84, M108, V111, and F122 have their side chains lining the heme-binding pocket. Taken together, these results might suggest that the heme-binding pocket is undergoing some type of breathing motion, accompanied by transient bending or fraying of the neighboring helices. In addition, comparison

between the NMR-data for cyanide-bound BvPgb1.2 and the crystal structures for both BvPgb1.2 WT and C86A showed agreement with the dimeric model [138].

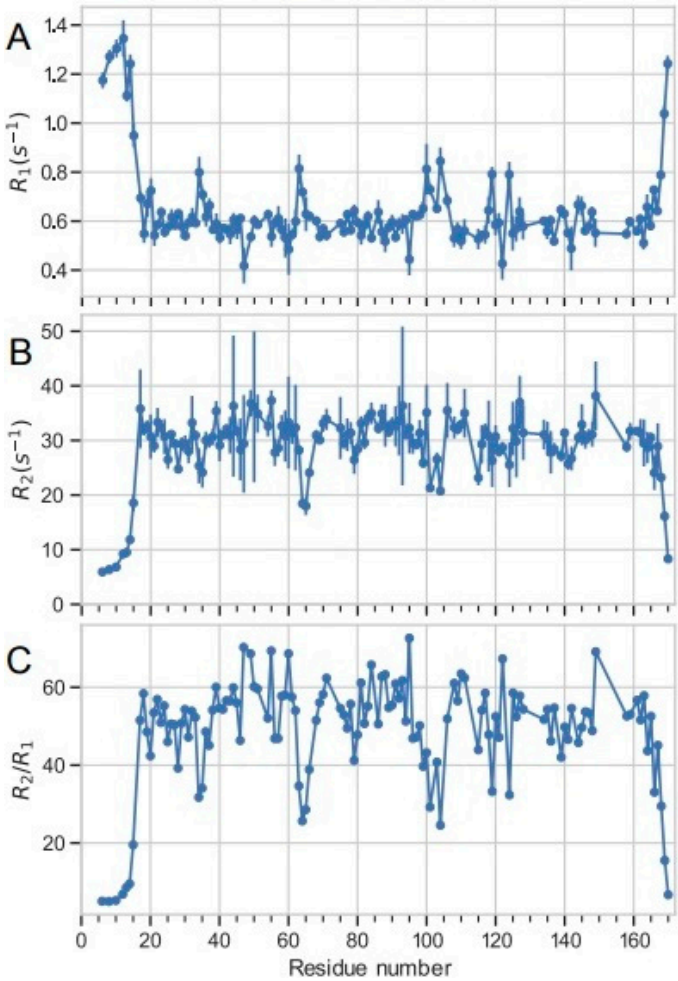


Figure 16. ^{15}N relaxation for residues in BvPgb1.2 showing high-quality data. **(A):** R_1, **(B):** R_2 and **(C):** R_2/R_1. In total 130 residues showed high quality relaxation data and used in the analysis, while seven (*N7, S11, Q72, K73, L74, L105 and E123)* residues were excluded due to spectral overlap. Adopted and modified from [138] (**Paper II**).

Based on the structural determination in **Paper I** and subsequent dynamic analysis in **Paper II**, we proposed the following dimer interactions in BvPgb1.2: hydrogen bonding primarily between T53, E120 and/or E123, in addition to a hydrophobic cluster possibly involving I47 and/or I50, F48, A51, V124 and F127, all interacting with the opposing subunit. These residues are essentially located in the BC loop and G helix, which is in accordance with previous observation in rice and barley nsPgbs [53], even though the precise locations of these residues are altered. In addition, a

T47-substituted mutant in nsPgb-1 from *Arabidopsis* showed a substantial increase in monomer content in comparison to the WT-protein [145]. This implicates the important hydrogen bonding for the corresponding T53 in BvPgb1.2, adding to the previous knowledge of dimeric interactions. These interactions are illustrated in more detail in Section 3.2 (**Paper III)**, where dimeric structure predictions using AlphaFold-software are presented.

We also investigated the dynamic features for the proposed residues important in electron transfer to maintain proper redox environment, namely C86 and Y115 (**Paper I**). The homodimer in C86A implicated the electrostatic and hydrophobic interactions in the dimer interface between helices B/C and helix G [145], applicable for the WT as well due to the monomeric resemblance. However, the relaxation rate for these residues did not stand out from the average, but still might be affected by the dynamic motions seen in helices E and F. It is plausible that bending of helix E would affect the solvent accessibility of the C86 sulfhydryl group and thus influencing its oxidation state. Similarly, conformational dynamics of helix F might well modulate the interactions between the Y115 side chain and the heme. Previous work regarding the ligand mechanism of nsPgbs has shown a ´piston´ movement of the E-helix caused by rearrangement loops of helices CD and EF [53]. This structural change gives rise to the dislocation of the distal histidine, allowing ligands to bind. Similar results were observed in **Paper II**, confirming the flexibility of helix E.

3.2 Structure Predictions and Evaluation of Oligomer Formation in Phytoglobins

One of the aspects regarding the implications of substituting the conserved cysteine residue analyzed in **Paper I** was its effect on dimerization. Here, it was found that the C86A mutant maintained its dimeric form at low protein concentrations (1 mg/ml) when investigated with size exclusion chromatography (SEC) [95], seen in Figure 17. The elution peak for both BvPgb1.2 WT (A) and C86A (B) were located at ~10.9 elution volume, corresponding to a molecular weight of ~39 kDa. If the monomer (19.2 kDa) was present, its peak would be detected ~12.7 ml elution volume. However, even though these results suggest a high degree of dimerization for both proteins, the tailing of the curve may implicate dynamic and interactive events occurring between the subunits. Thus, it would be of interest to further investigate the interactions in the proteins leading to this elution behavior.

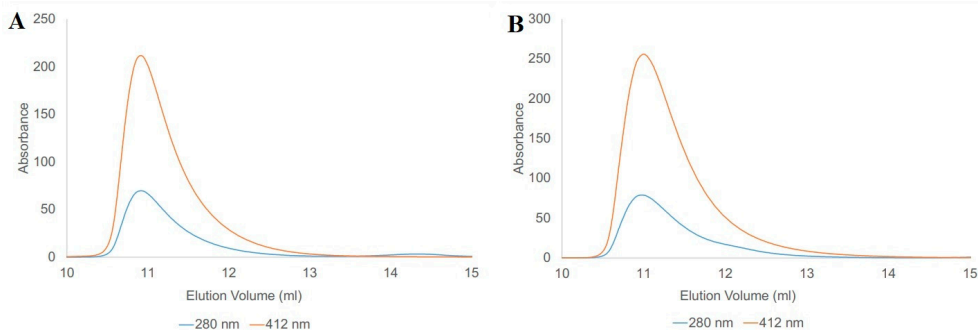


Figure 17. (A): Size exclusion chromatography of BvPgb1.2 WT and **(B):** C86A**.** The elution peak at 10.9 ml corresponds to a ~39 kDa protein, indicating the dimeric structure for both WT and mutant. Monomeric proteins were not observed for low protein concentrations (1 mg/ml). Adopted and modified from [95] (**Paper I**).

An attempt to examine the dynamic motions of the cysteine was made in **Paper II**, but its relaxation rate was not distinguishable from the average dynamic motions. Even though the results suggested fraying of the E-helix that might facilitate ligand accessibility to the heme pocket and maintain redox balance, little information regarding C86 in particular was found. Therefore, it would be of interest to use additional techniques to study these interactions in detail. One option would be to use small-angle X-ray scattering (SAXS) to study structural differences between BvPgb1.2 WT and C86A. In addition, structure prediction simulations would be a suitable tool to map the residues lacking electron density information in the crystal structures of BvPgb1.2 WT and C86A. Moreover, the dimeric structure of C86A was used to simulate and investigate the specific interactions involved in dimer stabilization (**Paper III**).

During recent years, AlphaFold developed by DeepMind, has become a powerful tool in predicting protein structures based on algorithms from known structures of similar proteins [149]. Since 26 amino acids were missing from the crystal structure of BvPgb1.2 WT (**Paper I**), it would be beneficial to estimate their predicted locations using AlphaFold. The missing residues were located in highly flexible regions, such as N- and C-termini and loops between helices CD/DE. The conducted simulation can be seen in Figure 18. Here, the structure of BvPgb1.2 WT and best model are presented in gray and blue/yellow, respectively. The dark blue color of most helices indicated high confidence positioning, while the light blue (CD/DE loops) and yellow (N- and C-termini) showed more uncertain locations of these residues. This was also evident in sequence identity summary (Figure 18B), where the model compared ~2500 related sequences for the most conserved regions but less for the termini and highlighted loops, as well as the EF-loop region (residues 97-103). The simulation had a low root-mean-square deviation (1.15), indicating an overall reliable model. Similar results were also observed for the C86A-based model were the remaining 24 residues were added to the structure (**Paper III**). The

resemblance indicated no/minor implications of the conserved cysteine on tertiary structure, in accordance with previous results (**Paper I**). The additional residue positions of the model are marked in the primary amino acid sequences in Table 4 (green):

Table 4. Primary amino acid sequence of BvPgb1.2 WT and added coverage of residues in AlphaFold (green).

#1/A	1	MSFTNVNYPASDGTVIFTEEQEALVVQSWNVMKKN
#1/A	36	SAELGLKLFLKIFEIAPTAKKMFSFVRDSDVPLEQ
#1/A	71	NQKLKGHAMSVFVMTCKSAAQLRKAGKVTFGESSL
#1/A	106	KHMGSVHLKYGVVDEHFEVTRFALLETIKEAVPEM
#1/A	141	WSPEMKNAWAEAFNHLVAAIKAEMQRLSTQP

Although, the precise positions of these residues might not be of great importance in terms of protein function, this model still proposes a complete protein folding with high probability, which is desirable. However, other dynamic simulations have linked the CD-region to increased thermal stability of globins, counteracting the resulting fluctuations from the thermal energy [129]. This high flexibility could contribute to quaternary state stability and prevent denaturation processes. Thus, using modern software based on artificial intelligence predictions, like AlphaFold, to predict protein structures may be more common in the future and would greatly decrease time and resources in comparison to X-ray crystallography, NMR and cryogenic electron microscopy methods or at least be a complement to these techniques.

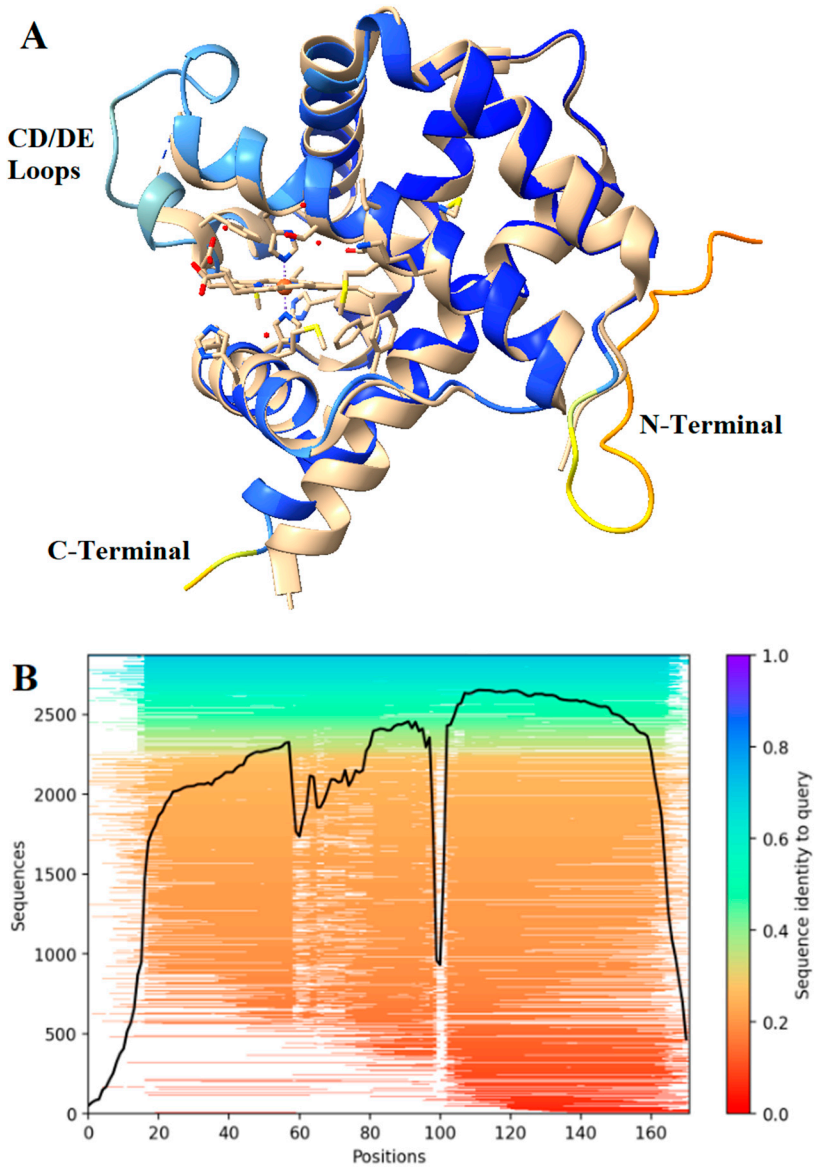


Figure 18. (A): Crystal structure of BvPgb1.2 WT (PDB-ID: 7ZOS) (gray) and simulated model using AlphaFold [149] (blue/yellow). **(B)** Sequence identity summary for AlphaFold model. Less sequence comparisons were found in the flexible the N- and C-termini and CD/DE/EF-loops. The low root-mean-square deviation (1.15) indicated a high-confidence simulated model.

