# The globins of cyanobacteria and green algae: An update

**Juliette T.J. Lecomte**[a,*] **and Eric A. Johnson**[b]
[a]T.C. Jenkins Department of Biophysics, Johns Hopkins University, Baltimore, MD, United States
[b]Department of Biology, Johns Hopkins University, Baltimore, MD, United States
*Corresponding author. e-mail address: lecomte_jtj@jhu.edu

## Contents

## Abstract

The globin superfamily of proteins is ancient and diverse. Regular assessments based on the increasing number of available genome sequences have elaborated on a complex evolutionary history. In this review, we present a summary of a decade of advances in characterising the globins of cyanobacteria and green algae. The focus is on haem-containing globins with an emphasis on recent experimental developments, which reinforce links to nitrogen metabolism and nitrosative stress response in addition to dioxygen management. Mention is made of globins that do not bind haem to provide an encompassing view of the superfamily and perspective on the field. It is reiterated that an effort toward phenotypical and in-vivo characterisation is needed to elucidate the many roles that these versatile proteins fulfil in oxygenic photosynthetic microbes. It is also proposed that globins from oxygenic organisms are promising proteins for applications in the biotechnology arena.

*Advances in Microbial Physiology*
ISSN 0065-2911, https://doi.org/10.1016/bs.ampbs.2024.04.004

## Abbreviations

| | |
|---|---|
| **5c** | five-coordinate |
| **6c** | six-coordinate |
| **BBAG** | bilin biosynthesis-associated globin |
| **BLAST** | basic local alignment search tool |
| **CtrHb** | haem domain of *Chlamydomonas eugametos* LI637 haemoglobin |
| **Fgb** | single domain haemoglobin related to flavohaemoglobin |
| **GCS** | globin-coupled sensor |
| **Hb** | haemoglobin |
| **LI** | light induced |
| **M** | myoglobin-like |
| **Mb** | myoglobin |
| **PDB** | protein data bank |
| **Pgb** | protoglobin |
| **PPR** | pentapeptide repeat |
| **PTM** | post-translational modification |
| **S** | sensor |
| **SSDgb** | sensor single domain globin related to the GCS globin domain |
| **STAS** | sulphate transporter and anti-sigma factor antagonist |
| **T** | truncated |

## 1. Introduction

The history of haemoglobin research is rooted in its importance to human health, and pragmatically, in the ease of extracting and purifying large quantities of coloured material from mammalian sources. By the middle of the twentieth century, globin studies had transformed biological sciences – proteins were finally viewed with a chemist's attention to the relationship between three-dimensional structure and functional properties. Decades of work in multiple laboratories yielded novel physical and biochemical knowledge about mammalian haemoglobins and myoglobins. The subsequent development of genome sequencing, recombinant technology, and protein purification techniques made possible the investigation of low-abundance proteins. As a result, the inventory of vertebrate globins saw an unexpected expansion with the addition of neuroglobin (Burmester, Weich, Reinhardt, & Hankeln, 2000), cytoglobin (Kawada et al., 2001), androglobin (Hoogewijs et al., 2011) and others (Koch & Burmester, 2016), while thousands of globins across all domains of life, including photosynthetic organisms, were also discovered. A century of mammalian physiologic investigations of haemoglobins has been linked to a growing, diverse genomic database of members of the globin superfamily. Now with our modern bioinformatic perspective this group of proteins offers unique insights on the evolution of life on earth spanning the early anoxic era, with oceans of plentiful reduced iron, through to the extensive diversity seen in the (generally) aerobic biome present today (Schuster, Salvatore, Moens, & Martí, 2024).

This article captures information available since our previous two summaries on the topic of cyanobacterial and chlorophytic globins (Johnson & Lecomte, 2013, 2015). To broaden the scope of the review, we briefly include non-haem binding globins as they contribute to the history of the superfamily. Throughout, we rely on several insightful reports that provide assessments of phylogeny (Schneider, Tan, Li, Fisher, & Zhang, 2022; Schuster et al., 2024), plant haemoglobins (Becana, Yruela, Sarath, Catalán, & Hargrove, 2020), and phycobiliproteins (Rockwell, Martin, & Lagarias, 2023). The reader is referred to these sources and the references they contain for in-depth descriptions.

## 1.1 Structural considerations

Before discussing the particulars of globins in cyanobacteria and chlorophytes, it is useful to take a structural perspective to organise the vast group of haem-binding globins. The globin domain is made of six to eight α helices of variable length and connected by turns and loops. In accordance with myoglobin labelling, the helices are sequentially referred to with letters from A at the N terminus to H at the C terminus (Perutz, 1979), with helix D missing from entire sets of proteins. Full-length globins adopt a sandwich topology by which three α helices fold over another three α helices (Fig. 1A). This "canonical" or 3/3 myoglobin fold (~150 residues) is occasionally decorated with extensions, e.g., for membrane anchoring (Blank et al., 2011; Ertas, Kiger, Blank, Marden, & Burmester, 2011; Hafideddine et al., 2023; Rhéault et al., 2015). The Protein Data Bank IDs of structural examples are provided in Table 1. Some globin domains are smaller (~120 residues) and referred to as "truncated" (Wittenberg, Bolognesi, Wittenberg, & Guertin, 2002). Their topology is a sandwich of two helices over two helices, 2/2 for short (Pesce et al., 2000) (Fig. 1B). Even smaller sequences are found in BLAST searches (Altschul, Gish, Miller, Myers, & Lipman, 1990), suggesting that 90–100 residues may be sufficient to wrap the helices in a globular shape and construct a spacious hydrophobic cavity. True haemoglobins have conserved a histidine, called proximal or F8, that anchors the haem group in the cavity. Haem-binding globins occur in a variety of quaternary structures (Flores et al., 2005; Royer, Knapp, Strand, & Heaslet, 2001), many cooperative and some large

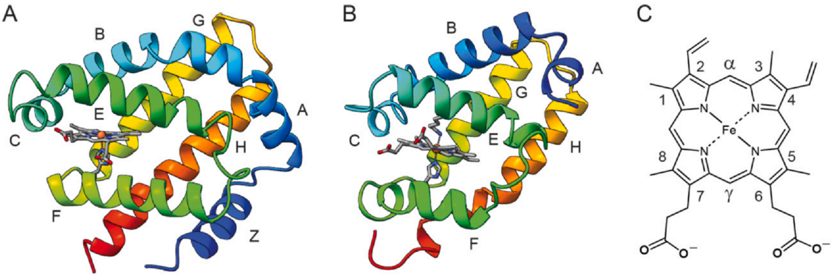


**Fig. 1** The two globin folds and the haem group. (A) Ribbon diagram of the 3/3 sandwich illustrated with *Thermosynechococcus elongatus* globin (PDB ID 8H17). Helices A, E, and F fold over helices B, G, and H. The haem group and proximal histidine are shown with sticks. This S globin has an additional helix, "pre-A" or Z. (B) Ribbon diagram of the 2/2 sandwich illustrated with *Chlamydomonas reinhardtii* THB4 (PDB ID 6BME). Helices B and E fold over helices G and H. The haem group, proximal histidine, and distal lysine are shown with sticks. In (A) and (B), the ribbons are coloured from N-terminus (blue) to C-terminus (red). Helices are labelled with the standard myoglobin nomenclature (Perutz, 1979). (C) The structure of the haem group. The haem is enveloped by the protein, its 6- and 7-carboxylate groups frequently the only functionalities exposed to solvent. Table 1 lists the PDB IDs of the proteins mentioned in the figures and text.

enough to warrant study by cryogenic electron microscopy (Afanasyev et al., 2017). They also appear as a module within polypeptide chains containing other globin and non-globin domains.

Over the years, repeated genome analyses have produced a robust tree composed of three monophyletic lineages (Schuster et al., 2024; Vinogradov et al., 2005). These lineages correspond to the myoglobin (M) family, the sensor globin (S) family, and the truncated globin (T) family, each appearing as single-domain or as multiple-domain proteins. The globin domain of M and S proteins adopts the 3/3 topology. The M family comprises mostly single domain proteins and two classes of chimeric proteins, androglobins and flavohaemoglobins. In contrast to the M family, the S family is mostly composed of multidomain proteins in which a globin, often N-terminal, serves as the sensing unit for one or more non-globin domains. Sensing units enable activities such as adenylate and diguanylate cyclase, histidine kinase, aerotaxis and oxygen sensing (Freitas, Saito, Hou, & Alam, 2005; Vávra, Sergunin, Jeřábek, Shimizu, & Martínková, 2022). The single domain proteins of the S family cluster into sensor globins (SSDgb) and protoglobins (Pgb). Their origin is likely the loss of non-globin domains from globin-coupled sensors over the course of evolution (Schuster et al., 2024). The T family is composed of 2/2 proteins grouped into

**Table 1** Structures mentioned in the text.

| Family | Organism | Protein | PDB or Alphafold2 ID | References |
|---|---|---|---|---|
| Haem-binding M family | | | | |
| SD | *Physeter catodon* | Oxymyoglobin | 1MBO | Phillips (1980) |
| | *Physeter catodon* | Biliverdin myoglobin | 1BVD | Wagner et al. (1995) |
| | *Riftia pachyptila* | C1 haemoglobin | 1YHU | Flores et al. (2005) |
| | *Lumbricus terrestris* | Haemoglobin (cryo-EM) | 5M3L | Afanasyev et al. (2017) |
| FH | *Escherichia coli* | Flavohaemoglobin | 1CQX | Ilari, Bonamore, Farina, Johnson, and Boffi (2002) |
| Fbg | *Aquifex aeolicus* | Thermoglobin | 7DIH | Muraki, Takeda, Nam, Muraki, and Aono (2021) |
| S family | | | | |
| GCS | *Geobacter sulfurreducens* | GCS globin domain | 2W31 | Pesce et al. (2009) |
| | *Bacillus subtilis* | HemAT globin domain | 1OR4 | Zhang and Phillips (2003) |
| SSDgb | *Thermosynechococcus elongatus* BP-1 | Single domain sensor | 8H17 | Mathur et al. (2023) |
| Pgb | *Methanosarcina acetivorans* C2A | Protoglobin | 2VEB | Nardini et al. (2008) |
| | *Aeropyrum pernix* | Protoglobin | 7UTE | Porter, Danelius, Gonen, and Arnold (2022) |

*(continued)*

**Table 1** Structures mentioned in the text. (*cont'd*)

| Family | Organism | Protein | PDB or Alphafold2 ID | References |
|---|---|---|---|---|
| T family | | | | |
| TrHb1 | *Chlamydomonas eugametos* | CTrHb | 1DLY | Pesce et al. (2000) |
| | *Synechocystis* sp. PCC 6803 | GlbN (NMR)<br>*bis*-His, no crosslink PTM | 1MWB | Falzone et al. (2002) |
| | *Synechocystis* sp. PCC 6803 | GlbN<br>*bis*-His, crosslink PTM | 1RTX | Hoy et al. (2004) |
| | *Synechocystis* sp. PCC 6803 | GlbN<br>His–Fe–CN, crosslink PTM | 1S69 | Trent et al. (2004) |
| | *Synechococcus* sp. PCC 7002 | GlbN<br>*bis*-His, crosslink PTM | 4MAX | Wenke, Lecomte, Héroux, and Schlessman (2014) |
| | *Synechococcus* sp. PCC 7002 | GlbN<br>His–Fe–CN, crosslink PTM | 4L2M | Wenke et al. (2014) |
| | *Chlamydomonas reinhardtii* | THB1<br>His–Fe–Lys | 4XDI | Rice et al. (2015) |
| | *Chlamydomonas reinhardtii* | THB1<br>His–Fe–CN | 6CII | Martinez Grundman et al. (2021) |

| | | | | |
|---|---|---|---|---|
| | *Chlamydomonas reinhardtii* | THB4<br>His–Fe–CN | 6BME | Johnson et al. (2018) |
| | *Chlamydomonas reinhardtii* | THB11<br>His–Fe–CN | 6TD7 | Huwald et al. (2020) |
| | *Shewanella benthica* | TrHb1 with *trans* haem *d* | 8VSH | |
| TrHb2 | *Mycobacterium tuberculosis* | Haemoglobin O | 1NGK | Milani et al. (2003) |
| TrHb3 | *Campylobacter jejuni* | trHbP | 2IG3 | Nardini et al. (2006) |
| Non-haem binding | | | | |
| Phycocyanin family | | | | |
| Phycocyanin | *Spirulina platensis* | Phycocyanin | 1HA7 | Padyana, Bhat, Madyastha, Rajashankar, and Ramakumar (2001) |
| | *Arthrospira platensis* | Allophycocyanin | 1ALL | Brejc, Ficner, Huber, and Steinbacher (1995) |
| | *Mastigocladus laminosus* | Phycoerythrocyanin | 2C7L | Schmidt, Krasselt, and Reuter (2006) |
| | *Griffithsia monilis* | Phycoerythin | 1B8D | Ritter et al. (1999) |

*(continued)*

**Table 1** Structures mentioned in the text. (*cont'd*)

| Family | Organism | Protein | PDB or Alphafold2 ID | References |
|---|---|---|---|---|
| Photoglobin | *Vibrio* sp. 10 N.286.49. B3 | B12-binding domain-containing protein, apo | A0A2N7DC63 | |
| RsbR family | | | | |
| RsbR | *Bacillus subtilis* | N-terminal RsbR | 2BNL | Murray et al. (2005) |
| | *Bacillus anthracis* | Sporulation inhibitor pXO1-118 | 3PMD | Stranzl et al. (2011) |

four clades, labelled TrHbN, TrHbO, TrHbP, and TrHbQ (Bustamante et al., 2016; Vuletich & Lecomte, 2006; Wittenberg et al., 2002) or, preferably, TrHb1 through TrHb4 (Vinogradov, Tinajero-Trejo, Poole, & Hoogewijs, 2013). Some T family proteins occur with a multiunit architecture and others as multidomain assemblies (Hade, Kaur, Chakraborti, & Dikshit, 2017; Nardini, Pesce, & Bolognesi, 2022).

Structural studies and sequence alignments have identified two additional groups of proteins that adopt the 3/3 globin fold but do not bind haem, represented by phycocyanins (Schirmer, Bode, Huber, Sidler, & Zuber, 1985) and environmental stress-signalling proteins (RsbR) (Murray, Delumeau, & Lewis, 2005) (Table 1). Phylogenetic studies support their relationship to the haem-binding globins and expand the superfamily to a five-branch tree (Schneider et al., 2022).

## 1.2 Chemical considerations

The haem-binding branches of the globin superfamily of proteins present a particular challenge for physiological assessment. The difficulty stems in large part from the chemical versatility of the haem group. The central iron is held in the nearly planar porphyrin core by bonds to each of the four pyrrole nitrogens (Fig. 1C). It can readily assume three oxidation states with different reactivities (Reeder, 2017): Fe(II), which is used to bind dioxygen reversibly, and two higher oxidation states, Fe(III) and Fe(IV), the latter encountered as a reaction intermediate. Fe(II) and Fe(III) haems can associate with and activate small molecules and ions; generate radicals and quench them; and act as a redox pair in electron transfer processes. The reactions that a particular globin may favour among many possibilities are ultimately selected and regulated by the protein matrix.

A common way by which haemoglobins adjust the chemistry of the haem group is through the coordination of the iron. Foremost is the stable ligation of the proximal histidine through its Nε atom. Histidine F8 occupies one of the two axial coordination sites, forming a five-coordinate (5c) square pyramid. Most haem-binding globins have the opposite or distal axial site open for an exogenous ligand, i.e., a non-proteinic molecule or ion. Given the propensity of Fe(II) and Fe(III) to complete their coordination sphere and achieve a six-coordinate (6c) octahedral geometry, it is not surprising that some proteins, either naturally (Couture et al., 2000; Johnson et al., 2014) or as a (mis)folding intermediate (Samuel, Smith, Phillips, & Olson, 2015; Samuel et al., 2020), provide a distal residue for one more ligation bond. Haemoglobins that have residues occupying both axial positions have been referred to as "hexacoordinate" in the literature (Kakar, Hoffman, Storz, Fabian, & Hargrove, 2010). Endogenous hexacoordination can alter the kinetics of exogenous ligand binding (Hafideddine et al., 2023; Smagghe, Sarath, Ross, Hilbert, & Hargrove, 2006) and tune the redox properties of the haem iron (Reedy, Elvekrog, & Gibney, 2008).

Regardless of whether a globin is involved in dioxygen pickup and delivery or chemical transformations, the makeup of the entire haem cavity is critical to establishing suitable rates of reaction so that substrate and electron fluxes mesh with the activities of other components of the cell. Four residues lining the distal side of the haem recur in groups of globins and influence ligand binding properties. According to Perutz notation, they occupy positions B10, E7, E11, and G8. In T globin, the residue at position E10 is also critical because it can serve as an axial ligand to the iron (Fig. 2A). Configurations with residues capable of hydrogen bonding, such as a tyrosine at B10 and a glutamine at E7 (Fig. 2B), tend to stabilise the dioxygen-bound state of the protein by decreasing the dissociation rate constant, $k_{\mathrm{off}}$, whereas hydrophobic residues at the same positions tend to accelerate ligand loss. Typical composition at essential positions of M, S, and T proteins are listed in Table 2.

Reactivity also depends subtly on the dynamic features of the protein (Estarellas et al., 2016). As observed for haemoglobin, reversible binding of

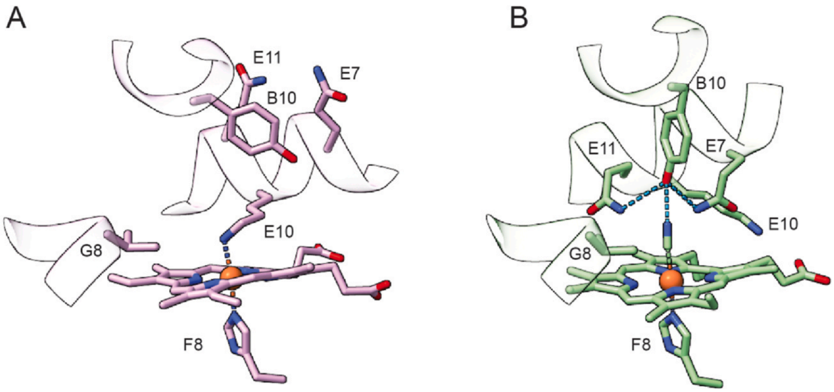


**Fig. 2** Key residues controlling haem chemistry. (A) The distal lysine-bound (hexacoordinate) state of ferric *C. reinhardtii* THB1 (PDB ID 4XDI). (B) The exogenous ligand-bound state of the same illustrated with the cyanide adduct (PDB ID 6CII). Residues B10, E7, E10, E11, F8 and G8 are shown with sticks along with the haem group. In (B), B10, E7, and E11 form a network of hydrogen bonds stabilising the cyanide ligand.

**Table 2** Typical composition at key locations in haem-binding globins.

| Family | Group | | B10 | CD1 | E7 | E10 | E11 | G8 | $O_2$ Affinity |
|---|---|---|---|---|---|---|---|---|---|
| M | Bacteria fungi | | Tyr | Phe | N/Q | | hp | hp | High |
| | Invertebrates vertebrates | | hp | Phe | H | | hp | hp | Moderate |
| | FHb | | Tyr | Phe | N/Q | | hp | hp | High |
| | Adgb | | hp | Phe | N/Q | | hp | hp | Moderate |
| S | Multidomain | | Tyr | Phe | hp | | N/Q/hp | | High |
| | Single domain | SSDgb | Tyr | Phe | R/K/- | | hp | | High-moderate |
| | | Pgb | Tyr | Phe | hp | | hp | | High-moderate |
| T | TrHb1 | Monophyletic | hp | Phe | hp | hp | hp | hp | Moderate-low |
| | | Paraphyletic | Y | Phe | N/Q/hp | H/K | N/Q/hp | hp | High |
| | TrHb2 | | Y | Tyr | S/T/hp | K/R | N/Q/hp | W | High |
| | TrHb3 | | Y | Phe | H | K | I | W | High-moderate |
| | TrHb4 | | Y | Phe | H | H/R | hp | R/K | High-moderate |

*hp*, Hydrophobic.
From Bustamante, J. P., Radusky, L., Boechi, L., Estrin, D. A., ten Have, A., & Martí, M. A. (2016). Evolutionary and functional relationships in the truncated hemoglobin family. *PLoS Computational Biology*, *12*(1), 26. https://doi.org/10.1371/journal.pcbi.1004701; Schuster, C. D., Salvatore, F., Moens, L., & Martí, M. A. (2024). Globin phylogeny, evolution and function, the newest update. *Proteins: Structure, Function, and Bioinformatics*, *2024*, 1–15. https://doi.org/10.1002/prot.26659.

ligands requires structural fluctuations (Perutz & Mathews, 1966) and so does cooperativity of ligand binding (Shibayama, 2020) and often enzymatic activity. Many microbial globins, in particular T globins (Milani et al., 2005) and protoglobins (Pesce, Bolognesi, & Nardini, 2013), possess networks of (gated) tunnels that allow access to the distal site and modulate the association rate constant, $k_{on}$, of exogenous ligands. The combination of tunnels and distal pocket composition results in variable affinity for $O_2$ and other molecules. Thus, inspection of structural models can discriminate between functions such as nitric oxide detoxification (high $O_2$ affinity and access for two substrates) and transport (moderate $O_2$ affinity) (Bustamante et al., 2016). Table 2 compiles anticipated $O_2$ affinity on the basis of structural features.

In the microbial world, where modern M, S, and T globins perform under the most diverse conditions and within host cells with the most diverse metabolism, it is expected that reversible dioxygen binding is only one of several possible roles that have evolved over billions of years. Even myoglobin, the textbook oxygen storage protein of vertebrates, assumes more complex roles than facilitated dioxygen diffusion. Its repertory now includes nitrite reduction and nitric oxide dioxygenation (Aboouf, Gorr, Hamdy, Gassmann, & Thiersch, 2023; Keppner et al., 2020). Presumably, these ancillary functions are tightly regulated in the crowded cellular environment. In general, active site composition, fluctuations and conformational rearrangements conspire to allow some processes and simultaneously inhibit others. Among confirmed microbial main cellular activities are the detoxification of nitric oxide by *Escherichia coli* flavohaemoglobin (Poole, 2020) and *Mycobacterium tuberculosis* haemoglobin N (a single domain T protein) (Carabet, Guertin, Lagüe, & Lamoureux, 2017), but others, listed in Table 3, have been demonstrated in vitro and proposed in vivo.

## 2. The haem-binding globins of cyanobacteria and green algae

Cyanobacteria (Bacteria) and unicellular green algae (Eukarya) harbour M, S, and T globins. For the purpose of this review, a search was performed for several haem-containing proteins expected to be found (or not) in cyanobacterial and green algal genomes. BLAST searches with default parameters (Altschul et al., 1990) were performed with the National Center for Biotechnology Information database of non-redundant protein

**Table 3** Some of the processes performed by haem-binding globins.

| | |
|---|---|
| Reversible $O_2$ binding | Hb–Fe(II) + $O_2$ ⇄ Hb–Fe(II)–$O_2$ |
| Autoxidation | Hb–Fe(II)–$O_2$ + 2 $H^+$ → Hb–Fe(III) + $H_2O_2$ |
| Nitrite reduction | Hb–Fe(II) + $NO_2^-$ + $H^+$ → Hb–Fe(III) + NO + $OH^-$ |
| NO dioxygenation | Hb–Fe(II)-$O_2$ + NO → Hb–Fe(III) + $NO_3^-$ |
| NO reduction | 2 Hb–Fe(II)–NO + 2 $H^+$ → 2 Hb–Fe(III) + $N_2O$ + $H_2O$ |
| Hydroxylamine reduction | 2 Hb–Fe(II) + $NH_2OH$ + 2 $H^+$ → 2 Hb–Fe (III) + $NH_4^+$ + $OH^-$ |
| Azanone (HNO) production | Hb–Fe(III)–NO + $H^-$ → Hb–Fe(II)–HNO |
| Catalase activity | Hb–Fe(III) + $H_2O_2$ → Hb–Fe(III) + $O_2$ + $H_2O$ |
| Peroxidase activity | Hb–Fe(III) + $H_2O_2$ → Hb–Fe(IV)=O + $H_2O$ |
| Hydrogen sulphide oxidation | Hb–Fe(III) + $H_2S$ → thiosulphate, persulphide, etc. |

sequences and limited to Cyanobacteria (taxid:1117), which included data for 921 organisms, or Chlorophyta (taxid:3041), which included data for 113 organisms.

## 2.1 Cyanobacterial haem-binding globins

### *2.1.1 T haemoglobins*

#### 2.1.1.1 *Nostoc commune* UTEX 584 GlbN

The discovery of cyanobacterial haemoglobin was reported in 1992 by Potts and colleagues (Potts, Angeloni, Ebel, & Bassam, 1992) who identified a truncated protein (group TrHb1) in the genome of *Nostoc commune* (order Nostocales). The gene, *glbN*, was found between two *nif* operons required for the synthesis of nitrogenase, and its expression was detected upon microaerobic nitrogen starvation. A protective role as a dioxygen scavenger was then proposed. Later characterisation of the purified protein and the cyanobacterium noted the moderate dioxygen affinity of GlbN (Thorsteinsson, Bevan, Ebel, Weber, & Potts, 1996) and its cytoplasmic localisation at the cell membrane (Hill et al., 1996). The nitrogenase-protective role of GlbN was then refined as delivering dioxygen to a terminal cytochrome oxidase.

The ability of GlbN to scavenge oxygen was recently applied to the phototrophic production of hydrogen. *Clostridium acetobutylicum*, among other anaerobic bacteria, is capable of relatively high hydrogen production using an [FeFe]-hydrogenase. This enzyme, like *N. commune*'s nitrogenase, is sensitive to oxygen (Hong & Pachter, 2012). Avilan and co-workers (Avilan et al., 2018) engineered *Nostoc* sp. PCC 7120 to express *C. acetobutylicum* hydrogenase in the cyanobacterial heterocysts, highly specialised cells that differentiate amidst vegetative cells when fixed nitrogen sources become limited. To deplete the heterocysts from oxygen generated by photosystem II in the adjacent vegetative cells, the authors included the gene for *N. commune* GlbN. The cyanobacterium engineered with GlbN indeed produced $H_2$ more efficiently than that without. This supports, albeit indirectly, the view that GlbN's function is to bind dioxygen reversibly for disposal.

Since the discovery of *N. commune* GlbN, several studies have focused on the relatives of this T protein, notably GlbN from *Synechocystis* sp. PCC 6803 and *Synechococcus* sp. PCC 7002, both model cyanobacteria. The two GlbN proteins, unlike many known only through genome sequencing, have been identified in the cell and are detected in omic data sets.

#### 2.1.1.2 *Synechocystis* sp. PCC 6803 GlbN

The first studies of *Synechocystis* sp. PCC 6803 GlbN (order Synechococcales, hereafter *Synechocystis* GlbN) appeared in 2000 (Couture et al., 2000; Scott & Lecomte, 2000). *Synechocystis* GlbN differed from *N. commune* GlbN in critical ways. Electronic absorption spectra indicated that the haem group was coordinated by two protein residues, which were identified as His46 (E10, distal) and His69 (F8, proximal), the former confirmed with the His46Ala replacement (Couture et al., 2000). Importantly, the oxygen dissociation constant was four times lower than that of *N. commune* GlbN, and because of this property, and the fact that *Synechocystis* does not fix nitrogen, a function different from oxygen scavenging was suspected.

Several studies of the purified protein followed. Besides the coordination of the iron by two histidine residues (yielding a *bis*-His, 6c complex), it was found that the protein prepared in the Fe(III) oxidation state, once reduced to the Fe(II) state, undergoes a spontaneous and irreversible post-translational modification (PTM). The result is a haem–protein crosslink preventing the loss of the cofactor (Vu, Jones, & Lecomte, 2002). This linkage is similar to that present in mature cytochrome *c*, except that the residue adding to a haem vinyl substituent is not a cysteine, but a histidine

(His117, at the end of the H helix). Structural information was obtained by nuclear magnetic resonance spectroscopy (Falzone, Vu, Scott, & Lecomte, 2002) (protein without the crosslink) and by X-ray diffraction (protein with the crosslink) (Hoy, Kundu, Trent, Ramaswamy, & Hargrove, 2004). GlbN exhibits the 2/2 fold with variations, for example a ~four-turn F helix and a kink in the H helix related to the haem–protein crosslink.