AlphaFold could also be used to investigate detailed interactions involved in dimerization between the monomers, in addition to the crystal structures and protein dynamics analysis (**Paper I** and **II**). Here, the dimeric structure of BvPgb1.2 C86A was used to predict the structures of opposing monomers forming the homodimer

(Figure 19). In Figure 19A, the complete structure predictions of both monomers are illustrated. The N-termini were located on opposite ends of the dimer, facing the solvent interface, while close contacts were mainly located in the BC helices with connected loop and helix G. Zoom-in of the dimer interface (Figure 19B) revealed hydrogen bonding (dotted lines) and hydrophobic interactions between aliphatic residues in close proximity. Hydrogen bond formation was primarily occurring between T53 and E123 in opposing subunits. In this figure, E120 was interacting with residues in the G helix itself, contributing to helix stabilization, even though it could possibly form a hydrogen bond with T53 as well.

Close contacts were observed between I50, A51 (BC loop) and V124, F127 (helix G) in opposing dimers. These are probably the most important residues involved in the hydrophobic cluster, stabilizing the dimer formation. Furthermore, I47 and F48 are located in helix B, further away from the dimer interface. Hence, these residues might not interact directly with the opposing subunit but have stabilizing effects in helix B and surrounding helices. These results are in line with previous reasoning regarding dimer interactions [145] and that the cysteine-substitution in the mutant did not affect the quaternary structure of BvPgb1.2 in terms of dimeric interactions.

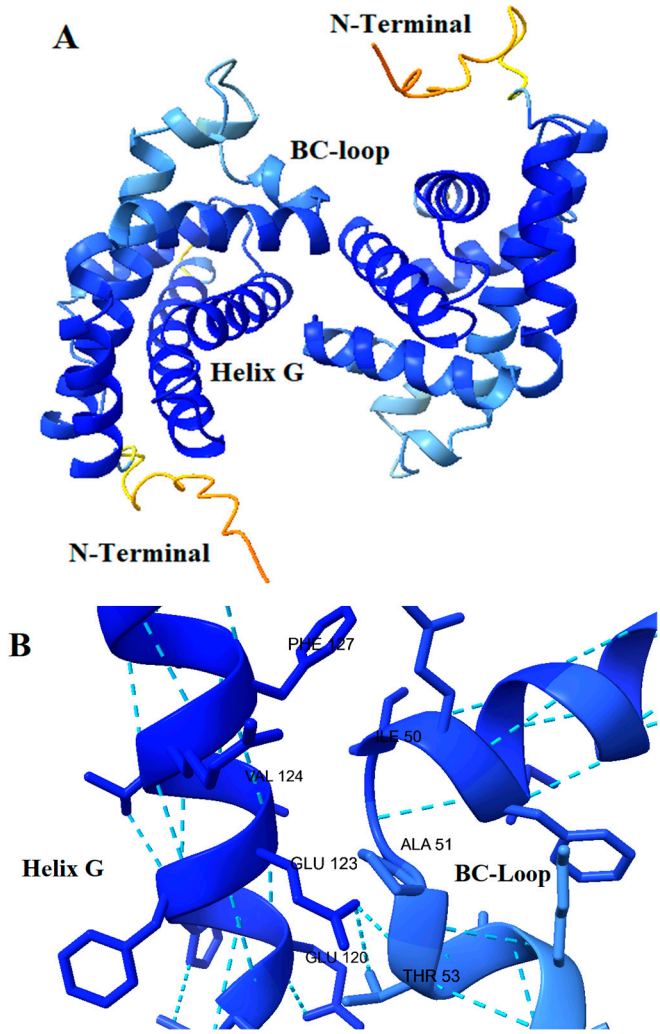


Figure 19. (A): Structure prediction of complete BvPgb1.2 C86A (PDB-ID:7Z1U) homodimer using AlphaFold [149]. The BC-helices, connecting loop, helix G and N-termini are depicted. **(B):** Zoom-in of dimer interface. Important residues for hydrogen bonding (T53 and E120/123, dotted lines) and close-contact hydrophobic interactions (I50, A51, V124 and F127) were observed. The three letter abbreviations in Figure 19 are equivalent to the one-letter counterparts used in the description.

As previously mentioned, an additional tool of importance regarding oligomeric states analysis and quaternary structures of proteins is SAXS [150]. Here, the theoretical models could be used as a complement to the experimental information from the SAXS measurements to provide valuable information of the protein quaternary structure in solution. Thus, by combining the theoretical results from the structure predictions, based on the crystal structures, with the SAXS analysis would

provide additional information regarding the degree of dimerization and oligomerization.

SAXS analysis has been used in previous studies of Hbs. For instance, structure integrity and stability imposed by modifications of residues on the surface of human fetal Hb were analyzed using this technique, which demonstrated maintained quaternary structure in relation to the WT protein [151]. To my knowledge, a comprehensive SAXS analysis has not been conducted for nsPgbs and would provide new information regarding dimeric and oligomeric tendencies in the BvPgb1.2 WT protein, as well as the implications of the cysteine residue substitution in C86A.

The dimeric structures for both WT and C86A at high/low protein concentrations were confirmed (ranging from ~19 to 0.15 mg/ml). The monomer content was estimated to range between 0-2% and 4-6% for WT and C86A, respectively, for the lowest protein concentration. This also included added salt concentrations (0-150 mM NaCl), indicating that ionic interactions from the buffer only had minor effects on the stability of the quaternary structure or oligomer formation. Furthermore, the radius of gyration (R_g) was determined to range between 22-24 Å for both proteins and was independent of salt concentration, further confirming the structural similarity between WT and C86A.

Although the dimeric configuration was confirmed, the observed oligomeric distribution showed different tendencies for the proteins and were dependent on protein concentration. The size distribution ranged between 39-48 kDa for the lowest to highest concentrations of BvPgb1.2 WT, in accordance with the 38.4 kDa dimer. In contrast, C86A ranged between 48-63 kDa, indicating higher order of oligomerization, especially for the higher protein concentrations. This is illustrated in Figure 20, where the association number for the proteins are presented as a function of protein concentration. Here, one association number equals the molecular weight of the monomer (19.2 kDa). Once again, the addition of salt (150 mM NaCl) only had minor effects regarding oligomer size distribution.

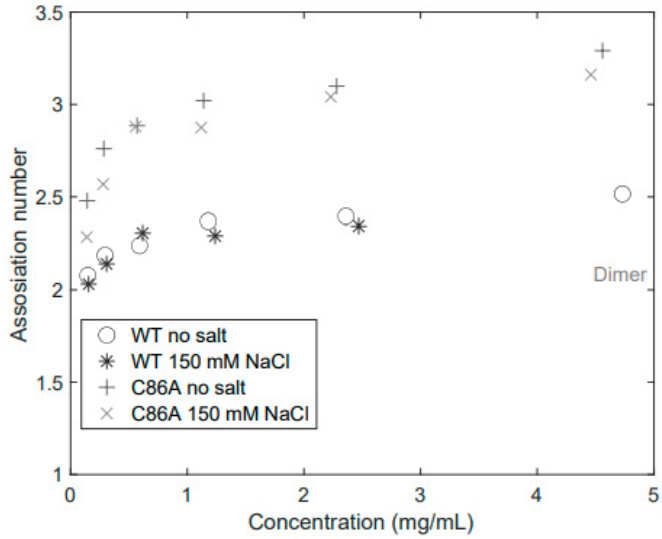


Figure 20. Association number of BvPgb1.2 WT and C86A as a function of protein concentration, where 1= 19.2 kDa (molecular weight of monomer). The molecular weight of the dimer is equivalent to association number 2. Size distributions for WT and C86A were between 39-48 kDa and 48-63 kDa, respectively. The last point for WT+150 mM NaCl (4.6 mg/ml) was not included due to radiation damage.

Even though the conserved cysteine residue seemed to be less important in quaternary structure stabilization, the substitution to alanine might offer different properties in terms hydrophobic interactions, facilitating oligomeric association. As mentioned before, these observations stands in contrast to previously reported results in barley nsPgb, where cysteine removal showed up to 50% monomer content [63]. Although, the substitution to serine, as in barley nsPgb, would give rise to other interactions compared to the hydrophobic alanine, the corresponding significant increase in monomer content was not observed in BvPgb1.2, even at low protein concentrations.

Furthermore, initial trials using size exclusion chromatography in combination with SAXS (SEC-SAXS) showed one major peak corresponding to the elution volume of the expected dimer molecular weight, while no/low amounts of monomer were detected, similar to Figure 17 (preliminary results not included in **Paper III**). In summary, the vast majority of both proteins displayed dimeric populations even at low protein concentrations, indicating minor implications of the cysteine-substitution in dimer formation for BvPgb1.2.

In summary, this chapter focused on the overall dynamic motions and interactions in BvPgb1.2 in particular, and the impact of removing the conserved cysteine residue regarding the quaternary structure of this class I nsPgb (**Paper II** and **III)**. In Chapter 4, the potential practical features of BvPgb1.2 are highlighted (**Paper IV** and **V**).

4 Practical Applications of Phytoglobins

4.1 Encapsulation of Phytoglobins

In order to utilize BvPgb1.2 and other Pgbs for practical applications, whether it is as iron supplements or O_2 therapeutics, these proteins would preferably not be administered directly in free form. Encapsulation of proteins in nanoparticles is a viable option, providing several benefits. Encapsulation can provide increased stability against denaturation and protease digestion, maintaining functionality and activity, improve solubility and absorption in tissues, as well as decrease degree of interfering moisture absorption [152, 153].

In **Paper IV**, we encapsulated BvPgb1.2 WT and Mb in a highly swollen lipid sponge phase system, previously used to efficiently encapsulate β-galactosidase [154]. The nanoparticles contained a mixture of mono-, di-and triglycerides of glyceryl oleate and diglycerol monooleate, stabilized with polysorbate 80 (P80). The advantages with this formulation are the high production volume under mild conditions without organic solvents, having a well-defined internal structure but capable of self-dispersing and being regarded as food grade [154]. In addition, we wanted to investigate the difference in encapsulation efficiency for these penta- and hexacoordinated proteins and what implications the iron coordination may have, in addition to other factors such as precipitation tendencies, buffer effects, pH and pI.

Bulk phases of the proteins were made with two different lipid compositions, where P80 constituted 25 or 30% of the total lipid content. BvPgb1.2 displayed the characteristic red color of the Fe^{2+} state, while Mb mostly appeared as brown, especially in the 30% P80 formulation, indicative of the Fe^{3+} state. Correlations between protein concentration and phase composition were also observed and analyzed with SAXS. Different phase behavior for the two lipid compositions were detected, where the characteristic sponge phase were most predominant in the 30% P80 samples. Moreover, the water channel diameter remained reasonably constant for both proteins with 30% P80 (9 nm), while higher Mb concentrations seemed to decrease the diameter for 25% P80.

An example of the encapsulation is shown in Figure 21, displaying Mb (A) and BvPgb1.2 (B) dispersions in H_2O. A visual difference between the nanoparticles

was observed, going from 0 mg/ml (left) to 60 mg/ml (right) of loaded protein, similar to the bulk phases in terms of oxidation states. However, precipitated particles were observed for BvPgb1.2 to a higher extent compared to Mb, indicating a higher aggregation tendency.

When the buffer effect was analyzed with dynamic light scattering (DLS), a different behavior was observed for the Mb dissolved in 50 mM Tris-HCl pH 8.5. While dispersions in H_2O had similar shapes for all protein concentrations, a trend was observed for Mb-Tris dispersions with increasing protein concentration. This could arise from the contribution of larger objects, for instance from sedimentation. In contrast, opposite behavior was observed for BvPgb1.2, where aggregation occurred in both Tris-HCl and H_2O, but was exacerbated in the latter solvent. In either case, a clear difference between buffer and H_2O effect regarding encapsulation and aggregation tendencies for both proteins were observed. Although, low buffer concentrations (10 mM Tris-HCl) were sufficient to prevent these tendencies for BvPgb1.2 in solution and appeared to be independent of added salt (**Paper III**), interactions from the lipids may add to the complexity of this encapsulating system.

As stated above, contribution from lipid interactions (without protein) would also affect the features of the nanoparticles and encapsulating capacity in the different solvents. In general, dispersions using Tris-HCl showed less well-defined and lower intensity peaks compared to H_2O dispersions, even without loaded protein. This highlighted the buffer effect in this lipid system and indicated that the presence of Tris-HCl could affect the order in the lipid bilayer, having an effect on lipid self-assembly.

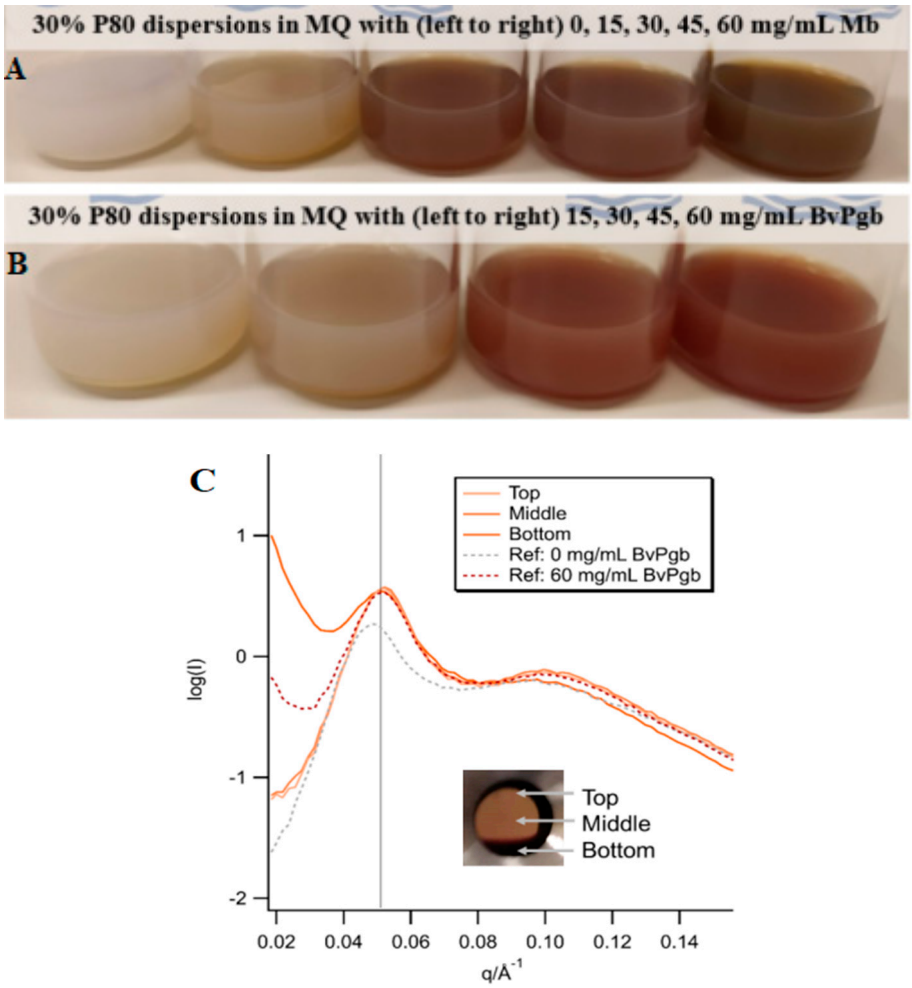


Figure 21. (A): Mb encapsulation and **(B):** BvPgb1.2 encapsulation in milliQ (MQ) H_2O and 30% P80. The protein concentrations ranged from zero or 15 mg/ml (left) to 60 mg/ml (right) of the loaded protein. Aggregation tendencies were observed for BvPgb1.2 to a greater extent. **(C):** SAXS measurments for BvPgb1.2 poor/rich (60 ml/ml in 50 mM Tris-HCl pH 8.5) showing phase separation in 30% P80 encapsulation.

Furthermore, peculiar behavior of the BvPgb1.2 particles was observed in terms of phase separation/sedimentation. As can be seen in Figure 21C, three different phases were detected. The phase separation showed accumulation in a protein rich region (bottom) and two sections with less protein concentrations (top and middle). The SAXS analysis showed that the sponge phase internal structure was maintained in all regions and when tilted horizontally, the protein/lipid matrix returned to the pre-separation condition. This was an interesting finding and might indicate reversible distribution of BvPgb1.2 in the bulk phase. This implied the transient behavior of the aggregates (possibly induced by confinement), weak interactions with the

internal structure and/or that these could exist alongside the sponge phase structure without disrupting it.

As previously mentioned, the aggregation tendencies for BvPgb1.2 were more profound in comparison to Mb in both lipid compositions. This behavior was correlated with protein concentration, where larger aggregation particles were formed in higher concentrations. These particles had to be filtered before the solutions were analyzed with SEC, resulting in removal of the majority aggregation. The results for 60 mg/ml-loaded particles in 30% P80 can be seen in Figure 22, showing Mb (A) and BvPgb1.2 (B) particles. The intact nanoparticles eluted ~7 ml in both cases, indicating the maximum separation capacity of the column. Leakage/excess Mb was detected at ~12.8 ml, corresponding to the monomer protein (16.9 kDa). This was not observed for the BvPgb1.2 encapsulation, where the expected elution volume would be ~10.9 ml for the dimer (see Figure 17). In addition, a minor peak ~11.9 ml was detected (Figure 22B), presumably a minor contamination from the encapsulation, rather than monomeric or dimeric BvPgb1.2. However, the protein aggregation certainly effected the amount of encapsulated BvPgb1.2, probably resulting in less encapsulated protein in the applied samples.

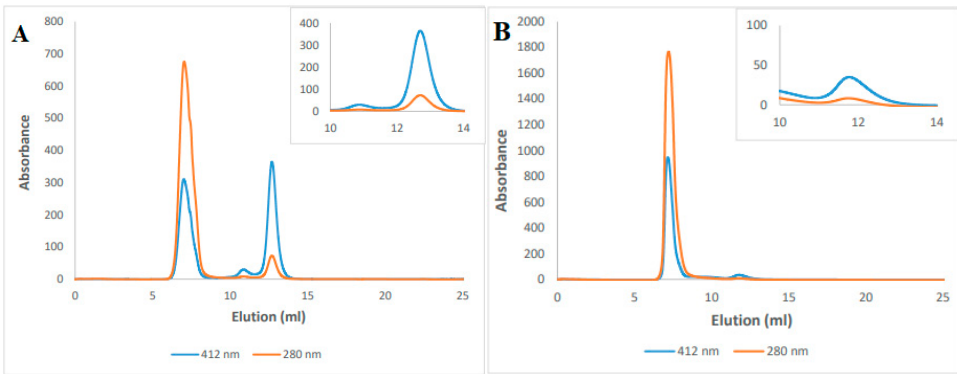


Figure 22. (A): Size exclusion chromatography of Mb and **(B)** BvPgb1.2 encapsulated nanoparticles (60 mg/ml) in 50 mM Tris-HCl pH 8.5 buffer. Free protein was observed for Mb (12.8 ml), indicating leakage and/or excess protein. No free BvPgb1.2 was detected.

Aforementioned, the largest aggregates were detected for BvPgb1.2 when high concentrations were used but seemed to be reversible. Upon agitation, the protein appeared to diffuse through the sponge phase as in the pre-separated state while maintaining the matrix structure. This indicates a transient behavior of the aggregates, showing weak interactions and/or the possibility to coincide with the sponge phase structure. In addition, the degree of aggregation might be linked to the pI of the proteins. The predicted pI:s for BvPgb1.2 and Mb were 7.85 and 6.8 [155], respectively. This could render BvPgb1.2 more sensitive to minor pH changes and

cause more extensive aggregation, especially if the pH is decreased. In this case, it is possible that the local pH drop close to/below the pI could induce aggregate formation.

Another potential implication might be the difference in iron coordination between BvPgb1.2 and Mb. Even though BvPgb1.2 showed high functional, thermal and structural stability (**Paper I**), this might not provide sufficient advantages in this lipid system, mainly constituted of unsaturated oleate. Heme-containing proteins are known to interact with lipids and display different interactions with saturated/unsaturated fatty acids [25, 156]. It is possible that disfavored interactions between the lipid sponge phase system and/or BvPgb1.2 occur due to the maintained Fe^{2+} oxidation state, in contrast to the Fe^{3+} state seen in Mb. Although, the stabilized Fe^{2+} state in the hexacoordinated nsPgb would potentially be less reactive compared to pentacoordinated globins. Furthermore, the formation of other internal structures, such as vesicles, might also contribute to different interactions between the monounsaturated lipids structure and proteins, which has to be investigated in more detail.

Finally, this study demonstrated the overall complexity of this system, with synergistic/competing effects of protein, buffer and lipid effects on aggregate size and self-assembled lipid structure. Even though characterization of bulk phases and dispersions were conducted in **Paper IV**, more actions are needed to investigate these interactions in order to facilitate efficient protein loading and delivery using similar lipid sponge phase systems.

Other Hbs have been successfully encapsulated in previous studies. For instance, vesicles from a mixture of saturated phospholipids, polyethylene glycol-conjugated phospholipids, cholesterol and negatively charged lipids have been used to efficiently encapsulate high concentrations of human Hb [157, 158]. Here, Hbs are normally saturated with CO to stabilize the protein, while this might not be needed for nsPgbs due to their high structural and redox stability. Moreover, efficient nanoparticle formation of bovine Hb through cross-linking reactions has also been reported. By co-precipitating Hb and $CaCO_3$, followed by cross-linking using glutaraldehyde and final dissolution of $CaCO_3$, formation of spherical Hb- particles with high O_2 affinity and low immunogenicity have been reported [159]. Similar methods have also been used to polymerize Hb and coat the exterior of the particles with human serum albumin, providing an efficient and cost effective technique while producing particles with desirable features, as previously mentioned [12, 160]. Thus, these methods might be suitable options to encapsulate BvPgb1.2 to provide a stable and low immunogenic formulation here as well. Since the current literature of nsPgb encapsulation is limited, additional studies must be conducted to find suitable techniques regarding efficient nanoparticle formation for this type of globin.