Purified GlbN has served as a platform for exploring the structural plasticity of the truncated fold. Displacement of the distal histidine (His46) by diatomic ligands such as cyanide illustrates a conformational rearrangement (Trent, Kundu, Hoy, & Hargrove, 2004) allowed by the ~9-residue loop connecting the E and F helices. Flexibility is also manifested by the His46Leu variant prepared in the Fe(III) state, which, despite the replacement, still forms a 6c, *bis*-His complex, but using His70 (proximal) and His117 (Nye & Lecomte, 2018). The unusual consequence of a single amino acid replacement emphasises the ability of the 2/2 fold to generate alternative haem binding sites. It also distinguishes at least some members of the T family from the 3/3 proteins, which tend to have relatively rigid structures and present few examples of coordination of the haem iron by a residue other than or in addition to F8.

The reactivity of *Synechocystis* GlbN was explored further for hints to potential function. Tests included NO dioxygenation to nitrate (Smagghe, Trent, & Hargrove, 2008), nitrite reduction to nitric oxide (Sturms, DiSpirito, & Hargrove, 2011), and hydroxylamine reduction to ammonia (Sturms, DiSpirito, Fulton, & Hargrove, 2011). These processes require a reducing agent or a reductase for multiple turnovers. In some situations, it is expected that the pool of ferredoxin serves the purpose (Scott et al., 2010). Recently, the existence of a cognate reductase was proposed (Uppal, Khan, & Kundu, 2020). The study used the sequence of the reductase that maintains soybean leghemoglobin in the Fe(II) state to query the genome of *Synechocystis* sp. PCC 6803. The search retrieved dihydrolipoamide dehydrogenase, the E3 component of pyruvate dehydrogenase. This ubiquitous class-I pyridine nucleotide-disulphide oxidoreductase, once purified, is capable of reducing Fe(III) GlbN and may provide another avenue for recycling oxidised GlbN to the Fe(II) state in the cell.

None of the above studies provide definitive in-the-cell information on the function of GlbN in *Synechocystis* sp. PCC 6803. However, recent engineering and transcriptomic analyses offer clues. In a study designed to probe the porphyrin biosynthetic pathway in *Synechocystis* sp. PCC 6803 and ultimately enhance the industrial production of haem and phycocyanin, Meng, Gao and coworkers created mutant strains of *Synechocystis* sp. PCC 6803 overexpressing genes associated with haem synthesis or use, either endogenous genes including *glbN*, or genes from *Synechococcus elongatus* PCC 7942 (Cao et al., 2023). Overexpression of *glbN* resulted in slower growth and a higher haem content than the wild-type strain by a factor of 2.5 in contrast to all other strains, which showed haem depletion. Overexpression of *glbN* also resulted in downregulation of *hemC* (hydroxymethylbilane synthase, now *HmbS* (Dailey et al., 2017)) and *hemN* (coproporphyrinogen dehydrogenase, now *CgdH*), whereas downregulation of *glbN* was observed with overexpression of *hemE* (uroporphyrinogen III decarboxylase, now *UroD*). The increase in haem content upon overexpression of *glbN* was attributed to the ability of the product to serve as a haem sink, although it was not specified whether the additional haem was cross-linked to the protein or not. It is perhaps relevant that neither of the two haem oxygenases, enzymes that process haem to produce photosynthetic pigments, were upregulated under those conditions. Regardless, this study suggests that GlbN participates in a feedback mechanism for haem synthesis.

In a transcriptomic study seeking to compare adaptation to different light spectra under anoxic conditions (Battistuzzi et al., 2024), a link was proposed between nitrogen metabolism and GlbN. When illuminated with far red light, *Synechocystis* sp. PCC 6803 upregulates the expression of *norB*, a nitric oxide reductase catalysing the conversion of NO to $N_2O$. A hypothesis to explain the need to remove NO despite the absence of an NO-producing nitrite reductase in the cyanobacterium (Büsch, Friedrich, & Cramm, 2002) implicates GlbN as responsible for NO generation. This is a documented activity of the protein in vitro under anoxic conditions (Sturms et al., 2011) and may well account for the *norB* observation.

#### 2.1.1.3 *Synechococcus* sp. PCC 7002 GlbN

*Synechococcus* sp. PCC 7002 (order Chroococcales) is an established model cyanobacterium that contains one *glbN* gene. The product (hereafter *Synechococcus* GlbN) is 59% identical in sequence to *Synechocystis* GlbN and shares many of its physico-chemical properties, including the ability to form a haem–protein crosslink as well as the coordination of the haem iron by both the proximal histidine and a distal histidine (His E10). The plasticity of the 2/2 fold is again apparent, revealed with the use of a modified haem cofactor. When dimethyl-esterified haem *b* is used instead of haem *b*, Lys E6 competes with His E10 for Fe(II) coordination (Nye, Preimesberger, Majumdar, & Lecomte, 2018).

Deletion of the *glbN* gene renders the synechococcal cells more susceptible to nitrate and nitric oxide challenges (Scott et al., 2010), and a role in the mediation of reactive nitrogen and oxygen species was therefore proposed. Analysis of cell extracts confirmed the formation of the haem–protein crosslink. Subsequent in vitro studies noted a difference in the reactivity of the protein with and without it (Preimesberger, Johnson, Nye, & Lecomte, 2017). Specifically, covalent attachment of the haem conveys protection from NO-mediated haem loss from the Fe(II) state. It also prevents the formation of azanone (HNO), a reactive compound that attacks thiols and is known to serve as a signalling agent in some organisms (Bianco, Toscano, Bartberger, & Fukuto, 2017).

In a network analysis of transcriptomic data (McClure et al., 2016), *glbN* is grouped with genes of the "defence and invasion systems." In the network, *glbN* has a single connection, to *gifB*, the product of which is the glutamine synthetase inactivating factor IF17 GifB (Bolay, Muro-Pastor, Florencio, & Klähn, 2018). Because glutamine synthetase is the key enzyme of nitrogen assimilation, the significance of the link to *glbN* may lie in a relationship to nitrogen metabolism as noted for *Synechocystis* GlbN.

#### 2.1.1.4 Other TrHbs

The cyanobacterial haemoglobins studied so far belong to the TrHb1 group. The available genomes of cyanobacteria contain fewer than twenty instances of TrHb2s, an abundant clade in bacteria (Vinogradov, Tinajero-Trejo, et al., 2013) and represented by *Mycobacterium tuberculosis* HbO (Ouellet, Milani, et al., 2007). HbO has a high affinity for dioxygen inconsistent with delivery and in line with its documented peroxidatic activity (Ouellet, Ranguelova, et al., 2007). To our knowledge, none of the cyanobacterial TrHb2s have been characterised. On the basis of amino acid composition at the essential positions mentioned in Section 1.2, cyanobacterial TrHb2s are expected to have similar catalase-peroxidase activity (Bustamante et al., 2016). Thus far, no instances of TrHb3 or TrHb4 has been detected in cyanobacteria.

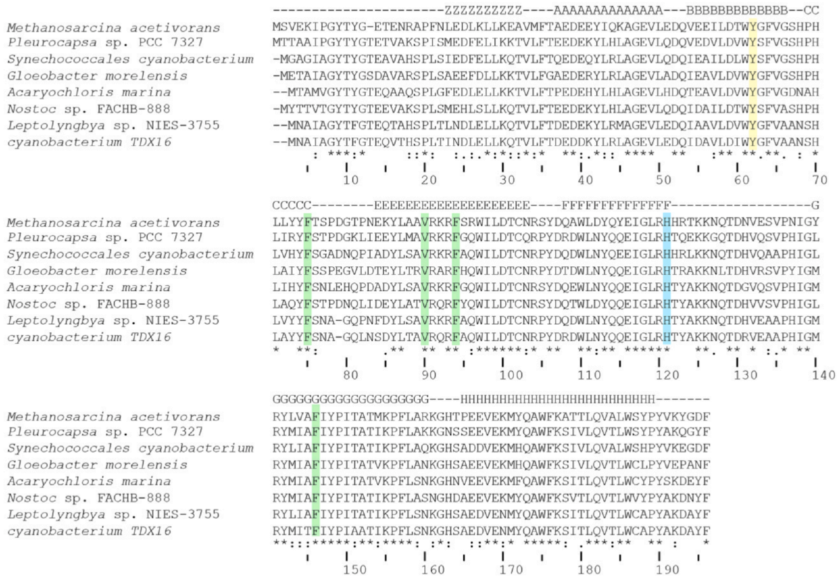
### *2.1.2 S haemoglobins*

Cyanobacteria are known to contain single domain globins from the S-family (Vinogradov, Tinajero-Trejo, et al., 2013). A search of cyanobacterial genomes was performed with the globin-sensor domain-containing protein of *Fischerella thermalis* JSC-11 (order Nostocales). The query (Uniprot accession G6FY51) returned a set of over 280 relatives. As expected, the

BLAST results show this group to align with the globin domain of the *Geobacter sulfurreducens* globin-coupled sensor (GCS). The sequences of the cyanobacterial proteins containing the GCS-like globin domain are highly conserved and only ~180-residue in length, i.e., single-domain globins (referred to as single domain sensor globins, SDSgb (Vinogradov et al., 2007), or sensor single domain globins, SSDgb (Schuster et al., 2024)). Their architecture is a ~150-residue globin portion bracketed by a short N-terminal segment and a C-terminal extension. AlphaFold2, a program designed to predict the three-dimensional structure of a protein on the basis of its amino acid sequence and selected Protein Data Bank structures (Jumper et al., 2021), has been remarkably successful at providing good accuracy three-dimensional models of proteins. The AlphaFold2 model of the query anticipates with very high confidence that the N-terminal and C-terminal extensions are helical. Residue conservation in these regions suggests that the secondary structure will be conserved across other cyanobacterial examples.

The S globin from the hyperthermophile *Thermosynechococcus elongatus* (order Acaryochloridales) shown in Fig. 1A belongs to the SSDgb group of globins. The protein was recently made recombinantly and crystallised in the imidazole-bound state (Mathur, Yadav, Yadav, Bhatt, & Kundu, 2023). The structure of the globin domain is consistent with that of *G. sulfurreducens* GCS (Pesce et al., 2009) (rmsd of 1.6 Å over 146/152 Cα, 32% identity) except that the presence of a phenylalanine at the distal position (E7) prohibits the endogenous hexacoordination observed in *G. sulfurreducens* GCS. The N-terminal and C-terminal extensions were not defined by the electron density, perhaps indicating that the AlphaFold2 model represents a plausible but sparsely populated conformer. In solution, *T. elongatus* globin is monomeric and capable of oxygen binding. The role of the protein in the organism remains to be elucidated.

A smaller group of about thirty single domain S proteins is identified by querying the cyanobacterial genomes with the protoglobin of the archaea *Methanosarcina acetivorans* (Freitas et al., 2004). Protoglobins are rare in cyanobacteria but have multiple representatives in archaebacteria. *M. acetivorans* protoglobin is a homodimer with an N-terminal extension preceding the pre-A (or Z) helix and a long loop between the F and G helices shielding the haem group from solvent (Nardini et al., 2008). Among the distinguishing characteristics of *M. acetivorans* protoglobin are a hydrophobic residue at position E7 (valine) and a ruffled haem group. *M. acetivorans* is a methanogenic archaebacterium that can use carbon monoxide as sole energy source for growth (Rother, Oelgeschläger, & Metcalf, 2007). The proposed functions of its protoglobin include sensing cellular levels of CO and detoxification of reactive oxygen and nitrogen species, the latter activity supported by the processing of nitrite by the Fe(III) protein to yield bound nitric oxide (Sgammato et al., 2023). Autoxidation of the oxygen-bound protein is rapid compared to globins dedicated to oxygen transport. The same traits are observed for the protoglobin of the extremophile archaea *Aeropyrum pernix* (Freitas et al., 2004), which also has unknown function. The *M. acetivorans* protoglobin-like proteins of cyanobacteria have highly conserved sequences and possess the essential valine at E7 (Fig. 3). It will be interesting to discover if the cyanobacterial protoglobins share chemical properties with archaeal proteins and how the structure adjusts these properties to different metabolic needs. Candidates for study include the proteins selected for Fig. 3.



```
                                ----------------------ZZZZZZZZZZ----AAAAAAAAAAAAAA---BBBBBBBBBBBBB--CC
Methanosarcina acetivorans      MSVEKIPGYTYG-ETENRAPFNLEDLKLLKEAVMFTAEDEEYIQKAGEVLEDQVEEILDTWYGFVGSHPH
Pleurocapsa sp. PCC 7327        MTTAAIPGYTYGTETVAKSPISMEDFELIKKTVLFTEEDEKYLHLAGEVLQDQVEDVLDVWYSFVGSHPH
Synechococcales cyanobacterium  -MGAGIAGYTYGTEAVAHSPLSIEDFELLKQTVLFTQEDEQYLRLAGEVLQDQIEAILDLWYSFVGSHPH
Gloeobacter morelensis          -METAIAGYTYGSDAVARSPLSAEEFDLLKKTVLFGAEDERYLRLAGEVLADQIEAVLDVWYGFVGSHPH
Acaryochloris marina            --MTAMVGYTYGTEQAAQSPLGFEDLELLKKTVLFTPEDEHYLHLAGEVLHDQTEAVLDVWYGFVGDNAH
Nostoc sp. FACHB-888            -MYTTVTGYTYGTEEVAKSPLSMEHLSLLKQTVLFTEEDEKYLHLAGEVLQDQIDAILDTWYSFVASHPH
Leptolyngbya sp. NIES-3755      --MNAIAGYTFGTEQTAHSPLTLNDLELLKQTVLFTDEDEKYLRMAGEVLEDQIAAVLDVWYGFVAANSH
cyanobacterium TDX16            --MNAIAGYTFGTEQVTHSPLTINDLELLKQTVLFTDEDDKYLRLAGEVLEDQIDAVLDIWYGFVAANSH
                                     : ***:*     ::*:  :.:.*:*::*:*  **:.*:: ****  **   :** **.**. : *
                                         10        20        30        40        50        60        70

                                CCCCC--------EEEEEEEEEEEEEEEEEEEE----FFFFFFFFFFFFFF------------------G
Methanosarcina acetivorans      LLYYFTSPDGTPNEKYLAAVRKRFSRWILDTCNRSYDQAWLDYQYEIGLRHHRTKKNQTDNVESVPNIGY
Pleurocapsa sp. PCC 7327        LIRYFSTPDGKLIEEYLMAVRKRFGQWILDTCQRPYDRDWLNYQQEIGLRHTQEKKGQTDHVQSVPHIGL
Synechococcales cyanobacterium  LVHYFSGADNQPIADYLSAVRKRFAQWILDTCNRPYDQDWLNYQEEIGLRHHRLKKNQTDHVQAAPHIGL
Gloeobacter morelensis          LAIYFSSPEGVLDTEYLTRVRARFHQWILDTCNRPYDTDWLNYQQEIGLRHTRAKKNLTDHVRSVPYIGM
Acaryochloris marina            LIHYFSNLEHQPDADYLSAVRKRFGQWILDTCNRSYDQEWLNYQQEIGLRHTYAKKNQTDGVQSVPHIGM
Nostoc sp. FACHB-888            LAQYFSTPDNQLIDEYLATVRQRFYQWILDTCNRSYDQTWLDYQQEIGLRHTYAKKNQTDHVVSVPHIGL
Leptolyngbya sp. NIES-3755      LVYYFSNA-GQPNFDYLSAVRKRFAQWILDTCNRPYDQDWLNYQQEIGLRHTYAKKNQTDHVEAAPHIGM
cyanobacterium TDX16            LAYYFSNA-GQLNSDYLTAVRQRFAQWILDTCNRPYDRDWLNYQQEIGLRHTYAKKNQTDRVEAAPHIGM
                                *  **:        .**  ** ** :******:* **  **:** ******   **. ** * :.* **
                                         80        90        100       110       120       130       140

                                GGGGGGGGGGGGGGGGGGGG----HHHHHHHHHHHHHHHHHHHHHHHHH-------
Methanosarcina acetivorans      RYLVAFIYPITATMKPFLARKGHTPEEVEKMYQAWFKATTLQVALWSYPYVKYGDF
Pleurocapsa sp. PCC 7327        RYMIAFIYPITATIKPFLAKKGNSSEEVEKMYQAWFKSIVLQVTLWSYPYAKQGYF
Synechococcales cyanobacterium  RYLIAFIYPITATIKPFLAQKGHSADDVEKMHQAWFKSIVLQVALWSHPYVKEGDF
Gloeobacter morelensis          RYMIAFIYPITATVKPFLANKGHSAEEVEKMHQAWFKSIVLQVTLWCLPYVEPANF
Acaryochloris marina            RYMIAFIYPITATVKPFLSNKGHNVEEVEKMFQAWFKSITLQVTLWCYPYSKDEYF
Nostoc sp. FACHB-888            RYMIAFIYPITATIKPFLASNGHDAEEVEKMYQAWFKSVTLQVTLWVYPYAKDNYF
Leptolyngbya sp. NIES-3755      RYLIAFIYPITATIKPFLSNKGHSAEDVENMYQAWFKSITLQVTLWCAPYAKDAYF
cyanobacterium TDX16            RYMITFIYPIAATIKPFLSNKGHSAEDVENMYQAWFKSITLQVTLWCAPYAKDAYF
                                **:::*****:**:****: :*:  ::**:* *****: .***:**  ** :   *
                                         150       160       170       180       190
```