4.2 Toxicological Studies of Phytoglobins in Zebrafish Model System

Another important aspect regarding the applicability of nsPgbs in biotechnical or biomedical innovations are the potential toxic effects they might have in biological systems. Even though these proteins can be encapsulated or protected in other formulations (**Paper IV**), potential toxicological implications could still arise. Since Hbs carry oxygen and other reactive ligands, it is possible to cause oxidative stress in tissues due to formation of ROS or other reactive molecules. Another reason is the reactive nature of the heme moiety itself and its ability to go between several oxidation steps, as previously described. If this process is not controlled, for instance through redox stability and/or antioxidant activity in the surrounding environment, this redox imbalance could cause additional oxidative damage. Thus, it is crucial to test the biological effects these proteins might have before they can be safely administered.

As mentioned before, class I nsPgbs are hexacoordinated hemeproteins which could be an advantage due to the high overall stability and low heme loss tendencies (**Paper I**). To investigate their biological impact, BvPgb1.2 WT and C86A were investigated using a zebrafish (*Danio rerio*) animal model. For comparative reasons, the pentacoordinated Mb was also used to evaluate the toxicological effect it may have (**Paper V).**

Zebrafish as a model system has been used as a relatively simplistic animal model to study different toxicological effects of different proteins and compounds. It has several benefits, such as its relative low cost, number of fishes that could be analyzed and similar anatomical/biological features as humans and other vertebrates, including the hematopoietic system (generation of blood cells) [161, 162]. Human hematopoietic pathologies have been studied in zebrafish by genetic engineering. Information regarding specific mechanisms on a cellular level and identification of possible targets to combat these conditions have given valuable insights in the design of future therapies [161]. In particular, erythropoiesis (generation of red blood cells) is conserved to a substantial degree between mammals and zebrafish, making it a suitable model to genetically study related human diseases [163]. Since heme and Hb are synthesized during this process, the heme and iron metabolism also show great resemblance and studies in zebrafish have provided important information for new therapies in this area [163, 164].

Similar to humans, zebrafish also expresses several hexacoordinated globins and have been used as a model system for human Ngb [165]. Ngb from zebrafish has approximately 50% sequence identity to the human version and does not possess similar inhibitory effect on G proteins as the human counterpart [165]. Moreover, a recently discovered Cgb in this fish has been linked to normal development of vertebrate embryos, especially in proper cell differentiation for neural crest-derived

tissues and organs [166]. Therefore, even though the functionalities of these hexacoordinated globins are not fully understood, they seem to have important implications in vertebrate development.

Due to the conserved iron metabolism and presence of similar hexacoordinated proteins between humans and zebrafish, it would be an appropriate initial model system to study the effect of nsPgbs. Thus, to evaluate the toxicological effects of the proteins, BvPgb1.2 WT, C86A and Mb were fluorescently labeled. This would enable detection of any protein-marker accumulation and where it would be situated in the fishes, as well as any associated oxidative stress in the tissues. One alternative is to use fluorescent markers equipped with an activated ester group. It is possible to link these markers to lysine residues located on the surface of the proteins, forming stable amide bonds [167]. BvPgb1.2 had 16 more or less surface-exposed residues in its tertiary structure (Figure 23A), making it suitable for this kind of modification. Similarly, 18 lysine residues were found in the Mb structure. In contrast, four lysine residues were found in human Ngb [98], further indicating the high variability in primary sequence found in globins.

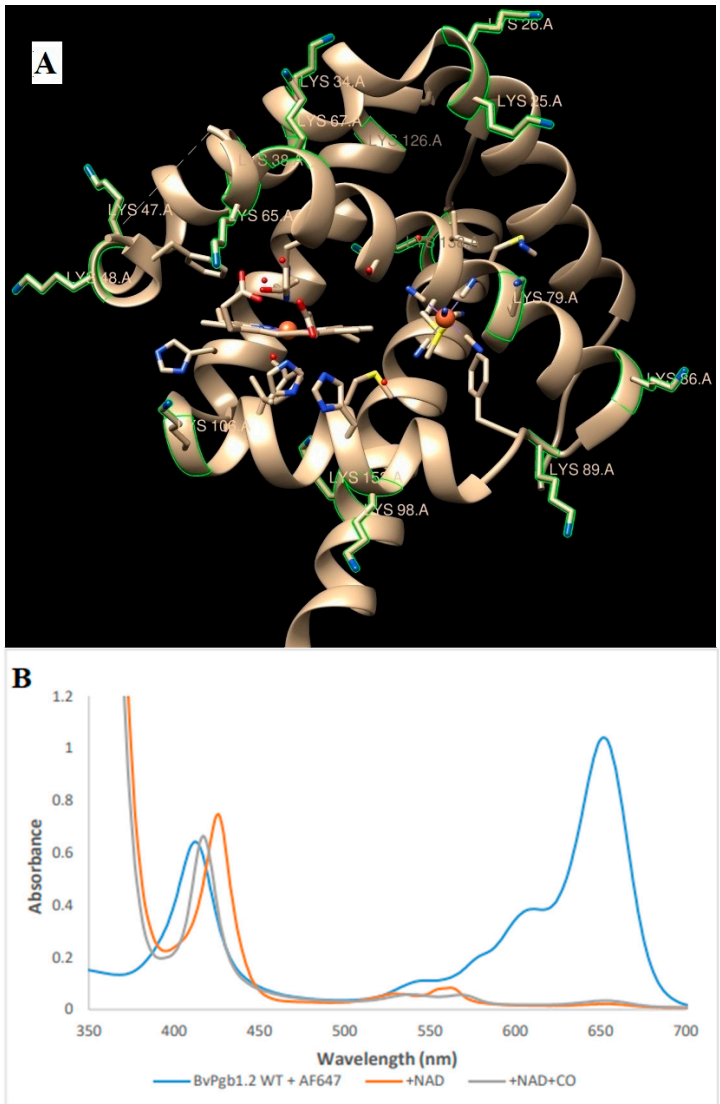


Figure 23. (A): BvPgb1.2 WT (PDB-ID: 7ZOS) with highlighted surface-exposed lysine residues (16 in total). The residues were used as handles to covalently link the fluorescent markers equipped with activated esters to the proteins. **(B):** Functionality of labeled BvPgb1.2 WT. Original protein-marker conjugation (blue), + sodium dithionite (NAD, orange) and CO (gray). The shifted Soret peak indicated functional protein after marker conjugation.

Even though the labeling conjugation would work in theory, modifications of proteins might alter the functionality and render them inactive. Therefore, the functionality of all proteins where analyzed with UV-Vis spectroscopy with different ligands. In Figure 23B, the results for BvPgb1.2 WT conjugated with a red fluorescent marker (AF647) can be seen. The characteristic Soret peak was located

at 412 nm, while the absorbance of the marker was detected at ~650 nm (blue). When reduced with NAD (orange), the absorption of the marker disappeared. However, this was expected due to the bleaching effect previously observed associated with NAD addition to fluorescent tags [168]. Moreover, the expected shift of the Soret peak was observed (~424 nm). When the nsPgb was saturated with CO (gray), the Soret peak shifted again (~416 nm), indicating a functioning Hb with maintained activity after conjugation (as in Figure 6). The same results were observed for both C86A and Mb, also with a second dye absorbing in the green spectrum (AF568).

The labeled proteins were injected into brain and lateral muscle tissues of zebrafishes and the effects were analyzed based on the behavior of the fishes in addition to protein localization. Granules of fluorescently labeled C86A and Mb were detected in brain tissue, especially in a subset of potentially injured neurons (**Paper V**). This tendency has been observed in a previous study using hemeproteins in zebrafish [169]. In both cases, no/low levels of oxidation were detected in surrounding tissues. In contrast, no accumulation of BvPgb1.2 WT could be detected in the brain.

All proteins were also injected into lateral muscle tissue. Similar results were seen for C86A and Mb, although WT was also detected here. C86A and WT accumulation are shown in Figure 24. Granules of AF647-C86A were found in a subset of muscle fibers. Scattered fluorescent granules were also associated with blood vessels (and/or nerves) along the muscle fiber bundles (arrow in Figure 24A). No excessive oxidative stress was associated with either C86A or Mb. Similar results were found for WT, although some of the most distinct protein granules showed elevated levels of oxidation. If this effect is consistent or a result of the microinjection procedure itself remains to be proven since all the granules were located close to site of injection. Thus, the microinjection itself might have contributed to observed effect. In any case, all globins showed accumulation of protein in either brain or muscle, mostly without associated oxidative stress in surrounding tissues. In addition, no altered behavior of the fishes was detected after injections. Thus, an adequate tolerance was observed for the globins and no significant differences were observed between the penta- and hexacoordinated proteins in this model system.

To further characterize the accumulation of the hemeproteins, future studies could involve multiple injections of the labeled proteins in the same/other tissues of the zebrafish to pinpoint a pattern of accumulation, as seen in **Paper V**. In addition, toxic effects for Fe^{2+}/Fe^{3+} iron salts and free heme could be added to exemplify the toxic effects to the free iron itself, while the protein-heme interactions in the globins would provide additional oxidative protection. The protective effect would be most prominent for the hexacoordinated proteins, due to their stable features and low heme loss tendencies. To my knowledge, this preliminary study is one of the first *in vivo* studies using nsPgbs and provided valuable insights regarding tolerance and

oxidative stress of nsPgbs in biological systems. Thus, more toxicological studies must be conducted to elucidate the effects of nsPgbs in similar and other model systems.

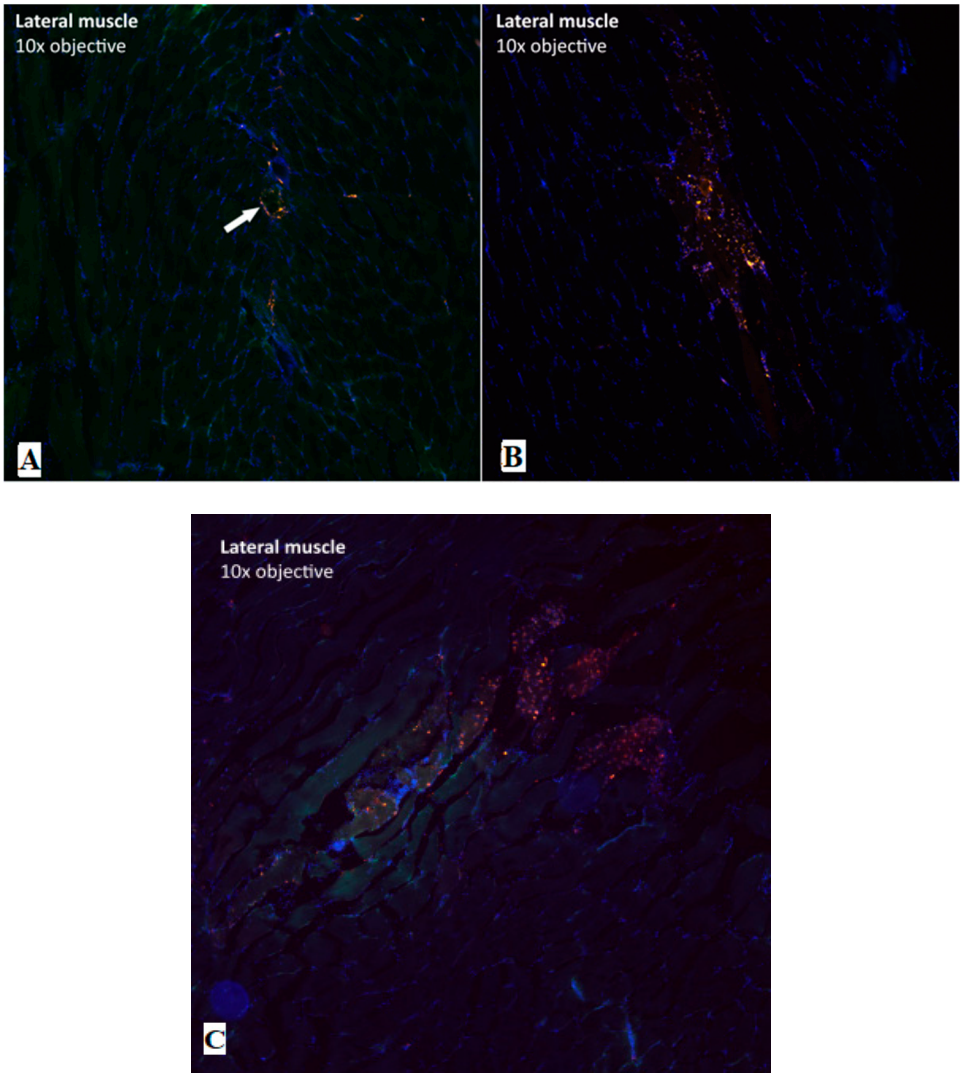


Figure 24. **(A and B):** Injections in lateral muscle tissue of BvPgb1.2 C86A and **(C):** WT injections in lateral muscle tissue. Granular AF647 fluorescence showed two distinct populations. **(A):** Scattered AF647-fluorescent granules associated with blood vessels (arrow: cross-sectioned blood vessel) between muscle fibers. **(B):** AF647-fluorescent granules accumulated in a subset of muscle fibers, presumably close to the injection site. **(C):** Distinct granular AF647-fluorescence was observed in a subset of muscle fibers, some of which showed marginally elevated diffuse fluorescence. No detection of AF647 was found in brain tissue. Strongly AF647-fluorescent granules also showed elevated oxidation levels. Redox staining: relative oxidation levels in green pseudocolor. Red pseudocolor shows the AF647 fluorescence.