**Fig. 3** Cyanobacterial protoglobin sequence alignment. The sequences were aligned with the sequence of *Methanosarcina acetivorans* protoglobin (Uniprot accession Q8TLY9) using Clustal Omega (Sievers & Higgins, 2021). The top row indicates the locations of helices. Note the insertion between the F and G helices. Essential positions are marked: B10 (yellow), CD1, E7, E11, G8 (green) and F8 (cyan). Sequences are *Methanosarcina acetivorans* (WP_011022833.1), *Pleurocapsa* sp. PCC 7327 (order Pleurocapsales; AFY77222.1), *Acaryochloris marina* (order Acaryochloridales; WP_261891958.1), *Synechococcales cyanobacterium* C42_A2020_086 (order Synechococcales; MBF2074950.1), *Gloeobacter morelensis* (order Gloeobacterales; WP_230841849.1), *Nostoc* sp. FACHB-888 (order Nostocales; MBD2247389.1), *Leptolyngbya* sp. NIES-3755 (order Leptolyngbiales; BAU15401.1), cyanobacterium TDX16 (OWY64285.1).

### *2.1.3 M haemoglobins*

Some cyanobacteria contain M globins (Vinogradov, Tinajero-Trejo, et al., 2013). The M lineage includes flavohaemoglobins, bacterial proteins specialising in the detoxification of nitric oxide (Poole, 2020) and made out of three modules for efficient catalysis, an N-terminal globin, an NAD-binding module, and a flavin-binding domain. Flavohaemoglobins as such are rare if existing in cyanobacteria; however, the globin domain of flavohaemoglobin is recognised within cyanobacterial genomes, either as single domains or fused to other domains. Hundreds such sequences are retrieved with BLAST searches.

The nitric oxide dioxygenase from *Fischerella thermalis* JSC-11 (Uniprot accession G6FNK7) is an example of cyanobacterial single domain M protein related to flavohaemoglobin or Fgb (Vinogradov, Tinajero-Trejo, et al., 2013). Among the relatives of this Fgb detected in a BLAST search are single domain cyanobacterial globins and the thermoglobin from the hyperthermophile *Aquifex aeolicus* (Miranda, Maillett, Soman, & Olson, 2005). This particular thermoglobin is monomeric and has a high affinity for oxygen. Speculation concerning its function includes detoxification of nitric oxide or excess dioxygen. Some of the single domain cyanobacterial globins with homology to nitric oxide dioxygenase *F. thermalis* G6FNK7 have been annotated as bacitracin resistance protein BacA. This enzyme is an undecaprenyl pyrophosphate phosphatase; it has an all-α helical fold and resides in the cytoplasmic membrane (El Ghachi et al., 2018). Comparison of genuine BacA to other protein folds using the program Dali (Holm & Rosenström, 2010) shows a low level of three-dimensional similarity to TrHb1 (e.g., the globin from *Shewanella benthica* KT99, PDB ID 7TT9, with a Z score of 3.0 and over only 62 residues). Relatives of G6FNK7 annotated as BacA are predicted to adopt a globin fold, suggesting that the automatic BacA annotation is misleading.

At the time of this writing, no experimental studies of cyanobacterial Fgbs seemed available. However, the high degree of sequence conservation and the relation to *A. aeolicus* thermoglobin offer insight. The consensus sequence derived from the 250 cyanobacterial entries aligning with *F. thermalis* G6FNK7 has highly conserved occupancies at the key positions controlling dioxygen affinity (Fig. 4). These are also found in thermoglobin. Thus, similar chemical properties may be expected, including a strong affinity for dioxygen.

An additional number of M-family globins retrieved with *F. thermalis* G6FNK7 as query are part of long polypeptide chains (>400 residues) ending

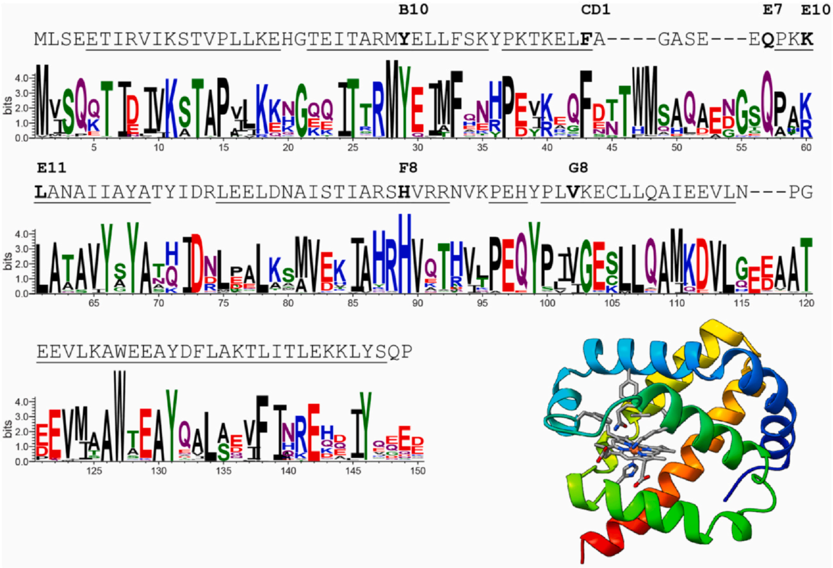


**Fig. 4** Consensus sequence of cyanobacterial single domain M globins. A total of 250 sequences were retrieved with a BLAST (Altschul et al., 1990) search using the nitric oxide dioxygenase of *Fischerella thermalis* JSC-11 (Uniprot accession G6FNK7) as query and aligned with Clustal Omega (Sievers & Higgins, 2021). The high level of conservation is illustrated with the logo plot (Crooks, Hon, Chandonia, & Brenner, 2004). The sequence of *Aquifex aeolicus* thermoglobin is shown above. Helical segments are underlined and key positions marked in bold face. The ribbon structure of *A. aeolicus* thermoglobin (PDB ID 7DIH) is included with key residues shown in sticks.

with a C-terminal pentapeptide repeat (PPR). PPRs are most commonly found in cyanobacteria (Bateman, Murzin, & Teichmann, 1998). They are ubiquitous and arise in different number of repeats and therefore lengths. They form right-handed β-helical structures with largely unknown function (Zhang & Kennedy, 2021). In certain contexts, the PPR domain may serve as a DNA mimic; in others, it may simply serve as a structural scaffold supporting intermolecular assemblies (Zhang & Kennedy, 2021). The globin domain is also found in association with an N-terminal cupin domain. Cupin domains have a β fold, typically bind a metal ion, most frequently iron but also others, and have diverse functions dependent on the presence and nature of the metal (Dunwell, Purvis, & Khuri, 2004). The cupin domain fused to cyanobacterial globins features two histidines found in the cupin metal binding motifs, and no other obvious potential metal ligand. Proteins containing the cupin domain are also annotated as flavohaemoglobins. Without further information, however, it is difficult to formulate hypotheses as to the purpose of these different multidomain M proteins.

M globin domains also appear in nitric oxide synthase oxygenase, proteins that oxidise arginine to synthesise nitric oxide, and in adenylate/guanylate cyclases, proteins that transform adenosine or guanosine 5′-triphosphate into the corresponding 3′,5′-cyclic monophosphate. In those contexts, it is possible that the globin module either performs chemistry or transfers electrons.

## 2.2 Chlorophytic haem-binding globins

Among green algae, *Chlamydomonas reinhardtii* has a long history of use as a model organism for photosynthesis and other biological processes (Dupuis & Merchant, 2023). Interest in eukaryotic algae has also been driven by their ability to synthesise diverse natural products, some extremely toxic. Whereas bacterial genome sequencing has progressed rapidly and yielded numerous putative globins, microalgal sequencing has proven to be technically challenging. Few chlorophytic whole genomes are available (O'Neill, 2020). The Plant Comparative Genomics portal of the Department of Energy's Joint Genome Institute, the major database for Chlorophyte genomes through Phytozome 13, lists only ten entries, from the classes Chlorophyceae, Trebouxiophyceae and Mamiellophyceae. With its 114 genomes the NCBI broadens species representation. The majority of chlorophytic globins are T globins (TrHb1s and TrHb2s) and single domain M globins. Thus far, no example of S globin has been found in green algae.

### *2.2.1 T haemoglobins*

#### 2.2.1.1 *Chlamydomonas* TrHb1s

The presence of haem-biding globins in the chloroplast of a green alga was noted thirty years ago when Couture and colleagues detected two light-inducible globin genes (LI410 and LI637) in *Chlamydomonas eugametos* (Couture, Chamberland, St-Pierre, Lafontaine, & Guertin, 1994). The atomic resolution structure of LI637, published in 2000 (Pesce et al., 2000), was not only the first of a chlorophytic globin but also that of a GlbN relative; it provided a glimpse into the T lineage of the superfamily. Years later, studies of the model organism *Chlamydomonas reinhardtii* revealed several genes for globin domains (Merchant et al., 2007; Vinogradov, Fernandez, Hoogewijs, & Arredondo-Peter, 2011). Proteomic datasets (Lechtreck et al., 2009; Mühlhaus, Weiss, Hemme, Sommer, & Schroda, 2011; Terashima, Specht, Naumann, & Hippler, 2010; Wang et al., 2014) detected the presence of the gene products in the cell. In all, *C. reinhardtii* expresses twelve globins, named THB1–12 for convenience (Hemschemeier et al., 2013), with architecture shown in Fig. 5. From this green alga, three additional structures of TrHb1s have been determined. These are from THB1 (Martinez Grundman et al., 2021; Rice et al., 2015), THB4 (Johnson, Russo, Nye, Schlessman, & Lecomte, 2018), and THB11 (Huwald et al., 2020). The *C. reinhardtii* THBs for which no experimental structure is available can, for the time being, be approximated by AlphaFold2, albeit as apoproteins biased to the holoprotein forms deposited in the Protein Data Bank and unable to inform on haem coordination.