5 Conclusions and Future Prospects

Ever since the first Pgbs in legumes were discovered in the late 1930s, there have been extensive efforts in understanding the role of these proteins in plants and their potential use in biotechnology. In addition, it is approximately 40 years ago nsPgbs were first observed. Since then, Pgbs have been found in essentially all plants in some form, with different numbers and characteristics.

Even though we currently know more of the origin and evolution of Pgbs and their features, continued efforts regarding the functionality and important interactions within and between the amino acids comprising the proteins are crucial in order to better understand and design these proteins for potential human use. This thesis has focused on two major areas, mainly characterization of class I nsPgbs and the impact of the conserved cysteine residue on a structural and functional level (**Paper I-III**), as well as their potential practical utility, mainly in terms of encapsulation and toxicological effects (**Paper IV** and **V**). Here, site-directed mutagenesis has been an important tool in generating new mutants and studying their characteristics in comparison with the WT protein in order to elucidate the impact of these residues.

In conclusion, this thesis has expanded the knowledge regarding the intrinsic functions of the conserved cysteine residue and its implications. This work has deepened the understanding regarding dimeric interactions and flexible regions in nsPgbs and provided important clues for potential practical applicability. The focus was to further characterize important features of nsPgbs, understand the interactions within and between the predominately dimeric nature of these proteins and provide insights in how to design and use them in biotechnology to solve current problems. However, despite the gained understanding of these processes, several obstacles remain to be resolved in order to be an important tool in the future. In other words, what actions would be beneficial to conduct for future considerations?

In terms of globin features, more research is needed regarding the structure and function of class II and trPgbs, to add to current and limited knowledge. For example, no crystal structures of class II nsPgbs exist, which would add valuable information of structural and functional characteristics for this class. We have tried to crystallize BvPgb2 but without any success, further indicating the non-crystallizing propensity of class II nsPgbs. Even though structure predictions (AlphaFold) have been conducted for this protein, the models will predominately be based on class I nsPgbs and important structural features might be lost. Thus, missing electron density

information of class II nsPgbs could result in less accurate models. In addition, structural studies of trPgb, especially from legumes, have not been studied to a great extent. This could also be interesting to investigate further, adding to the current limited knowledge regrading this type of globin. Furthermore, less is known about the intracellular localization of Pgbs, even though they have been found in both cytoplasm and chloroplasts in plant cells. What other proteins do they interact with and how do these interactions affect the functionality of Pgbs? These are a few examples that would provide greater understanding of these proteins and possibly show even more versatility than we currently know.

In terms of structure-function relationships for class I nsPgbs, the ability to maintain (or increase) the stability of the cysteine-substituted mutant in combination with its inert characteristics could make it a suitable candidate as a food additive or supplement. This would relieve some of the current problems using inorganic iron salts and their related side effects while providing a bioavailable source of dietary iron. Moreover, the previously mentioned concerns regarding micro- and anti-nutrient content in meat substitutes would need new approaches to relieve these current limitations [10]. Using a stable and bioavailable source of heme iron, as in nsPgbs, could be a suitable solution in new formulations of these products, especially considering their propensity to maintain the heme group in comparison to pentacoordinated Hbs (**Paper I**).

Another potential use of nsPgbs could be as a novel O_2 therapeutic, to be used for severely O_2-deprived tissues in particular. Even though unmodified nsPgbs might be of benefit here as well, they could potentially be less suitable due to their slow O_2 release and NO scavenging. However, the general high stability and robustness of nsPgbs might provide valuable properties that current solutions lack. Although the lipid system used in **Paper IV** was not efficient for encapsulation of the hemeproteins under the studied conditions, other techniques have been used to encapsulate similar proteins. As mentioned before, alternative encapsulation formulations of mainly pentacoordinated Hbs have been conducted with positive results. Using saturated phospholipids, cross-linking the nsPgbs to each other/different proteins (such as human serum albumin) could be suitable techniques to obtain stable and non-immunogenic nanoparticles with increased retention times. Here, the final application would ultimately determine the formulation. In addition, the number of surface-exposed lysine residues in BvPgb1.2 would provide suitable polymerization sites for glutaraldehyde cross-linking and would be interesting to investigate in the future.

Nevertheless, regardless if these proteins will be used for nutritional or therapeutic purposes, these still have to be clinically tested for human use. Although we provided some clues regarding the non-toxic effects of these proteins in zebrafish **(Paper V),** nsPgbs need to be more extensively studied in other biological systems. For instance, rodents might be the next animal model in order to evaluate if these proteins can be regarded as safe before more extensive clinical trials might be conducted. This is

something that has to be explored in detail before the use of these proteins can be frequently adopted for mentioned purposes. Furthermore, another obstacle that we are facing is the current legislation regarding genetically modified organisms (GMOs), including genes and proteins thereof. Currently, the most common approach to obtain high quantities of nsPgbs is through recombinant production in expression hosts, such as *E. coli* or yeast. Since the early 2000s, when the European Food Safety Authority (EFSA) was formed, Europe has had one of the most stringent regulations regarding GMOs for food and/or feed. All non-traditional foods, called "novel foods", has to be approved on a case-by-case basis regarding four main categories: safety, freedom of choice, labeling and tractability source [170]. This slow and circumstantial process of approval has been criticized on several occasions for risking the future food safety in Europe. Thus, the utilization of these techniques for nutritional applications are limited.

Hopefully, this thesis has highlighted the current knowledge of Pgbs but also the future work needed to elevate it to the next level. For instance, the generation of new mutants of these proteins with altered features and characteristics are crucial to expand the possibilities of usage. In our lab, we have generated a library of mutants of BvPgb1.2 with one or more cysteine residues located on the surface of the protein but still maintaining its functionality. It would be possible to use the thiol groups as handles for linking desired molecules, such as lipids, which could lead to increased retention time in biological systems. Moreover, if the effects of the mutations can be assessed more accurately, for instance with reliable dynamic simulations and approximations, this would save important resources in the future regarding generation of new mutant libraries. This was touched upon in **Paper III**, but more work is needed regarding the evaluation of monomeric/dimeric characteristics from future simulations and confirmation with other methods, such as SAXS, circular dichroism and cryogenic electron microscopy.

Finally, overexpressing or introducing nsPgbs or other globins within the plants themselves may fortify the iron content, which have been done in previous work [103, 113]. Furthermore, expression of Hb from *Vitreoscilla* bacteria in tobacco plants (*Nicotiana tabaccum*) yielded transgenic plants with increased productivity, higher growth rates, faster development and altered metabolism [171]. The same protein was also overexpressed in cabbage (*Brassica oleracea* var. *capitata*), *Arabidopsis* and maize (*Zea mays*), where similar results were obtained in addition to increased tolerance against immersion stress and waterlogging [172, 173]. This could lead to a more efficient solution of utilizing the iron directly in the plants. Thus, the need of downstream purification and other resource-consuming steps would be mitigated. However, a more comprehensive understanding regarding overexpression and/or suppression of nsPgbs *in vivo* is needed to better engineer and design future plants with different characteristics. For instance, overexpression of barley nsPgb has been shown to delay growth and development, as well as decreasing seed yield [174], while beneficial effects have been observed for the

same and other plants, as previously mentioned [88]. This lack of consistency would be crucial to elucidate in order to start extensive future cultivation programs. Furthermore, more knowledge of other genes induced by stress responses would be of interest and how these genes (or translated proteins) relate and interact with nsPgbs. Therefore, further investigations regarding similar effects of overexpressing these proteins in monocots/dicots, as observed in barley, are needed to evaluate the upscaling and feasibility potential. In any case, the legislation of using transgenic plants for non-research purposes is still restrictive, limiting their potential use as well.

Even though there are plenty of areas where more efforts are needed, the possibilities are vast regarding nsPgbs due to their versatility and valuable features. These proteins might be of great use in the future as long as researches are given continued support to explore and utilize them for the greater good, facilitating innovative biotechnical solutions to current problems.

6 Acknowledgements

Finally, it is time to thank all of you who have supported me and made this thesis possible to withstand. This would not have been possible without you and I am truly grateful for your support.

First of all, I want to thank my supervisor and mentor **Leif Bülow** for his support and belief in me. Without you, the world of phytoglobins would not have opened up for me, of that I am sure. Thank you for taking me in under your wings and for all the conversations we have had over the years, both scientific and sport-related ones. Deep down I know you like Frölunda as much as I do but you really have to look for it.

I want to thank **Mats Larsson** for his support during the later stages of my PhD. I am grateful for our conversations, both professionally and personally. You are a true son of Alvesta and it is always a pleasure talking to you.

Thank you to the people at TBK for your support and company over the years. Thank you **Lei Ye** for constructive conversations and good collaboration, especially in the course lab. **Alfredo Zambrano**, who has been my guiding light in the division. The conversations we have had over the years would be inappropriate to be mentioned here! Thank you **Cedric Dicko** for (almost) always lighten up the mood even in the darkest times. It has been an honor to get to know you and your dark and shitty corner. I am looking forward to our business idea; it is so brilliant I will not mention it here. It is between us my friend.

Thank you **Liselotte Andersson** and **Per-Olof Larsson** for your help at the division over the years. Ever since I tasted your glögg for the first time P-O, I cannot go back. Thanks **Lieselotte Cloetens**, **Solmaz Hajizadeh** and **Sofia Marmon** for the nice conversations and company over the years. Too bad Corona put Cake Maffia on-hold but I am sure it will make a comeback in the near future. Thank you **Alfia Khairullina** for the support and discussions. Soon we are both finished and then nothing can stop us! Also, thank you **Nélida Leiva-Eriksson** for the collaboration and guidance during my Master Degree/beginning of my doctoral studies.

I would also like to thank the members of the Gene Group. To **Johan** for helping me in the lab, especially in the course lab. Without you, Protein Engineering would not be such a nice course and it has been fun teaching with you. In addition, thank you for always keeping your door open and wanting to help. Thanks to **Leonard**,

Anna and **Mona-Lissa** for your presence in the division and nice work. **Leonard:** even though I am leaving and you have to carry the heavy burden of phytoglobins on your own, remember that I am always in your heart when you need me! In addition, I would like to thank **Karin Kettisen** for the help with Hb-related questions over the years. However, Rangers will always be better than Avalanche, no matter what you or Tobias say…

I would also like to thank the people at CropTailor. Thank you **Lars**, **Olof**, **Andreas** and **Varma** for the company during my PhD.

I would also like to thank all collaborators and co-authors that I have interacted with during my studies. A special thanks to **Marie Skepö** and her group (**Mona, Amanda** and **Henrik**). We had a "svineflott" time in Grenoble for sure! In addition, thank you **Jennifer** and **Tommy** for the nice collaboration with the nsPgb/Mb encapsulation.

A big thanks to my friends outside of TBK. Even though a few years have passed since we studied together, it feels like time has stopped when we meet again. This work would not have been possible without your support and company. As we say in Profilerna: JAAAOOOO!

Last but certainly not least, thank you to my loving family. Life would not mean anything without you, your love and comfort. Mamma, pappa, Mudi, Tobbe, Teddan, Therese and the rest of my family back home in Alvesta: Jag älskar er så otroligt mycket och detta vore absolut inte möjligt utan ert stöd!

7 References

1. Pettersen, E.F., et al., *UCSF Chimera--a visualization system for exploratory research and analysis.* J Comput Chem, 2004. **25**(13): p. 1605-12.
2. Goddard, T.D., et al., *UCSF ChimeraX: Meeting modern challenges in visualization and analysis.* Protein Sci, 2018. **27**(1): p. 14-25.
3. Hill, R., M. Hargrove, and R. Arredondo-Peter, *Phytoglobin: a novel nomenclature for plant globins accepted by the globin community at the 2014 XVIII conference on Oxygen-Binding and Sensing Proteins.* F1000Res, 2016. **5**: p. 212-212.
4. Warner, M.a.K., MT. *Iron Defieciency Anemia* 2022 2022-08/08 [cited 202320/1];Available from: https://www.ncbi.nlm.nih.gov/books/NBK448065/.
5. Cappellini, M.D. and I. Motta, *Anemia in Clinical Practice-Definition and Classification: Does Hemoglobin Change With Aging?* Semin Hematol, 2015. **52**(4): p. 261-9.
6. Organization, W.H. *Anaemia* 2023 [cited 2023 12/1]; Available from: https://www.who.int/health-topics/anaemia#tab=tab_1.
7. van der Merwe, R., et al., *Improving iron and zinc bioaccessibility through food-to-food fortification of pearl millet with tropical plant foodstuffs (moringa leaf powder, roselle calyces and baobab fruit pulp).* J Food Sci & Tech, 2019. **56**(4): p. 2244-2256.
8. Chouraqui, J.P., *Dietary Approaches to Iron Deficiency Prevention in Childhood-A Critical Public Health Issue.* Nutrients, 2022. **14**(8): p. 1604.
9. Piskin, E., et al., *Iron Absorption: Factors, Limitations, and Improvement Methods.* ACS Omega, 2022. **7**(24): p. 20441-20456.
10. Mayer Labba, I.-C., et al. *Nutritional Composition and Estimated Iron and Zinc Bioavailability of Meat Substitutes Available on the Swedish Market*. Nutrients, 2022. **14**(19): p. 3903, DOI: 10.3390/nu14193903.
11. Chen, J.Y., M. Scerbo, and G. Kramer, *A review of blood substitutes: examining the history, clinical trial results, and ethics of hemoglobin-based oxygen carriers.* Clinics (Sao Paulo), 2009. **64**(8): p. 803-13.
12. Bäumler, H., et al., *Novel Hemoglobin Particles—Promising New-Generation Hemoglobin-Based Oxygen Carriers.* Artificial Organs, 2014. **38**(8): p. 708-714.
13. Sen Gupta, A., *Hemoglobin-based Oxygen Carriers: Current State-of-the-art and Novel Molecules.* Shock, 2019. **52**(1S Suppl 1): p. 70-83.
14. Kubo, H., *Uber das Hamoprotein aus den Wurzelknollchen von Leguminosen.* Acta Phyto Tokyo, 1939. **11**(1): p. 195-200.
15. Bogusz, D., et al., *Functioning haemoglobin genes in non-nodulating plants.* Nature, 1988. **331**(6152): p. 178-80.

16. Appleby, C.A., J.D. Tjepkema, and M.J. Trinick, *Hemoglobin in a nonleguminous plant, parasponia: possible genetic origin and function in nitrogen fixation.* Science, 1983. **220**(4600): p. 951-3.
17. Dordas, C., *Nonsymbiotic hemoglobins and stress tolerance in plants.* Plant Sci, 2009. **176**(4): p. 433-40.
18. Riquelme, A. and P. Hinrichsen, *Non-symbiotic hemoglobin and its relation with hypoxic stress.* Chil J Agri Res, 2015. **75**: p. 80-89.
19. Edelheit, O., A. Hanukoglu, and I. Hanukoglu, *Simple and efficient site-directed mutagenesis using two single-primer reactions in parallel to generate mutants for protein structure-function studies.* BMC Biotech, 2009. **9**(1): p. 61.
20. Jensen, F.B., A. Fago, and R.E. Weber, *1 - Hemoglobin Structure and Function*, in *Fish Physiology*, S.F. Perry and B.L. Tufts, Editors. 1998, Academic Press. p. 1-40.
21. Stryer, L.B., J.M; Gatto Jr, G.J; Tymoczko, J.L *Biochemistry*. 9th ed. Biochemistry, ed. W.H.F.a. Company. 2019, New York: Macmillan International Higher Education. 207-226.
22. Perutz, M.F., et al., *Structure of Hæmoglobin: A Three-Dimensional Fourier Synthesis at 5.5-Å. Resolution, Obtained by X-Ray Analysis.* Nature, 1960. **185**(4711): p. 416-422.
23. Kendrew, J.C., et al., *Structure of Myoglobin: A Three-Dimensional Fourier Synthesis at 2 Å. Resolution.* Nature, 1960. **185**(4711): p. 422-427.
24. Hoeger, U.a.H., J.R, *Vertebrate and Invertebrate Respitory Proteins, Lipoproteins and other Body Fluid Proteins*, U.a.H. Hoeger, J.R, Editor. 2020, Springer Cham, Switzerland p. 345-382.
25. Ahmed, M.H., M.S. Ghatge, and M.K. Safo, *Hemoglobin: Structure, Function and Allostery.* Subcell Biochem, 2020. **94**: p. 345-382.
26. Vázquez-Limón, C., et al., *The evolution of land plant hemoglobins.* Plant Sci, 2012. **191-192**: p. 71-81.
27. Fermi, G., et al., *The crystal structure of human deoxyhaemoglobin at 1.74 A resolution.* J Mol Biol, 1984. **175**(2): p. 159-74.
28. Poulos, T.L., *Heme Enzyme Structure and Function.* Chem Rev, 2014. **114**(7): p. 3919-3962.
29. Everse, J., *Heme Proteins*, in *Encyclopedia of Biological Chemistry (Second Edition)*, W.J. Lennarz and M.D. Lane, Editors. 2013, Academic Press: Waltham. p. 532-538.
30. Gupta, K.J., et al., *Plant hemoglobins: Important players at the crossroads between oxygen and nitric oxide.* FEBS Lett, 2011. **585**(24): p. 3843-3849.
31. Li, T., H.L. Bonkovsky, and J.-t. Guo, *Structural analysis of heme proteins: implications for design and prediction.* BMC Struc Bio, 2011. **11**(1): p. 13.
32. Kakar, S., et al., *Structure and reactivity of hexacoordinate hemoglobins.* Biophy Chem, 2010. **152**(1-3): p. 1-14.
33. Burmester, T. and T. Hankeln, *What is the function of neuroglobin?* J Exp Biol, 2009. **212**(Pt 10): p. 1423-8.

34. Halligan, K.E., F.L. Jourd'heuil, and D. Jourd'heuil, *Cytoglobin is expressed in the vasculature and regulates cell respiration and proliferation via nitric oxide dioxygenation.* J Biol Chem, 2009. **284**(13): p. 8539-47.
35. Hodges, N.J., et al., *Cellular protection from oxidative DNA damage by over-expression of the novel globin cytoglobin in vitro.* Mutagenesis, 2008. **23**(4): p. 293-8.
36. Dordas, C., et al., *Expression of a stress-induced hemoglobin affects NO levels produced by alfalfa root cultures under hypoxic stress.* Plant J, 2003. **35**(6): p. 763-70.
37. Sowa, A.W., et al., *Altering hemoglobin levels changes energy status in maize cells under hypoxia.* PNAS, 1998. **95**(17): p. 10317-21.
38. Nie, X. and R.D. Hill, *Mitochondrial Respiration and Hemoglobin Gene Expression in Barley Aleurone Tissue.* Plant Physiol, 1997. **114**(3): p. 835-840.
39. Wu, G., L.M. Wainwright, and R.K. Poole, *Microbial globins.* Adv Microb Physiol, 2003. **47**: p. 255-310.
40. Vinogradov, S.N., et al., *A model of globin evolution.* Gene, 2007. **398**(1-2): p. 132-42.
41. Hoogewijs, D., et al., *Androglobin: a chimeric globin in metazoans that is preferentially expressed in Mammalian testes.* Mol Biol Evol, 2012. **29**(4): p. 1105-14.
42. Vinogradov, S.N., et al., *A phylogenomic profile of globins.* BMC Evol Bio, 2006. **6**: p. 31-31.
43. Hoogewijs, D., et al., *A phylogenetic analysis of the globins in fungi.* PLoS One, 2012. **7**(2): p. e31856.
44. Becana, M., et al., *Plant hemoglobins: a journey from unicellular green algae to vascular plants.* New Phyt, 2020. **227**(6): p. 1618-1635.
45. Vinogradov, S., D. Hoogewijs, and R. Arredondo-Peter, *What are origins and phylogeny of plant hemoglobins?* Com & Int Biology, 2011. **4**: p. 443-5.
46. Vinogradov, S.N., et al., *Phylogenetic Relationships of 3/3 and 2/2 Hemoglobins in Archaeplastida Genomes to Bacterial and Other Eukaryote Hemoglobins.* Mol Plant, 2011. **4**(1): p. 42-58.
47. Smagghe, B.J., et al., *Review: correlations between oxygen affinity and sequence classifications of plant hemoglobins.* Biopol, 2009. **91**(12): p. 1083-96.
48. Sturms, R., et al., *Trema and Parasponia Hemoglobins Reveal Convergent Evolution of Oxygen Transport in Plants.* Biochemistry, 2010. **49**(19): p. 4085-4093.
49. Clarke, J.T., R.C.M. Warnock, and P.C.J. Donoghue, *Establishing a time-scale for plant evolution.* New Phytol, 2011. **192**(1): p. 266-301.
50. Timme, R.E., T.R. Bachvaroff, and C.F. Delwiche, *Broad Phylogenomic Sampling and the Sister Lineage of Land Plants.* PLoS One, 2012. **7**(1): p. e29696.
51. Jiao, Y., et al., *Ancestral polyploidy in seed plants and angiosperms.* Nature, 2011. **473**(7345): p. 97-100.
52. Vessey, J.K., K. Pawlowski, and B. Bergman, *Root-based N_2-fixing symbioses: Legumes, actinorhizal plants, Parasponia sp. and cycads.* Plant & Soil, 2004. **266**(1/2): p. 205-230.
53. Hoy, J.A., et al., *Plant hemoglobins: a molecular fossil record for the evolution of oxygen transport.* J Mol Biol, 2007. **371**(1): p. 168-79.