THB1 binds the haem with the proximal histidine and a distal lysine (Johnson et al., 2014), an interesting feature because it requires the lysine to be in the neutral state under physiological conditions and raises questions of pH response and protein dynamics. In-depth experimental and computational studies have implicated a complex network of interactions among ionisable side chains and haem carboxylates in the control of the Lys-on/Lys-off equilibrium and the redox potential of the iron (Julió Plana, Grundman, Estrin, Lecomte, & Capece, 2021; Martinez Grundman et al., 2021). It is with THB1 that the versatility of haem-binding globins is apparent. In the presence of oxygen, THB1 acts as a dioxygenase by converting NO into $NO_3^-$ (Johnson & Lecomte, 2015; Johnson et al., 2014; Sanz-Luque, Ocaña-Calahorro, de Montaigu, et al., 2015). Under

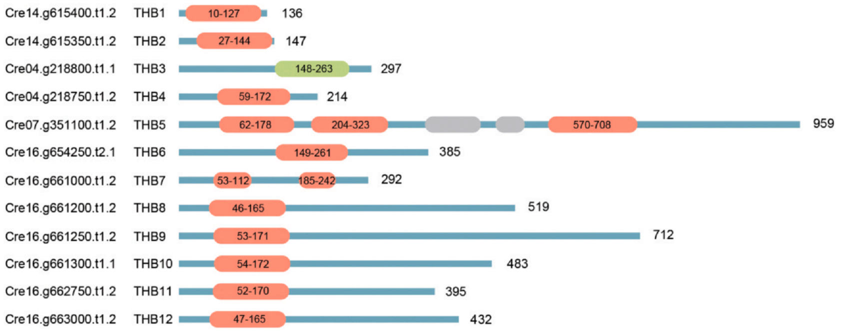


**Fig. 5** The truncated haemoglobins of *C. reinhardtii*. The transcripts are from Phytozome 13, along with total chain length and location of the globin domains. Only THB1–4 are named as such in the database; the names THB5–12 are used for convenience (Hemschemeier et al., 2013). The red domains have the proximal histidine. The green domain does not. The sequence of THB5 contains a short region (380–466) also annotated as a globin and possibly completed by region 489–534. This additional domain is in grey.

anaerobic conditions, THB1 is able to reduce nitrite to NO in vitro with high efficiency (Ciaccio et al., 2015).

NO dioxygenation and nitrite reduction begin with Fe(II) THB1 and leave the globin in the Fe(III) state. The coordination of the haem iron with His F8 and Lys E10 facilitates reduction by setting the redox potential to a moderately negative value and minimising protein and solvation rearrangement during electron exchange (Johnson et al., 2014, 2018). NAD(P)H fails to reduce THB1, but proteins such as cytochrome $b_5$ reductase, dihydrolipoamide dehydrogenase and monodehydroascorbate reductase may serve the purpose (Shandilya et al., 2021). Of particular significance is that nitrate reductase, an NADH-dependent multidomain enzyme, can shunt electrons to THB1 (Johnson et al., 2018; Sanz-Luque, Chamizo-Ampudia, Llamas, Galvan, & Fernandez, 2015; Sanz-Luque, Ocaña-Calahorro, de Montaigu, et al., 2015) before they reach the nitrate reduction active site, essentially modulating the production of nitrite. Added to these in-vitro observations, *THB1* expression is upregulated by both NO (Sanz-Luque, Ocaña-Calahorro, Galván, & Fernández, 2015) and nitrite (Plouviez et al., 2017). Control of nitrate, nitrite, and the messenger NO (León & Costa-Broseta, 2020) is essential for economical nitrogen utilisation (Bellido-Pedraza et al., 2020). A picture emerges by which THB1 participates in regulating the balance of nitrogen species in the cell (Calatrava, Tejada-Jimenez, Sanz-Luque, Fernandez, & Galvan, 2023) by engaging in a nitrate-to-nitrate cycle with nitrite and NO as intermediates. In this scheme, the NO-forming nitrite reductase (NOFNiR) generates NO.

THB2 has properties similar to those of THB1. It is soluble, coordinates the iron with His F8 and Lys E10, and displays NO dioxygenase activity (Johnson et al., 2018). THB2 also has nitrite reductase activity although it is half as efficient as THB1 under the same conditions (Ciaccio et al., 2015). In contrast to *THB1*, *THB2* expression is downregulated by NO (Sanz-Luque, Ocaña-Calahorro, Galván, et al., 2015) and nitrite (Plouviez et al., 2017). Under phosphorus deficiency, THB1 and THB2 control intracellular NO levels (Filina, Grinko, & Ermilova, 2019; Grinko, Alqoubaili, Lapina, & Ermilova, 2021), presumably through an NO-dependent signalling pathway. In another quantitative RT-PCR study, *THB1*, *THB2*, and *THB12* were found to be induced under conditions not only of phosphorus deprivation but also of sulphur deprivation (Minaeva, Zalutskaya, Filina, & Ermilova, 2017) indicating that THB1 plays a role in the induction of the S-limitation response. Coordination of sulphate limitation and nitrate assimilation emerges as an additional function of THB1, whereas the function of THB2 appears strongly linked to phosphorus limitation (Grinko et al., 2021).

THB3 has so far resisted functional speculation. The revised sequence of the protein (Uniprot accession R9S068) is 297 residues in length, with the globin domain spanning positions 148 to 268. Although it aligns well with other TrHb1s, the presence of a glutamine at F8, where the proximal histidine normally resides, raises questions about expression, prosthetic group binding and ultimately function. A genome-wide expression analysis of the singlet oxygen acclimation response of *C. reinhardtii* detected transcripts of THB3 (Fischer et al., 2012). Compared to wild-type, an attenuation of *THB3* expression was measured when the singlet oxygen acclimation protein SAK1 was knocked-out. Even though no explanation is offered for the observation, the study confirms that the cell produces *THB3* transcripts and that further investigation of the protein is warranted.

The apoglobin domain of THB3 (apoTHB3) was expressed recombinantly in *E. coli* (Johnson et al., 2018). It is highly helical, thermally stable, and associates readily with haem, but with only micromolar affinity. Replacement of His191 at position E10 with alanine weakens this affinity and argues for the haem occupying the canonical binding site (Johnson et al., 2018). However, the identity of the cofactor present in cellular THB3 has not been determined. The micromolar haem affinity could signify a function as haem transporter or temporary store. It is also possible that the protein binds linear tetrapyrroles or lipids. Interestingly, the non-globin N-terminal portion has homology only to a few Chlamydomonales sequences. The role of this extension, predicted to be partially helical, is unknown.

THB3 has algal relatives also lacking the proximal histidine, and large scale genomic analysis of the entire haemoglobin superfamily reveals that 8% of S proteins and a small percentage of M proteins also have the same feature (Schuster et al., 2024). Thus, if sequencing accuracy is not in question and the genes are expressed, there exists a type of proteins that deviates from the standard haem-binding globins principally by the replacement of the proximal histidine, and it is safe to assume that these proteins are functionally relevant. THB3 offers an opportunity to inspect possible roles for this large group of proteins.

The globin domain of THB4 was investigated in vitro as a plastid-located (Terashima et al., 2010) orthologue of *C. eugametos* LI637 associated with the diurnal cycle of *C. reinhardtii* (Zones, Blaby, Merchant, & Umen, 2015). Along with THB1 and THB2, THB4 shows His–Lys coordination of the haem iron (Johnson et al., 2018), but the Fe(III) state is remarkably insensitive to pH despite the ligation of the distal lysine. THB4 may be under the control of the NIT2 transcription factor (Sanz-Luque, Ocaña-Calahorro, de Montaigu, et al., 2015); it has NO dioxygenation activity, although it is more readily damaged in the process than THB1 and THB2 (Johnson et al., 2018). THB4 is also an efficient nitrite reductase (Ciaccio et al., 2015). Differences in gene regulation, reduction rate, reduction potential, stability of Fe(III) and Fe(II) states, and catalytic activity suggest that the three proteins are not redundant in the cell.

The gene encoding THB8 is strongly up-regulated under anoxic conditions in the light. Knocking down THB8 inhibits cell growth under the same conditions, but the effect is not remedied by NO scavengers, which points to an NO-dependent signal transduction role rather than dioxygenation like THB1 and THB2 (Hemschemeier et al., 2013). Besides the globin domain, THB8 has a long (>300 residues) C-terminal extension. Although local helix formation may occur, it appears that this part of the protein is disordered (Johnson & Lecomte, 2015). The role of this extension and its influence on the globin domain are not known.

THB11 also has terminal extensions. In this 395-residue protein, the globin domain spans positions 52–170 (Fig. 5). Both N- and C-terminal regions play a role in oligomerization according to size-exclusion chromatography and native polyacrylamide gel electrophoresis data (Huwald et al., 2020). The same study compared the nitrite reduction rate of the globin within the full-length protein and truncated versions and determined that the C-terminal extension enhances reactivity (Huwald et al., 2020). Nitrite reduction, an activity of the non-symbiotic *Arabidopsis thaliana* haemoglobin 2 (Tiso, Tejero, Kenney, Frizzell, & Gladwin, 2012), is a suspected function for THB11.

Spectral characterisation of THB11 shows that the protein adopts a 5c haem ligation (Huwald et al., 2020) even though it does have a lysine at position E10. This observation emphasises that iron coordination needs to be determined experimentally for each protein under study (Martinez Grundman, Johnson, & Lecomte, 2023) with allowance for ligand switching behaviour as illustrated by *C. eugametos* LI637 (Das et al., 1999) and GlbN (Nye et al., 2018). The crystal structure of the globin domain of THB11 with bound cyanide (PDB ID 6TD7) (Huwald et al., 2020) differs from the structure of THB1 in the same ligation state (PDB ID 6CII). The haem group, which has pseudo-$C_{2v}$ symmetry, is rotated by 180° about the $\alpha$–$\gamma$ axis (Fig. 1), directing its 2-vinyl group toward position G8. Perhaps related to this is the presence of a glycine at G8, where a preponderance of valine, isoleucine, and phenylalanine makes contact with the haem in other TrHb1s. Whereas the E helix in TrHb1s is generally regular, the E helix of the crystalline THB11 globin is interrupted in the middle and forms a ~60° bend. The origin of this unusual feature is unclear, but it is interesting that the distortion is located where the GH corner of a symmetry-related molecule makes contact in the crystal lattice.

The distinct features of the haem environment within THB11 are likely to affect the chemical properties of the protein. To perform functions dependent on substrate or ligand binding to the haem, access paths must be available. The porous structure of T globins facilitating the entry and exit of small molecules (Milani et al., 2004) was mentioned in Section 1.2. Although tunnel networks vary from protein to protein, three main branches are typically observed in structures determined with a bound exogenous ligand such as cyanide. The branches are referred to as the long tunnel (LT), the short tunnel G8 (STG8) and the E7 gate (E7G) (Bustamante et al., 2016) (Fig. 6). The residues at H5, B2, H9, E15, E11 and G8 control passage through the long tunnel, which is parallel to the H helix; the residues at H9, G8 and G9 control the short tunnel, perpendicular to the long tunnel and exiting between helices G and H; the E7 gate derives its properties from the positions at B10, CD1, E7 and E11, and runs opposed to the short tunnel. In THB11, the lining of the tunnels is mostly hydrophobic as in other TrHb1s. However, the glycine at G8 is expected to lower the barrier to exit through the long tunnel (Bustamante et al., 2016). The short tunnel is blocked by a

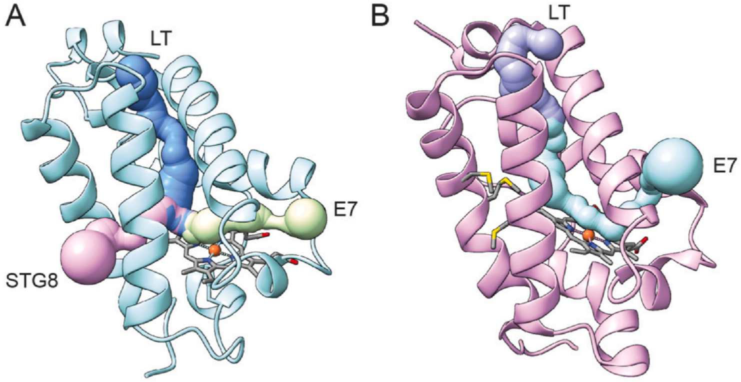


**Fig. 6** Access tunnels in *M. tuberculosis* HbN (PDB ID 1RTE, chain A) and *C. reinhardtii* THB11 (PDB ID 6TD7). (A) The three tunnels of HbN, LT (blue), STG8 (pink), and E7 (green). (B) The two tunnels of THB11, LT (blue) and E7 (cyan). A cluster of three methionines (sticks) blocks the STG8 branch.

cluster of three methionines (at G9, H6, and H9), which suggests limited access to the haem group with potential kinetic consequences.

Comparatively little is known about the remaining *C. reinhardtii* THBs. The *THB5* gene codes for a string of three or four TrHb1 domains with intervening regions for a total length of 958 residues (Fig. 5). The four domains have pairwise sequence identity between 45% and 20%. AlphaFold2, which returns plausible structures for other TrHb1s, predicts the 2/2 topology in the first two domains with reasonable confidence, whereas the quality of the result is low for the last two. These predictions nevertheless support a multiunit TrHb architecture. THB6 has a single globin domain located toward the middle of a 385-residue protein. The structure of the domain itself is predicted with high confidence, whereas the extensions may sample some helical structure. AlphaFold2 also generates a believable representation of THB7 despite a ~70-residue insertion between the E and F helices. Experimental data support haem binding (Huwald, Schrapers, Kositzki, Haumann, & Hemschemeier, 2015), but with uncharacteristic Fe (III) electronic absorption spectrum reminiscent and guanylate cyclase (Stone, Sands, Dunham, & Marletta, 1996). Likewise, THB9, THB10 and THB12 all are predicted to have a folded globin domain and disordered extensions. Expression of *THB7* and *THB10* increases under dark anoxic conditions and does not depend on nitrate. The globin domain of THB9 has been prepared and has electronic absorption spectra consistent with 5c ligation in the Fe(III) and Fe(II) states (Shandilya et al., 2021). The same was observed for THB10 (Huwald et al., 2015). Deciphering further the roles of each THB will require the use of multiple strains and conditions to uncover relationship to specific branches of metabolism.

Studies of heterologously expressed and purified proteins are needed. However, they typically rest on several assumptions. First, the nature of the cofactor is taken to be haem *b*, but this requires verification given the propensity for apoglobins to bind different molecules including biliverdin (Wagner, Muller, Schmitzberger, Falk, & Kratky, 1995) and chlorophyll derivatives (Marković, Pröll, Bubenzer, & Scheer, 2007) or to modify their own haem *b* (e.g., haem–protein crosslink in GlbN and formation of haem *d* in *Shewanella benthica* TrHb1). For THB1, it was possible to extract material from *C. reinhardtii* cells and detect the haem group (Johnson & Lecomte, 2014). Such information is not available for other THBs and is especially important for THB3, the atypical TrHb1. Other assumptions are also made. For example, THB1 is N-acetylated (Lechtreck et al., 2009) and others may be as well. The modification can have consequences for the behaviour of the protein (Drazic, Myklebust, Ree, & Arnesen, 2016). The fact that several THBs contain disordered extensions is suspicious. These regions have a number of experimentally detected phosphorylation sites (Wang et al., 2014) and more may yet be found (Thapa et al., 2021). The (de)phosphorylation of these intrinsically disordered regions could play a regulatory role (Newcombe, Delaforge, Hartmann-Petersen, Skriver, & Kragelund, 2022). Lastly, reactivity in vitro is only a guide for potential functions in vivo. Setting aside these caveats, the demonstrated in-vitro NO dioxygenase and nitrite reductase activities of at least some THBs are in agreement with participation in *C. reinhardtii* nitrogen metabolism and response to nitrosative stress. The model organism *C. reinhardtii* transformed biology (Sasso, Stibor, Mittag, & Grossman, 2018). It is fitting that with twelve haemoglobins, it adds novel information to our understanding of the truncated lineage of the superfamily.