54. Gopalasubramaniam, S.K., et al., *Cloning and characterization of a caesalpinoid (Chamaecrista fasciculata) hemoglobin: The structural transition from a nonsymbiotic hemoglobin to a leghemoglobin.* Proteins: Struc, Func & Bioinfo, 2008. **72**(1): p. 252-260.
55. Hebelstrup, K.H., A.U. Igamberdiev, and R.D. Hill, *Metabolic effects of hemoglobin gene expression in plants.* Gene, 2007. **398**(1-2): p. 86-93.
56. Igamberdiev, A.U., N.V. Bykova, and R.D. Hill, *Structural and functional properties of class 1 plant hemoglobins.* IUBMB Life, 2011. **63**(3): p. 146-52.
57. Garrocho-Villegas, V., S.K. Gopalasubramaniam, and R. Arredondo-Peter, *Plant hemoglobins: what we know six decades after their discovery.* Gene, 2007. **398**(1-2): p. 78-85.
58. Baudouin, E., *The language of nitric oxide signalling.* Plant Biol (Stuttg), 2011. **13**(2): p. 233-42.
59. Gupta, K.J. and A.U. Igamberdiev, *The anoxic plant mitochondrion as a nitrite: NO reductase.* Mitochondrion, 2011. **11**(4): p. 537-43.
60. Gupta, K.J., et al., *On the origins of nitric oxide.* Trends Plant Sci, 2011. **16**(3): p. 160-8.
61. Cantrel, C., et al., *Nitric oxide participates in cold-responsive phosphosphingolipid formation and gene expression in Arabidopsis thaliana.* New Phytol, 2011. **189**(2): p. 415-27.
62. Hargrove, M.S., et al., *Crystal structure of a nonsymbiotic plant hemoglobin.* Structure, 2000. **8**(9): p. 1005-14.
63. Bykova, N.V., et al., *Identification of an intermolecular disulfide bond in barley hemoglobin.* Biochem & Biophy Res Com, 2006. **347**(1): p. 301-309.
64. Arredondo-Peter, R., et al., *Rice hemoglobins. Gene cloning, analysis, and O2-binding kinetics of a recombinant protein synthesized in Escherichia coli.* Plant Physiol, 1997. **115**(3): p. 1259-66.
65. Perazzolli, M., et al., *Arabidopsis nonsymbiotic hemoglobin AHb1 modulates nitric oxide bioactivity.* Plant Cell, 2004. **16**(10): p. 2785-94.
66. Igamberdiev, A.U., N.V. Bykova, and R.D. Hill, *Nitric oxide scavenging by barley hemoglobin is facilitated by a monodehydroascorbate reductase-mediated ascorbate reduction of methemoglobin.* Planta, 2006. **223**(5): p. 1033-40.
67. The, S.V., R. Snyder, and M. Tegeder, *Targeting Nitrogen Metabolism and Transport Processes to Improve Plant Nitrogen Use Efficiency.* Front Plant Sci, 2021. **11:** p. 1-22.
68. Igamberdiev, A.U. and R.D. Hill, *Nitrate, NO and haemoglobin in plant adaptation to hypoxia: an alternative to classic fermentation pathways.* J Exp Bot, 2004. **55**(408): p. 2473-82.
69. Hoy, J.A. and M.S. Hargrove, *The structure and function of plant hemoglobins.* Plant Physio & Biochem : PPB, 2008. **46**(3): p. 371-379.

70. Vigeolas, H., D. Hühn, and P. Geigenberger, *Nonsymbiotic Hemoglobin-2 Leads to an Elevated Energy State and to a Combined Increase in Polyunsaturated Fatty Acids and Total Oil Content When Overexpressed in Developing Seeds of Transgenic Arabidopsis Plants.* Plant Phys, 2011. **155**(3): p. 1435-1444.
71. Wang, Y.H., et al., *Two tomato non-symbiotic haemoglobin genes are differentially expressed in response to diverse changes in mineral nutrient status.* Plant, Cell & Envir, 2003. **26**(5): p. 673-680.
72. Watts, R.A., et al., *A hemoglobin from plants homologous to truncated hemoglobins of microorganisms.* PNAS, 2001. **98**(18): p. 10119-10124.
73. Parrilli, E., et al., *The role of a 2-on-2 haemoglobin in oxidative and nitrosative stress resistance of Antarctic Pseudoalteromonas haloplanktis TAC125.* Biochimie, 2010. **92**(8): p. 1003-9.
74. Vieweg, M.F., N. Hohnjec, and H. Küster, *Two genes encoding different truncated hemoglobins are regulated during root nodule and arbuscular mycorrhiza symbioses of Medicago truncatula.* Planta, 2005. **220**(5): p. 757-66.
75. Hargrove, M.S., et al., *Characterization of recombinant soybean leghemoglobin a and apolar distal histidine mutants11Edited by K. Nagai.* J Mol Bio, 1997. **266**(5): p. 1032-1042.
76. Becana, M., et al., *Recent insights into antioxidant defenses of legume root nodules.* New Phyt, 2010. **188**(4): p. 960-976.
77. Ott, T., et al., *Symbiotic leghemoglobins are crucial for nitrogen fixation in legume root nodules but not for general plant growth and development.* Curr Biol, 2005. **15**(6): p. 531-5.
78. Ott, T., et al., *Absence of symbiotic leghemoglobins alters bacteroid and plant cell differentiation during development of Lotus japonicus root nodules.* Mol Plant Micro Interact, 2009. **22**(7): p. 800-8.
79. Wang, L., et al., *CRISPR/Cas9 knockout of leghemoglobin genes in Lotus japonicus uncovers their synergistic roles in symbiotic nitrogen fixation.* New Phytol, 2019. **224**(2): p. 818-832.
80. Simontacchi, M., et al., *Plant Survival in a Changing Environment: The Role of Nitric Oxide in Plant Responses to Abiotic Stress.* Front Plant Sci, 2015. **6**.
81. Sun, C., et al., *Molecular functions of nitric oxide and its potential applications in horticultural crops.* Hort Res, 2021. **8**(1): p. 71.
82. Hill, R.D., *Non-symbiotic haemoglobins-What's happening beyond nitric oxide scavenging?* AoB PLANTS, 2012. **2012**: p. pls004.
83. Brouquisse, R., *Multifaceted roles of nitric oxide in plants.* J Exp Bot, 2019. **70**(17): p. 4319-4322.
84. Hill, R.D., *What are hemoglobins doing in plants?* Can J Bot, 1998. **76**(5): p. 707-712.
85. Hunt, P.W., et al., *Increased level of hemoglobin 1 enhances survival of hypoxic stress and promotes early growth in Arabidopsis thaliana.* PNAS, 2002. **99**(26): p. 17197-17202.
86. Hunt, P.W., et al., *Expression and evolution of functionally distinct haemoglobin genes in plants.* Plant Mol Biol, 2001. **47**(5): p. 677-92.

87. Igamberdiev, A.U., et al., *The haemoglobin/nitric oxide cycle: involvement in flooding stress and effects on hormone signalling.* Ann Bot, 2005. **96**(4): p. 557-64.
88. Zhang, J., et al., *Phytoglobin overexpression promotes barley growth in the presence of enhanced level of atmospheric nitric oxide.* J Exp Bot, 2019. **70**(17): p. 4521-4537.
89. Fleming, A.I., et al., *The purification, characterization and ligand-binding kinetics of hemoglobins from root nodules of the non-leguminous Casuarina glauca — Frankia symbiosis.* BBA - Pro Stru & Mol Enzym, 1987. **911**(2): p. 209-220.
90. Taylor, E.R., et al., *A cereal haemoglobin gene is expressed in seed and root tissues under anaerobic conditions.* Plant Mol Biol, 1994. **24**(6): p. 853-862.
91. Duff, S.M., J.B. Wittenberg, and R.D. Hill, *Expression, purification, and properties of recombinant barley (Hordeum sp.) hemoglobin. Optical spectra and reactions with gaseous ligands.* J Biol Chem, 1997. **272**(27): p. 16746-52.
92. Trevaskis, B., et al., *Two hemoglobin genes in Arabidopsis thaliana: the evolutionary origins of leghemoglobins.* PNAS, 1997. **94**(22): p. 12230-4.
93. Mukhi, N., et al., *X-ray crystallographic structural characteristics of Arabidopsis hemoglobin I and their functional implications.* BBA, 2013. **1834**(9): p. 1944-56.
94. Kakar, S., et al., *Crystal structures of Parasponia and Trema hemoglobins: differential heme coordination is linked to quaternary structure.* Biochemistry, 2011. **50**(20): p. 4273-80.
95. Christensen, S., et al., *Oxidative Implications of Substituting a Conserved Cysteine Residue in Sugar Beet Phytoglobin BvPgb 1.2.* Antioxidants, 2022. **11**(8): p. 1615.
96. Fronticelli, C., et al., *Chloride ion independence of the Bohr effect in a mutant human hemoglobin beta (V1M+H2deleted).* J Biol Chem, 1994. **269**(39): p. 23965-9.
97. Hubbard, S.R., et al., *X-ray crystal structure of a recombinant human myoglobin mutant at 2.8 A resolution.* J Mol Biol, 1990. **213**(2): p. 215-8.
98. Pesce, A., et al., *Human Brain Neuroglobin Structure Reveals a Distinct Mode of Controlling Oxygen Affinity.* Structure, 2003. **11**(9): p. 1087-1095.
99. de Sanctis, D., et al., *Crystal structure of cytoglobin: the fourth globin type discovered in man displays heme hexa-coordination.* J Mol Biol, 2004. **336**(4): p. 917-27.
100. Zhang, J., et al., *Optimization of the heme biosynthesis pathway for the production of 5-aminolevulinic acid in Escherichia coli.* Sci Reps, 2015. **5**(1): p. 8584.
101. Trent, J.T., 3rd and M.S. Hargrove, *A ubiquitously expressed human hexacoordinate hemoglobin.* J Biol Chem, 2002. **277**(22): p. 19538-45.
102. Ishchuk, O.P., et al., *Improved production of human hemoglobin in yeast by engineering hemoglobin degradation.* Met Engin, 2021. **66**: p. 259-267.
103. Carlsson, M.L.R., et al., *Plant based production of myoglobin - a novel source of the muscle heme-protein.* Sci Reps, 2020. **10**(1): p. 920.
104. Shao, Y., et al., *High-level secretory production of leghemoglobin in Pichia pastoris through enhanced globin expression and heme biosynthesis.* Biores Tech, 2022. **363**: p. 127884.

105. Leiva Eriksson, N., et al., *Sugar beet hemoglobins: reactions with nitric oxide and nitrite reveal differential roles for nitrogen metabolism.* Biochem J, 2019. **476**(14): p. 2111-2125.
106. Kettisen, K., et al., *Site-directed mutagenesis of cysteine residues alters oxidative stability of fetal hemoglobin.* Redox Biol, 2018. **19**: p. 218-225.
107. Holmberg, N., U. Ryde, and L. Bülow, *Redesign of the coenzyme specificity in L-lactate dehydrogenase from bacillus stearothermophilus using site-directed mutagenesis and media engineering.* Protein Eng, 1999. **12**(10): p. 851-6.
108. Cummins, P.M., K.D. Rochfort, and B.F. O'Connor, *Ion-Exchange Chromatography: Basic Principles and Application*, in *Protein Chromatography: Methods and Protocols*, D. Walls and S.T. Loughran, Editors. 2017, Springer New York: New York, NY. p. 209-223.
109. Panella, L., *Sugar Beet as an Energy Crop.* Sugar Tech, 2010. **12**(3): p. 288-293.
110. Leiva-Eriksson, N., et al., *Differential expression patterns of non-symbiotic hemoglobins in sugar beet (Beta vulgaris ssp. vulgaris).* Plant Cell Physiol, 2014. **55**(4): p. 834-44.
111. Leiva-Eriksson, N., *Biochemical and Physiological Characterization of Nonsymbiotic Plant Hemoglobins*, in *Division of Pure and Applied Biochemistry*. 2014, Lund University Lund. p. 54.
112. Chekin, F., et al., *Direct electrochemistry and bioelectrocatalysis of a class II non-symbiotic plant haemoglobin immobilised on screen-printed carbon electrodes.* Anal Bioanal Chem, 2010. **398**(4): p. 1643-9.
113. Gisbert, C., et al., *Overexpression of BvHb2, a Class 2 Non-Symbiotic Hemoglobin from Sugar Beet, Confers Drought-Induced Withering Resistance and Alters Iron Content in Tomato.* Agronomy, 2020. **10**(11): p. 1754.
114. Sosna, M., et al., *Electrochemical Characterization and Bioelectrocatalytic H2O2 Sensing of Non-Symbiotic Hexa-Coordinated Sugar Beet Hemoglobins.* Chem Electro Chem, 2020. **7**(9): p. 2114-2122.
115. Madeira, F., et al., *The EMBL-EBI search and sequence analysis tools APIs in 2019.* Nuc Acids Res, 2019. **47**(W1): p. W636-w641.
116. Garrocho-Villegas, V. and R. Arredondo-Peter, *Molecular cloning and characterization of a moss (Ceratodon purpureus) nonsymbiotic hemoglobin provides insight into the early evolution of plant nonsymbiotic hemoglobins.* Mol Biol Evol, 2008. **25**(7): p. 1482-7.
117. Waterman, M.R., *Role of cysteine residues in hemoglobin structure and function: transfer of p-mercuribenzoate from alpha subunits to beta subunits during tetramer formation.* BBA, 1974. **371**(1): p. 159-67.
118. Storz, J.F., R.E. Weber, and A. Fago, *Oxygenation properties and oxidation rates of mouse hemoglobins that differ in reactive cysteine content.* Comp Biochem Physiol A Mol Integr Physiol, 2012. **161**(2): p. 265-70.
119. Kan, H.-I., et al., *Subunit Disassembly Pathway of Human Hemoglobin Revealing the Site-Specific Role of Its Cysteine Residues.* J Phys Chem B, 2013. **117**(34): p. 9831-9839.