#### 2.2.1.2 Other TrHbs

Single or multiple TrHb1s are detected in each of the ten Phytozome 13 fully annotated genomes mentioned in Section 2.2, except that of the Mamiellophycea *Ostreococcus lucimarinus* as previously noted (Johnson & Lecomte, 2015; Vinogradov et al., 2011). Additional instances are found in another ~30 chlorophytes when the NCBI database is queried with *C. reinhardtii* TrHb1 sequences. In contrast to TrHb1s, TrHb2s are rare in green algae (Becana et al., 2020), with only a few examples detected in *Chlorella* sp. The reverse applies to plants, TrHb2s being common and TrHb1s rare (Vinogradov et al., 2011). *Arabidopsis thaliana* Glb-3 (Watts et al., 2001) is a well-studied plant TrHb2. Glb-3 has unusual, concentration-independent CO rebinding kinetics following photolysis (Watts et al., 2001). Compared to other TrHb2s, Glb-3 has an N-terminal extension that secures the dimeric quaternary structure (Reeder & Hough, 2014). The origin of the two types of T globins in chlorophytes derives from different bacterial ancestors (Vinogradov et al., 2011), and generalisation from Glb-3 may not be warranted. *M. tuberculosis* HbO mentioned in Section 2.1.1 may be a better model and suggests catalase-peroxidase activity.

### *2.2.2 M haemoglobins*

Within Phytozome 13, M globins of the Fgb group are present in the two Micromonas representatives. A broader search through the NCBI Chlorophyta phylum detected three additional species (*Bathycoccus prasinos*, *Edaphochlamys debaryana*, and *Gonium pectorale*) carrying Fgb genes. Thus, chlorophytic

M globins are still found only in members of the Mamiellophyceae, Prasinophyceae, and Trebouxiophyceae classes (Vinogradov, Bailly, et al., 2013). Knowledge of chlorophytic M globins is limited to genetic sequence, and confirmation at the protein level will be required to understand their physiologic significance.

## 3. The linear tetrapyrrole-binding globins of cyanobacteria and algae

Cyanobacteria were the first known organisms to perform oxygenic photosynthesis. To harvest the energy of the light and transfer it to the photosystems, these bacteria resort to phycobilisomes, enormous supramolecular assemblies of ordered phycobiliproteins (Bryant & Gisriel, 2024). The light-harvesting pigment within phycobiliproteins is formed from phycocyanobilin, a linear tetrapyrrole derived from haem and covalently attached to the proteins through one or two cysteines. Phycobiliproteins include phycocyanins, allophycocyanins, and phycoerythrins, each containing one, two, or three bilins anchored to the protein via cysteine(s). The proposal that phycoyanins are related to globins originated with the structural work of Schirmer and colleagues. C-phycocyanins from the cyanobacteria *Mastigocladus laminosus* (Schirmer et al., 1985) and *Agmenellum quadruplicatum* (now *Synechococcus* sp. PCC 7002) (Schirmer et al., 1986) were found to adopt a globular shape composed of six α helices preceded by an additional pair of helices responsible for dimerisation of an α subunit with a β subunit (Fig. 7A). The structural similarity of globins and phycobiliproteins immediately raised the question of evolutionary relationship versus common solution to forming a globular three-dimensional structure capable of accommodating a large hydrophobic prosthetic group.

A detailed analysis of the structure of *M. laminosus* phycocyanin by Pastore & Lesk explored the issue of common ancestry (Pastore & Lesk, 1990). These authors noted that the short $3_{10}$ C-helix, a distinctive element of globin structure with no apparent functional role, was present in phycocyanin. The conservation of such an idiosyncratic feature could only be explained by an evolutionary relationship. Interestingly, Isogai transformed the α chain of *M. laminosus* phycocyanin into a haem binding protein with a design based on the myoglobin from the sea hare *Aplysia limacina*. Replacement of eight residues in the haem cavity (including the distal and

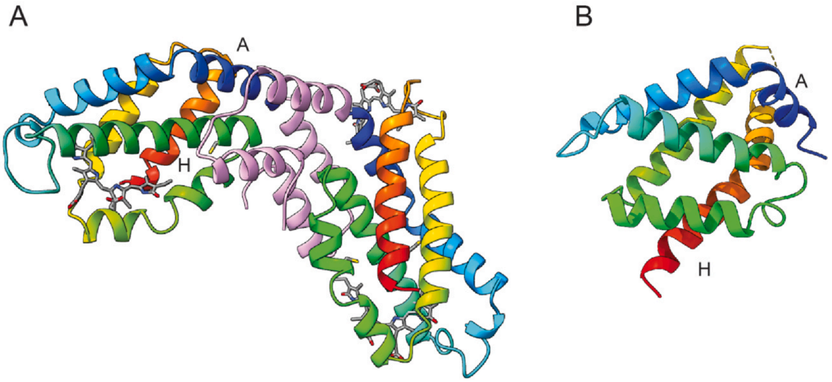


**Fig. 7** The structure of phycocyanin and RsbR. (A) C-Phycocyanin αβ dimer from *Spirulina platensis* (PDB ID 1HA7, chains A & B) with phycocyanobilin molecules in sticks. (B) RsbR (PDB ID 2BNL, chain A). The globin domains are in rainbow colours. The A and H helices are indicated. The additional pairs of helices securing the phycocyanin assembly are in pink.

proximal histidines) was sufficient to confer the ability to bind haem (Isogai & Ishida, 2009). Of note is that the transformation did not produce a reliable oxygen transport protein. Phycobiliproteins, like haemoglobins, excel at controlling the properties of their attached prosthetic group (Gisriel et al., 2023). In the case of phycobiliproteins, evolution allowed for the selection of specific linear tetrapyrroles and the tuning of absorption wavelength.

There is yet another subdivision of the phycocyanin family, called photoglobins (Schneider et al., 2022). This group was recognised as such with large scale genomic analyses. No three-dimensional structure is available yet, although homology predictions are sufficiently robust to describe photoglobins as adopting the 3/3 globin fold. The architecture of multidomain photoglobins vary. Most are associated with a B12-binding domain, module typically involved in light sensing or methyl group transfer. Photoglobins have the properties required for binding bilins and may represent an early form of light sensor. Unlike phycobiliproteins, they are rarely found in cyanobacteria.

Recently, Lagarias and co-workers produced a comprehensive analysis of phycobiliproteins (Rockwell et al., 2023) and clarified the relationship of photoglobins and phycocyanins. The last step in phycocyanobilin biosynthesis is performed by the ferredoxin-dependent bilin reductase, PcyA. PcyAs descended from Pre-PcyAs, which are associated with proteins having homology to phycobiliproteins and globins. The globin part of the tree identifies bilin biosynthesis-associated globins (BBAGs), with the conclusion that cyanobacterial BBAGs are phylogenetically distinct from light-harvesting phycobiliproteins. BBAGs cluster in three clades. They all bind bilin but do not seem to be efficient at light harvesting. The BBAG-2 lineage includes photoglobins.

Interestingly, the phytozome entry of *C. reinhardtii* CC-4532 lists all twelve THBs, and also a globin domain in glutathione S-transferase (Cre01.g044700_4532.2). On inspection, this C-terminal domain has an AphaFold2 predicted fold closely matching that of a rhodophytan phycoerythrin (Ritter, Hiller, Wrench, Welte, & Diederichs, 1999).

## 4. The RsbR-like proteins

The founding member of the RsbR group is the N-terminal domain of an environmental stress-response regulator, N-RsbR, from *Bacillus subtilis* (Murray et al., 2005). This particular stressosome domain has the overall topology of a globin except for the relative displacement of structural elements such that the haem binding cavity does not form. A few years after the observation, the same fold was noted in two sensor proteins from *Bacillus anthracis*, sporulation inhibitors pXO1–118 and pXO2–61 (Stranzl et al., 2011). Sporulation inhibitor pXO1–118 crystallised with a fatty acid molecule bound in a long hydrophobic tunnel. A coevolutionary question arises again, that is, whether the RsbR proteins are a manifestation of the topological ease with which helices can fold into a globular shape or they have an evolutionary relationship with the globin superfamily. A recent hypothesis arising from current genomic data organises RsbR-like proteins in several distinct clusters distantly related to haem-binding globins (Schneider et al., 2022). As an aside, *Vibrio* RsbR proteins have an N-terminal domain that binds a haem group (Jia, Wang, Rivera, Duong, & Weinert, 2016). These domains sense dioxygen and are related to the aerotactic protein HemAT (a GCS) rather than non-haem binding proteins. Hereafter, RsbR refers to the *B. subtilis* type.

A BLAST search using the globin domain of *B. subtilis* RsbR (PDB ID 2BNL) as the query found approximately thirty instances of related proteins in cyanobacteria where the N-terminal globin domain is followed by a sulphate transporter and anti-sigma factor antagonist (STAS) domain. Multiple sequence alignment, comparison with the structures of *B. subtilis* and *B. anthracis* RsbRs, and use of Alphafold2 models support the presence of an arginine where the proximal histidine resides. Position B10 is occupied by a tryptophan. Also conspicuous is a conserved cysteine in the E helix. These cyanobacterial relatives are unlikely to bind haem, and their role in the cell remains a mystery.

## 5. Phylogeny of the five-branch globin superfamily

The constant expansion of genomic data has led to several revisions of the superfamily history. The latest phylogenetic assessment of haem-binding globins (Schuster et al., 2024) posits that the superfamily originated with the 3/3 fold and places M, S, and T proteins in the pre-last universal common ancestor (Schuster et al., 2024). The S and M lineages split first from the ancestral globin, and the T family split later from the M family. The nature of the most ancestral haemoprotein is not known, but the presence of thermoglobin in hyperthermophiles and thermoglobin-like proteins in modern cyanobacteria may favour an M-type origin (Miranda et al., 2005). The S branch has been less well understood (Pesce et al., 2013; Schuster et al., 2024; Vinogradov et al., 2007) in part because of an insufficient number of sequences displaying a relationship to globin-coupled sensors. Several groups of S proteins are recognised according to their multidomain architecture and composition (Schuster et al., 2024; Vávra et al., 2022) whereas the single domain protoglobins and sensor single domain globins cluster in two different clades. Ramification was also adjusted for the T family, with the early split of the Q subfamily, followed by the split of the P subfamily, then O and N, N being paraphyletic. Early branching from the ancestral M/S proteins likely generated bilin-binding proteins, phycocyanins and photoglobins. It was initially proposed that RsbR was a recycled globin sensor domain and surmised that this was the first instance of what would constitute yet another branch of the globin superfamily (Murray et al., 2005; Vinogradov, Tinajero-Trejo, et al., 2013). Rather than a recycled globin domain, RsbR-like proteins may be ancestral (Schneider et al., 2022). In this scenario, these proteins occurred first to bind to other proteins, evolved to bind fatty acids, then linear tetrapyrroles, and further to bind haem (Fig. 8).

Cyanobacteria have globins belonging to all five branches of the superfamily. Parsing further, TrHb3s, and TrHb4s are absent as are photoglobins except for one instance (Schneider et al., 2022). Phycobiliproteins are of course present and essential for cyanobacterial photosynthesis.

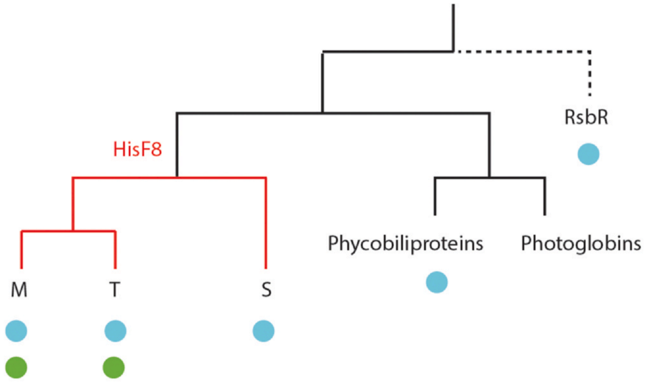


**Fig. 8** Simplified globin superfamily tree. The appearance of HisF8 marks the ability to bind haem. Some S proteins do not have this residue. Cyan and green dots indicate presence in cyanobacteria and chlorophytes, respectively. RsbR is the outgroup, possibly ancestral to the superfamily. *Adapted from (Schuster et al., 2024) and (Rockwell et al., 2023).*

Chlorophytes harbour mostly TrHb1s and a small number of Fgb-like proteins, and perhaps the phycocyanin-like domain mentioned in Section 3. These are likely derived from endosymbiosis with a cyanobacterium (Vinogradov, Tinajero-Trejo, et al., 2013).

As mentioned in Section 3, cyanobacteria formed the first, core group of organisms capable of oxygenic photosynthesis. They were instrumental instigators of the Great Oxidation Event that altered the earth's biosphere from the early anoxic environment favouring abundant soluble reduced iron, to the modern aerobic and oxidised earth (Kořený, Oborník, Horáková, Waller, & Lukeš, 2022; Schirrmeister, Gugger, & Donoghue, 2015). The diversity of the globins from cyanobacteria and their endosymbiotic relationship to chlorophytes illustrates deep evolutionary ties within the superfamily and may help trace how fundamental chemistries adapt to changing environments.

## 6. Applications

The early origin of the globin superfamily and the functional diversity of its members make for a challenging phylogenetic organisation, and with a distinct lack of physiological information on functional diversity, annotation is frequently inaccurate. With the rise of deep learning approaches (Al Fahoum, Zyout, Alquran, & Abu-Qasmieh, 2023; Bileschi et al., 2022), it is hoped that the pattern recognition required for classification and annotation can be enhanced and facilitate curation of the growing number of sequences in the databases. In fact, the ancient lineage and diverse chemistries of the haem-binding globins and the entire globin superfamily can be considered as a stringent test case for artificial intelligence applications to bioinformatics.

In the cyanobacterial Section 2.1.1.1, mention was made of using *N. commune* GlbN to boost protection from dioxygen in photosynthetic organisms and *Synechocystis* sp. PCC 6803 GlbN as a route to the production of haem. TrHbs can also serve as templates for exploring how to improve the properties of dioxygen binders. For example, the haem–protein crosslink of *Synechocystis* GlbN was engineered in the globin domain of *C. eugametos* LI637 in a demonstration of a simple approach to stabilising the protein and preventing the loss of toxic haem (Rice, Preimesberger, Johnson, & Lecomte, 2014). Other applications, for example use of GlbN-like proteins as coloured tags facilitating expression and purification (Khan, Uppal, & Kundu, 2021), are conceivable. In a recent study, directed evolution of the *A. pernix* protoglobin was used to develop catalysts for the formation of drug precursors (Schaus et al., 2023). Other applications are based on microbial globins, for example the production of leghaemoglobin as a "food colouring" agent, the production of the triterpenoid botulinic acid in *Saccharomyces cerevisiae* with the aid of *Vitreoscilla* haemoglobin (Zhao, Zhou, Du, & Chen, 2021), and the synthesis of limonene in *E. coli* with the help of the β-proteobacterium *Sphaerotilus natans* truncated haemoglobins (Li et al., 2024).

Several species of Chlorophytes, including the model organism *C. reinhardtii*, are being explored as platforms for synthetic biology and commercial application (Goold, Moseley, & Lauersen, 2024), owing to their potential for autotrophic growth, use of readily available nutrient streams, and a diversity of possible metabolic chemistries. Among the applications under consideration is biological production of hydrogen as a fuel source (Hemschemeier, Posewitz, & Happe, 2023) driven by the [FeFe]-hydrogenase enzymes found in *C. reinhardtii*. Several of the TrHbs of *C. reinhardtii* have been found to be linked to hypoxic growth and NO signalling, with both pathways link to hydrogen production (Hemschemeier et al., 2013). Given the exquisite versatility of haem chemistry and the plasticity of the globin fold, globins of oxygenic photosynthetic organisms have potential for further biotechnological innovations.

## 7. Concluding remarks

The ancient origins, broad functional diversity, and a century of mammalian physiologic research make the globin superfamily an excellent system to investigate, and globins of cyanobacteria and chlorophytes contribute valuable information about the range of its possible chemistries. The rise of powerful genomic sequencing is outpacing our investigations into the physiologic chemical balance cells use for survival, and the study of these globins can bridge the divide between computational predictions to physiological evidence. It is then sobering that, since Vinogradov and coworkers published their revised perspective on bacterial globin (Vinogradov, Tinajero-Trejo, et al., 2013), the lack of experimental evidence pointed out a decade ago is still problematic today.

The DNA-encoded information still provides vital studies into globin function. Advances in molecular biology have made heterologous protein expression routine, simplifying in vitro studies of chemical and physical properties. Computational approaches have also progressed to the point that useful approximations of three-dimensional structures of these hypothetical proteins can be retrieved from libraries or calculated from sequence in minutes of CPU time. Other computational tools are being developed to predict physical properties such as oxygen binding kinetics with knowledge of the primary structure and existing structural models (Bustamante et al., 2016). These avenues are promising and, once broadly validated, will improve our understanding of the relationship between structure and function within the superfamily.

Meanwhile, the "gold standard" remains the exploration of globin in-vivo function by mutagenesis and phenotypic characterisation (Vinogradov, Tinajero-Trejo, et al., 2013). There is a broadening gap between the growing number of available DNA-derived amino acid sequences and the number of investigations into the physiologically relevant properties of the hypothetical proteins encoded within these genes. Only foundational physiologic studies, performed on judiciously chosen model organisms spanning different metabolic needs and living conditions, will be able to define confidently the necessity of globin chemistry in photosynthetic organisms.

## Acknowledgement

The authors acknowledge the financial support provided by the National Science Foundation, grant CHE-2003950.

## References

Aboouf, M. A., Gorr, T. A., Hamdy, N. M., Gassmann, M., & Thiersch, M. (2023). Myoglobin in brown adipose tissue: A multifaceted player in thermogenesis. *Cells, 12*(18), https://doi.org/10.3390/cells12182240.

Afanasyev, P., Seer-Linnemayr, C., Ravelli, R. B. G., Matadeen, R., De Carlo, S., Alewijnse, B., ... van Heel, M. (2017). Single-particle cryo-EM using alignment by classification (ABC): The structure of *Lumbricus terrestris* haemoglobin. *IUCrJ, 4*(5), 678–694. https://doi.org/10.1107/S2052252517010922.

Al Fahoum, A., Zyout, A., Alquran, H., & Abu-Qasmieh, I. (2023). A novel multi-stage bispectral deep learning method for protein family classification. *Computers, Materials & Continua, 76*(1), 1173–1193. https://doi.org/10.32604/cmc.2023.038304.

Altschul, S. F., Gish, W., Miller, W., Myers, E. W., & Lipman, D. J. (1990). Basic local alignment search tool. *Journal of Molecular Biology, 215*(3), 403–410. https://doi.org/10.1016/S0022-2836(05)80360-2.

Avilan, L., Roumezi, B., Risoul, V., Bernard, C. S., Kpebe, A., Belhadjhassine, M., ... Latifi, A. (2018). Phototrophic hydrogen production from a clostridial [FeFe] hydrogenase expressed in the heterocysts of the cyanobacterium *Nostoc* PCC 7120. *Applied Microbiology and Biotechnology, 102*(13), 5775–5783. https://doi.org/10.1007/s00253-018-8989-2.

Bateman, A., Murzin, A. G., & Teichmann, S. A. (1998). Structure and distribution of pentapeptide repeats in bacteria. *Protein Science,* 7(6), 1477–1480. https://doi.org/10.1002/pro.5560070625.

Battistuzzi, M., Morlino, M. S., Cocola, L., Trainotti, L., Treu, L., Campanaro, S., ... La Rocca, N. (2024). Transcriptomic and photosynthetic analyses of *Synechocystis* sp. PCC 6803 and *Chlorogloeopsis fritschii* sp. PCC 6912 exposed to an M-dwarf spectrum under an anoxic atmosphere. *Frontiers in Plant Science, 14*, 1322052. https://doi.org/10.3389/fpls.2023.1322052.

Becana, M., Yruela, I., Sarath, G., Catalán, P., & Hargrove, M. S. (2020). Plant hemoglobins: A journey from unicellular green algae to vascular plants. *New Phytologist, 227*(6), 1618–1635. https://doi.org/10.1111/nph.16444.

Bellido-Pedraza, C. M., Calatrava, V., Sanz-Luque, E., Tejada-Jiménez, M., Llamas, Á., Plouviez, M., ... Galván, A. (2020). *Chlamydomonas reinhardtii*, an algal model in the nitrogen cycle. *Plants, 9*(7), https://doi.org/10.3390/plants9070903 Article 7.

Bianco, C. L., Toscano, J. P., Bartberger, M. D., & Fukuto, J. M. (2017). The chemical biology of HNO signaling. *Archives of Biochemistry and Biophysics, 617*, 129–136. https://doi.org/10.1016/j.abb.2016.08.014.

Bileschi, M. L., Belanger, D., Bryant, D. H., Sanderson, T., Carter, B., Sculley, D., ... Colwell, L. J. (2022). Using deep learning to annotate the protein universe. *Nature Biotechnology, 40*(6), 932–937. https://doi.org/10.1038/s41587-021-01179-w.

Blank, M., Wollberg, J., Gerlach, F., Reimann, K., Roesner, A., Hankeln, T., ... Burmester, T. (2011). A membrane-bound vertebrate globin. *PLoS One, 6*(9), e25292. https://doi.org/10.1371/journal.pone.0025292.

Bolay, P., Muro-Pastor, M. I., Florencio, F. J., & Klähn, S. (2018). The distinctive regulation of cyanobacterial glutamine synthetase. *Life (Chicago, Ill.: 1978), 8*(4), https://doi.org/10.3390/life8040052.

Brejc, K., Ficner, R., Huber, R., & Steinbacher, S. (1995). Isolation, crystallization, crystal structure analysis and refinement of allophycocyanin from the cyanobacterium *Spirulina platensis* at 2.3 Å resolution. *Journal of Molecular Biology, 249*(2), 424–440. https://doi.org/10.1006/jmbi.1995.0307.

Bryant, D. A., & Gisriel, C. J. (2024). The structural basis for light harvesting in organisms producing phycobiliproteins. *The Plant Cell*. https://doi.org/10.1093/plcell/koae126 (in press.).

Burmester, T., Weich, B., Reinhardt, S., & Hankeln, T. (2000). A vertebrate globin expressed in the brain. *Nature, 407*(6803), 520–523. https://doi.org/10.1038/35035093.

Büsch, A., Friedrich, B., & Cramm, R. (2002). Characterization of the *norB* gene, encoding nitric oxide reductase, in the nondenitrifying cyanobacterium *Synechocystis* sp. strain PCC6803. *Applied and Environmental Microbiology, 68*(2), 668–672. https://doi.org/10.1128/AEM.68.2.668-672.2002.

Bustamante, J. P., Radusky, L., Boechi, L., Estrin, D. A., ten Have, A., & Martí, M. A. (2016). Evolutionary and functional relationships in the truncated hemoglobin family. *PLoS Computational Biology, 12*(1), 26. https://doi.org/10.1371/journal.pcbi.1004701.

Calatrava, V., Tejada-Jimenez, M., Sanz-Luque, E., Fernandez, E., & Galvan, A. (2023). Nitrogen metabolism in *Chlamydomonas*. In A. R. Grossman, & F.-A. Wollman (Eds.). *The chlamydomonas sourcebook* (pp. 99–128)(third ed.). Academic Press. https://doi.org/10.1016/B978-0-12-821430-5.00004-3.

Cao, K., Wang, X., Sun, F., Zhang, H., Cui, Y., Cao, Y., … Gao, Z. (2023). Promoting heme and phycocyanin biosynthesis in *Synechocystis* sp. PCC 6803 by overexpression of porphyrin pathway genes with genetic engineering. *Marine Drugs, 21*(7), https://doi.org/10.3390/md21070403.

Carabet, L. A., Guertin, M., Lagüe, P., & Lamoureux, G. (2017). Mechanism of the nitric oxide dioxygenase reaction of *Mycobacterium tuberculosis* hemoglobin N. *Journal of Physical Chemistry B, 121*(37), 8706–8718. https://doi.org/10.1021/acs.jpcb.7b06494.

Ciaccio, C., Ocaña-Calahorro, F., Droghetti, E., Tundo, G. R., Sanz-Luque, E., Polticelli, F., … Coletta, M. (2015). Functional and spectroscopic characterization of *Chlamydomonas reinhardtii* truncated hemoglobins. *PLoS One, 10*(5), e0125005. https://doi.org/10.1371/journal.pone.0125005.

Couture, M., Chamberland, H., St-Pierre, B., Lafontaine, J., & Guertin, M. (1994). Nuclear genes encoding chloroplast hemoglobins in the unicellular green alga *Chlamydomonas eugametos*. *Molecular Genetics and Genomics, 243*(2), 185–197. https://doi.org/10.1007/BF00280316.

Couture, M., Das, T. K., Savard, P. Y., Ouellet, Y., Wittenberg, J. B., Wittenberg, B. A., … Guertin, M. (2000). Structural investigations of the hemoglobin of the cyanobacterium *Synechocystis* PCC 6803 reveal a unique distal heme pocket. *European Journal of Biochemistry, 267*(15), 4770–4780. https://doi.org/10.1046/j.1432-1327.2000.01531.x.

Crooks, G. E., Hon, G., Chandonia, J. M., & Brenner, S. E. (2004). WebLogo: A sequence logo generator. *Genome Research, 14*(6), 1188–1190. https://doi.org/10.1101/gr.849004.

Dailey, H. A., Dailey, T. A., Gerdes, S., Jahn, D., Jahn, M., O'Brian, M. R., & Warren, M. J. (2017). Prokaryotic heme biosynthesis: Multiple pathways to a common essential product. *Microbiology and Molecular Biology Reviews, 81*(1), https://doi.org/10.1128/mmbr.00048-16.

Das, T. K., Couture, M., Lee, H. C., Peisach, J., Rousseau, D. L., Wittenberg, B. A., … Guertin, M. (1999). Identification of the ligands to the ferric heme of *Chlamydomonas* chloroplast hemoglobin: Evidence for ligation of tyrosine-63 (B10) to the heme. *Biochemistry, 38*(46), 15360–15368. https://doi.org/10.1021/bi991237e.

Drazic, A., Myklebust, L. M., Ree, R., & Arnesen, T. (2016). The world of protein acetylation. *Biochimica et Biophysica Acta (BBA) – Proteins and Proteomics, 1864*(10), 1372–1401. https://doi.org/10.1016/j.bbapap.2016.06.007.

Dunwell, J. M., Purvis, A., & Khuri, S. (2004). Cupins: The most functionally diverse protein superfamily? *Phytochemistry, 65*(1), 7–17. https://doi.org/10.1016/j.phytochem.2003.08.016.

Dupuis, S., & Merchant, S. S. (2023). *Chlamydomonas reinhardtii*: A model for photosynthesis and so much more. *Nature Methods, 20*(10), 1441–1442. https://doi.org/10.1038/s41592-023-02023-6.

El Ghachi, M., Howe, N., Huang, C.-Y., Olieric, V., Warshamanage, R., Touzé, T., ... Caffrey, M. (2018). Crystal structure of undecaprenyl-pyrophosphate phosphatase and its role in peptidoglycan biosynthesis. *Nature Communications, 9*(1), 1078. https://doi.org/10.1038/s41467-018-03477-5.

Ertas, B., Kiger, L., Blank, M., Marden, M. C., & Burmester, T. (2011). A membrane-bound hemoglobin from gills of the green shore crab *Carcinus maenas*. *Journal of Biological Chemistry, 286*(5), 3185–3193. https://doi.org/10.1074/jbc.M110.160341.