120. Mot, A.C., et al., *Redox control and autoxidation of class 1, 2 and 3 phytoglobins from Arabidopsis thaliana.* Sci Reps, 2018. **8**(1): p. 13714.
121. Silkstone, G.G.A., et al., *Engineering tyrosine electron transfer pathways decreases oxidative toxicity in hemoglobin: implications for blood substitute design.* Biochem J, 2016. **473**(19): p. 3371-3383.
122. Moser, C.C., et al., *Nature of biological electron transfer.* Nature, 1992. **355**(6363): p. 796-802.
123. Cooper, C.E., et al., *Haptoglobin binding stabilizes hemoglobin ferryl iron and the globin radical on tyrosine β145.* Antiox Redox Signal, 2013. **18**(17): p. 2264-73.
124. Silkstone, R.S., et al., *The βLys66Tyr Variant of Human Hemoglobin as a Component of a Blood Substitute.* Adv Exp Med Biol, 2016. **876**: p. 455-460.
125. Reeder, B.J., et al., *Tyrosine residues as redox cofactors in human hemoglobin: implications for engineering nontoxic blood substitutes.* J Biol Chem, 2008. **283**(45): p. 30780-30787.
126. Hargrove, M.S., et al., *His64(E7)-->Tyr apomyoglobin as a reagent for measuring rates of hemin dissociation.* J Biol Chem, 1994. **269**(6): p. 4207-14.
127. Ku, T., et al., *Predicting melting temperature directly from protein sequences.* Comp Biol & Chem, 2009. **33**(6): p. 445-450.
128. Wan, L., et al., *In vitro evolution of horse heart myoglobin to increase peroxidase activity.* PNAS, 1998. **95**(22): p. 12825-31.
129. Julió Plana, L., et al., *Thermal Stability of Globins: Implications of Flexibility and Heme Coordination Studied by Molecular Dynamics Simulations.* J Chem Inf Model, 2019. **59**(1): p. 441-452.
130. Kvist, M., et al., *An investigation of the peroxidase activity of Vitreoscilla hemoglobin.* J Biol Inorg Chem, 2007. **12**(3): p. 324-34.
131. Everse, J., M.C. Johnson, and M.A. Marini, *Peroxidative activities of hemoglobin and hemoglobin derivatives.* Meth Enzymol, 1994. **231**: p. 547-61.
132. Violante-Mota, F., et al., *Analysis of peroxidase activity of rice (Oryza sativa) recombinant hemoglobin 1: implications for in vivo function of hexacoordinate non-symbiotic hemoglobins in plants.* Phytochem, 2010. **71**(1): p. 21-6.
133. Marion, D., *An introduction to biological NMR spectroscopy.* Mol & Cell Prot: MCP, 2013. **12**(11): p. 3006-3025.
134. Bax, A. and S. Grzesiek, *Methodological advances in protein NMR.* Acc Chem Res, 1993. **26**(4): p. 131-138.
135. Bax, A., *Multidimensional nuclear magnetic resonance methods for protein studies.* Cur Opi Stru Biol, 1994. **4**(5): p. 738-744.
136. Kay, L.E. and K.H. Gardner, *Solution NMR spectroscopy beyond 25 kDa.* Cur Opi Stru Biol, 1997. **7**(5): p. 722-731.
137. Clore, G.M. and A.M. Gronenborn, *Determining the structures of large proteins and protein complexes by NMR.* Trend Biotech, 1998. **16**(1): p. 22-34.
138. Christensen, S., et al., *Conformational Dynamics of Phytoglobin BvPgb1.2 from Beta vulgaris ssp. vulgaris.* Int J Mol Sci, 2023. **24**(4): p. 3973.

139. al, O.e., *Deuterium induces a distinctive Escherichia coli proteome that correlates with the reduction in growth rate.* J Biol Chem, 2018. **294**(7): p. 2279-2292.
140. Wong, L.E., et al., *Sensitivity-Enhanced Four-Dimensional Amide–Amide Correlation NMR Experiments for Sequential Assignment of Proline-Rich Disordered Proteins.* J Amer Chem Soc, 2018. **140**(10): p. 3518-3522.
141. Parodi, F., *19 - Physics and Chemistry of Microwave Processing*, in *Comprehensive Polymer Science and Supplements*, G. Allen and J.C. Bevington, Editors. 1989, Pergamon: Amsterdam. p. 669-728.
142. Wüthrich, K., et al., *1H NMR relaxation in high-spin ferrous hemoproteins.* J Mag Res (1969), 1975. **19**(1): p. 111-113.
143. Banci, L., et al., *1H-NMR study of reduced heme proteins myoglobin and cytochrome P450.* Eur J Biochem, 1993. **215**(2): p. 431-7.
144. Zhao, J., et al., *Dynamics of dehaloperoxidase-hemoglobin A derived from NMR relaxation spectroscopy and molecular dynamics simulation.* J Inorg Biochem, 2018. **181**: p. 65-73.
145. Astegno, A., et al., *Structural Insights into the Heme Pocket and Oligomeric State of Non-Symbiotic Hemoglobins from Arabidopsis thaliana.* Biomolecules, 2020. **10**(12).
146. Savard, P.Y., et al., *Structure and dynamics of Mycobacterium tuberculosis truncated hemoglobin N: insights from NMR spectroscopy and molecular dynamics simulations.* Biochemistry, 2011. **50**(51): p. 11121-30.
147. Song, X.J., et al., *A comparative NMR study of the polypeptide backbone dynamics of hemoglobin in the deoxy and carbonmonoxy forms.* Biochemistry, 2007. **46**(23): p. 6795-803.
148. Kleckner, I.R. and M.P. Foster, *An introduction to NMR-based approaches for measuring protein dynamics.* BBA, 2011. **1814**(8): p. 942-68.
149. Mirdita, M., et al., *ColabFold: making protein folding accessible to all.* Nature Meth, 2022. **19**(6): p. 679-682.
150. Korasick, D.A. and J.J. Tanner, *Determination of protein oligomeric structure from small-angle X-ray scattering.* Protein Sci, 2018. **27**(4): p. 814-824.
151. Kettisen, K., et al., *Site-Specific Introduction of Negative Charges on the Protein Surface for Improving Global Functions of Recombinant Fetal Hemoglobin.* Front Mol Biosci, 2021. **8**: p. 649007.
152. Liu, Y. and A. Cao, *Encapsulating Proteins in Nanoparticles: Batch by Batch or One by One.* Meth Enzymol, 2017. **590**: p. 1-31.
153. Aguilar-Toalá, J.E., et al., *Encapsulation of bioactive peptides: a strategy to improve the stability, protect the nutraceutical bioactivity and support their food applications.* RSC Adv, 2022. **12**(11): p. 6449-6458.
154. Gilbert, J., et al., *Immobilisation of β-galactosidase within a lipid sponge phase: structure, stability and kinetics characterisation.* Nanoscale, 2019. **11**(44): p. 21291-21301.
155. Zhao, F., et al., *Joule Heating and Thermal Denaturation of Proteins in Nano-ESI Theta Tips.* J Am Soc Mass Spectrom, 2017. **28**(10): p. 2001-2010.

156. Szebeni, J., et al., *Encapsulation of hemoglobin in phospholipid liposomes: characterization and stability.* Biochemistry, 1985. **24**(12): p. 2827-32.
157. Kure, T. and H. Sakai, *Preparation of Artificial Red Blood Cells (Hemoglobin Vesicles) Using the Rotation–Revolution Mixer for High Encapsulation Efficiency.* ACS Biomat Sci & Eng, 2021. **7**(6): p. 2835-2844.
158. Sakai, H., *Overview of Potential Clinical Applications of Hemoglobin Vesicles (HbV) as Artificial Red Cells, Evidenced by Preclinical Studies of the Academic Research Consortium.* J Funct Biomater, 2017. **8**(1): p. 10.
159. Xiong, Y., et al., *Hemoglobin-Based Oxygen Carrier Microparticles: Synthesis, Properties, and In Vitro and In Vivo Investigations.* Biomacro, 2012. **13**(10): p. 3292-3300.
160. Xiong, Y., et al., *Nonvasoconstrictive hemoglobin particles as oxygen carriers.* ACS Nano, 2013. **7**(9): p. 7454-61.
161. Zizioli, D., et al., *Zebrafish disease models in hematology: Highlights on biological and translational impact.* BBA Mol Basis Dis, 2019. **1865**(3): p. 620-633.
162. Choi, T.-Y., et al., *Zebrafish as an animal model for biomedical research.* Exp & Mol Med, 2021. **53**(3): p. 310-317.
163. Zhang, Y., M. Chen, and C. Chen, *Using the Zebrafish as a Genetic Model to Study Erythropoiesis.* Int J Mol Sci, 2021. **22**(19).
164. Zhang, J. and I. Hamza, *Zebrafish as a model system to delineate the role of heme and iron metabolism during erythropoiesis.* Mol Genet Metab, 2019. **128**(3): p. 204-212.
165. Wakasugi, K. and I. Morishima, *Identification of Residues in Human Neuroglobin Crucial for Guanine Nucleotide Dissociation Inhibitor Activity.* Biochemistry, 2005. **44**(8): p. 2943-2948.
166. Takahashi, K., et al., *A globin-family protein, Cytoglobin 1, is involved in the development of neural crest-derived tissues and organs in zebrafish.* Dev Biol, 2021. **472**: p. 1-17.
167. Das, A. and P. Theato, *Activated Ester Containing Polymers: Opportunities and Challenges for the Design of Functional Macromolecules.* Chem Revs, 2016. **116**(3): p. 1434-1495.
168. Andra, K.K., et al., *Structural mapping of fluorescently-tagged, functional nhTMEM16 scramblase in a lipid bilayer.* J Biol Chem, 2018. **293**(31): p. 12248-12258.
169. Nässel, D.R., *Extensive labelling of injured insect neurons with seven different heme peptides.* Histochemistry, 1983. **79**(1): p. 95-104.
170. Compass, G. *Genetic Engineering plants and food: The European Regulatory System*. 2012 [cited 2023 15/1]; Available from: https://web.archive.org/web/20120814025652/http://www.gmo-compass.org/eng/regulation/regulatory_process/156.european_regulatory_system_genetic_engineering.html.

171. Holmberg, N., et al., *Transgenic tobacco expressing Vitreoscilla hemoglobin exhibits enhanced growth and altered metabolite production.* Nature Biotechnology, 1997. **15**(3): p. 244-247.
172. Li, X., et al., *Vitreoscilla hemoglobin overexpression increases submergence tolerance in cabbage.* Plant Cell Rep, 2005. **23**(10-11): p. 710-5.
173. Du, H., et al., *Overexpression of Vitreoscilla hemoglobin increases waterlogging tolerance in Arabidopsis and maize.* BMC Plant Biology, 2016. **16**(1): p. 35.
174. Hebelstrup, K.H., et al., *An assessment of the biotechnological use of hemoglobin modulation in cereals.* Physiol Plant, 2014. **150**(4): p. 593-603.

8 Papers

Printed by Media-Tryck, Lund 2023 NORDIC SWAN ECOLABEL 3041 0903

LUND UNIVERSITY

Division of Pure and Applied Biochemistry
Department of Chemistry
Faculty of Engineering

ISBN 978-91-7422-936-3