Estarellas, C., Capece, L., Seira, C., Bidon-Chanal, A., Estrin, D. A., & Luque, F. J. (2016). Structural plasticity in globins: Role of protein dynamics in defining ligand migration pathways. In C. Z. Christov (Vol. Ed.), *Advances in protein chemistry and structural biology: Vol. 105*, (pp. 59–80). Academic Press. https://doi.org/10.1016/bs.apcsb.2016.07.002.

Falzone, C. J., Vu, B. C., Scott, N. L., & Lecomte, J. T. J. (2002). The solution structure of the recombinant hemoglobin from the cyanobacterium *Synechocystis* sp. PCC 6803 in its hemichrome state. *Journal of Molecular Biology, 324*, 1015–1029. https://doi.org/10.1016/S0022-2836(02)01093-8.

Filina, V., Grinko, A., & Ermilova, E. (2019). Truncated hemoglobins 1 and 2 are implicated in the modulation of phosphorus deficiency-induced nitric oxide levels in *Chlamydomonas*. *Cells, 8*(9), 947. https://doi.org/10.3390/cells8090947.

Fischer, B. B., Ledford, H. K., Wakao, S., Huang, S. G., Casero, D., Pellegrini, M., ... Niyogi, K. K. (2012). SINGLET OXYGEN RESISTANT 1 links reactive electrophile signaling to singlet oxygen acclimation in *Chlamydomonas reinhardtii*. *Proceedings of the National Academy of Sciences of the United States of America, 109*, E1302–E1311. https://doi.org/10.1073/pnas.1116843109.

Flores, J. F., Fisher, C. R., Carney, S. L., Green, B. N., Freytag, J. K., Schaeffer, S. W., & Royer, W. E. (2005). Sulfide binding is mediated by zinc ions discovered in the crystal structure of a hydrothermal vent tubeworm hemoglobin. *Proceedings of the National Academy of Sciences of the United States of America, 102*(8), 2713–2718. https://doi.org/10.1073/pnas.0407455102.

Freitas, T. A. K., Hou, S., Dioum, E. M., Saito, J. A., Newhouse, J., Gonzalez, G., ... Alam, M. (2004). Ancestral hemoglobins in *Archaea*. *Proceedings of the National Academy of Sciences of the United States of America, 101*(17), 6675–6680. https://doi.org/10.1073/pnas.0308657101.

Freitas, T. A. K., Saito, J. A., Hou, S., & Alam, M. (2005). Globin-coupled sensors, protoglobins, and the last universal common ancestor. *Journal of Inorganic Biochemistry, 99*(1), 23–33. https://doi.org/10.1016/j.jinorgbio.2004.10.024.

Gisriel, C. J., Elias, E., Shen, G., Soulier, N. T., Brudvig, G. W., Croce, R., & Bryant, D. A. (2023). Structural comparison of allophycocyanin variants reveals the molecular basis for their spectral differences. *Photosynthesis Research*. https://doi.org/10.1007/s11120-023-01048-4.

Goold, H. D., Moseley, J. L., & Lauersen, K. J. (2024). The synthetic future of algal genomes. *Cell Genomics, 4*(3), https://doi.org/10.1016/j.xgen.2024.100505.

Grinko, A., Alqoubaili, R., Lapina, T., & Ermilova, E. (2021). Truncated hemoglobin 2 modulates phosphorus deficiency response by controlling of gene expression in nitric oxide-dependent pathway in *Chlamydomonas reinhardtii*. *Planta, 254*(2), 39. https://doi.org/10.1007/s00425-021-03691-4.

Hade, M. D., Kaur, J., Chakraborti, P. K., & Dikshit, K. L. (2017). Multidomain truncated hemoglobins: New members of the globin family exhibiting tandem repeats of globin units and domain fusion. *IUBMB Life, 69*(7), 479–488. https://doi.org/10.1002/iub.1630.

Hafideddine, Z., Loier, T., Brempt, N. V., Henau, S. D., Ching, H. Y. V., Neukermans, S., ... Doorslaer, S. V. (2023). GLB-3: A resilient, cysteine-rich, membrane-tethered globin expressed in the reproductive and nervous system of *Caenorhabditis elegans*. *Journal of Inorganic Biochemistry, 238*, 112063. https://doi.org/10.1016/j.jinorgbio.2022.112063.

Hemschemeier, A., Duner, M., Casero, D., Merchant, S. S., Winkler, M., & Happe, T. (2013). Hypoxic survival requires a 2-on-2 hemoglobin in a process involving nitric oxide. *Proceedings of the National Academy of Sciences of the United States of America, 110*(26), 10854–10859. https://doi.org/10.1073/pnas.1302592110.

Hemschemeier, A., Posewitz, M. C., & Happe, T. (2023). Hydrogenases and hydrogen production. In A. R. Grossman, & F.-A. Wollman (Eds.). *The chlamydomonas sourcebook* (pp. 343–367)(third ed.). Academic Press. https://doi.org/10.1016/B978-0-12-821430-5.00008-0.

Hill, D. R., Belbin, T. J., Thorsteinsson, M. V., Bassam, D., Brass, S., Ernst, A., … Potts, M. (1996). GlbN (cyanoglobin) is a peripheral membrane protein that is restricted to certain *Nostoc* spp. *Journal of Bacteriology, 178*(22), 6587–6598. https://doi.org/10.1128/jb.178.22.6587-6598.1996.

Holm, L., & Rosenström, P. (2010). Dali server: Conservation mapping in 3D (PMC) *Nucleic Acids Research, 38*(Web Server issue), W545–W549. https://doi.org/10.1093/nar/gkq366.

Hong, G., & Pachter, R. (2012). Inhibition of biocatalysis in [Fe–Fe] hydrogenase by oxygen: Molecular dynamics and density functional theory calculations. *ACS Chemical Biology, 7*(7), 1268–1275. https://doi.org/10.1021/cb3001149.

Hoogewijs, D., Ebner, B., Germani, F., Hoffmann, F. G., Fabrizius, A., Moens, L., … Hankeln, T. (2011). Androglobin: A chimeric globin in metazoans that is preferentially expressed in mammalian testes. *Molecular Biology and Evolution, 29*(4), 1105–1114. https://doi.org/10.1093/molbev/msr246.

Hoy, J. A., Kundu, S., Trent, J. T., 3rd, Ramaswamy, S., & Hargrove, M. S. (2004). The crystal structure of *Synechocystis* hemoglobin with a covalent heme linkage. *Journal of Biological Chemistry, 279*(16), 16535–16542. https://doi.org/10.1074/jbc.M313707200.

Huwald, D., Duda, S., Gasper, R., Olieric, V., Hofmann, E., & Hemschemeier, A. (2020). Distinctive structural properties of THB11, a pentacoordinate *Chlamydomonas reinhardtii* truncated hemoglobin with N- and C-terminal extensions. *Journal of Biological Inorganic Chemistry, 25*(2), 267–283. https://doi.org/10.1007/s00775-020-01759-2.

Huwald, D., Schrapers, P., Kositzki, R., Haumann, M., & Hemschemeier, A. (2015). Characterization of unusual truncated hemoglobins of *Chlamydomonas reinhardtii* suggests specialized functions. *Planta, 242*(1), 167–185. https://doi.org/10.1007/s00425-015-2294-4.

Ilari, A., Bonamore, A., Farina, A., Johnson, K. A., & Boffi, A. (2002). The X-ray structure of ferric *Escherichia coli* flavohemoglobin reveals an unexpected geometry of the distal heme pocket. *Journal of Biological Chemistry, 277*(26), 23725–23732. https://doi.org/10.1074/jbc.M202228200.

Isogai, Y., & Ishida, M. (2009). Design of a novel heme protein with a non-heme globin scaffold. *Biochemistry, 48*(34), 8136–8142. https://doi.org/10.1021/bi900518q.

Jia, X., Wang, J., Rivera, S., Duong, D., & Weinert, E. E. (2016). An $O_2$-sensing stressosome from a Gram-negative bacterium. *Nature Communications, 7*(1), 12381. https://doi.org/10.1038/ncomms12381.

Johnson, E. A., & Lecomte, J. T. J. (2013). The globins of cyanobacteria and algae. *Advances in Microbial Physiology, 63*, 195–272. https://doi.org/10.1016/B978-0-12-407693-8.00006-6.

Johnson, E. A., & Lecomte, J. T. J. (2014). Characterization of the truncated hemoglobin THB1 from protein extracts of *Chlamydomonas reinhardtii. F1000Research, 3*, 294. https://doi.org/10.12688/f1000research.5873.1.

Johnson, E. A., & Lecomte, J. T. J. (2015). The haemoglobins of algae. *Advances in Microbial Physiology, 67*, 177–234. https://doi.org/10.1016/bs.ampbs.2015.08.003.

Johnson, E. A., Rice, S. L., Preimesberger, M. R., Nye, D. B., Gilevicius, L., Wenke, B. B., … Lecomte, J. T. J. (2014). Characterization of THB1, a *Chlamydomonas reinhardtii* truncated hemoglobin: Linkage to nitrogen metabolism and identification of lysine as the distal heme ligand. *Biochemistry, 53*(28), 4573–4589. https://doi.org/10.1021/bi5005206.

Johnson, E. A., Russo, M. M., Nye, D. B., Schlessman, J. L., & Lecomte, J. T. J. (2018). Lysine as a heme iron ligand: A property common to three truncated hemoglobins from *Chlamydomonas reinhardtii*. *Biochimica et Biophysica Acta (BBA) – General Subjects, 1862*(12), 2660–2673. https://doi.org/10.1016/j.bbagen.2018.08.009.

Julió Plana, L., Grundman, J. E. M., Estrin, D. A., Lecomte, J. T. J., & Capece, L. (2021). Distal lysine (de)coordination in the algal hemoglobin THB1: A combined computer simulation and experimental study. *Journal of Inorganic Biochemistry, 220*, 111455. https://doi.org/10.1016/j.jinorgbio.2021.111455.

Jumper, J., Evans, R., Pritzel, A., Green, T., Figurnov, M., Ronneberger, O., … Adler, J., Hassabis, D. (2021). Highly accurate protein structure prediction with AlphaFold. *Nature, 596*(7873), https://doi.org/10.1038/s41586-021-03819-2 Article 7873.

Kakar, S., Hoffman, F. G., Storz, J. F., Fabian, M., & Hargrove, M. S. (2010). Structure and reactivity of hexacoordinate hemoglobins. *Biophysical Chemistry, 152*(1–3), 1–14. https://doi.org/10.1016/j.bpc.2010.08.008.

Kawada, N., Kristensen, D. B., Asahina, K., Nakatani, K., Minamiyama, Y., Seki, S., & Yoshizato, K. (2001). Characterization of a stellate cell activation-associated protein (STAP) with peroxidase activity found in rat hepatic stellate cells. *Journal of Biological Chemistry, 276*(27), 25318–25323. https://doi.org/10.1074/jbc.M102630200.

Keppner, A., Maric, D., Correia, M., Koay, T. W., Orlando, I. M. C., Vinogradov, S. N., & Hoogewijs, D. (2020). Lessons from the post-genomic era: Globin diversity beyond oxygen binding and transport. *Redox Biology, 37*, 101687. https://doi.org/10.1016/j.redox.2020.101687.

Khan, M. A., Uppal, S., & Kundu, S. (2021). *Novel hemoglobin from Synechocystis sp. PCC 6803: Shedding light on the structure-function relationship and its biotechnological applications. Cyanobacteria-recent advances in taxonomy and applications.* IntechOpen. https://doi.org/10.5772/intechopen.97699.

Koch, J., & Burmester, T. (2016). Membrane-bound globin X protects the cell from reactive oxygen species. *Biochemical and Biophysical Research Communications, 469*(2), 275–280. https://doi.org/10.1016/j.bbrc.2015.11.105.

Kořený, L., Oborník, M., Horáková, E., Waller, R. F., & Lukeš, J. (2022). The convoluted history of haem biosynthesis. *Biological Reviews, 97*(1), 141–162. https://doi.org/10.1111/brv.12794.

Lechtreck, K.-F., Johnson, E. C., Sakai, T., Cochran, D., Ballif, B. A., Rush, J., … Witman, G. B. (2009). The *Chlamydomonas reinhardtii* BBSome is an IFT cargo required for export of specific signaling proteins from flagella. *Journal of Cell Biology, 187*(7), 1117–1132. https://doi.org/10.1083/jcb.200909183.

León, J., & Costa-Broseta, Á. (2020). Present knowledge and controversies, deficiencies, and misconceptions on nitric oxide synthesis, sensing, and signaling in plants. *Plant, Cell & Environment, 43*(1), 1–15. https://doi.org/10.1111/pce.13617.

Li, M., Yang, Z., Chen, S., Liu, Z., Tong, L., Zheng, S., & Yang, D. (2024). *Sphaerotilus natans* hemoglobins have an NADH oxidation activity and promote the yield of limonene in an engineered *E. coli* strain. *International Journal of Biological Macromolecules, 254*, 128112. https://doi.org/10.1016/j.ijbiomac.2023.128112.

Marković, D., Pröll, S., Bubenzer, C., & Scheer, H. (2007). Myoglobin with chlorophyllous chromophores: Influence on protein stability. *Biochimica et Biophysica Acta (BBA) – Bioenergetics, 1767*, 897–904. https://doi.org/10.1016/j.bbabio.2007.03.011.

Martinez Grundman, J. E., Johnson, E. A., & Lecomte, J. T. J. (2023). Architectural digest: Thermodynamic stability and domain structure of a consensus monomeric globin. *Biophysical Journal, 122*(15), 3117–3132. https://doi.org/10.1016/j.bpj.2023.06.016.

Martinez Grundman, J. E., Julió Plana, L., Schlessman, J. L., Capece, L., Estrin, D., & Lecomte, J. T. J. (2021). Control of distal lysine coordination in a monomeric hemoglobin: A role for heme peripheral interactions. *Journal of Inorganic Biochemistry*, 111437. https://doi.org/10.1016/j.jinorgbio.2021.111437.

Mathur, S., Yadav, S. K., Yadav, K., Bhatt, S., & Kundu, S. (2023). A novel single sensor hemoglobin domain from the thermophilic cyanobacteria *Thermosynechococcus elongatus* BP-1 exhibits higher pH but lower thermal stability compared to globins from mesophilic organisms. *International Journal of Biological Macromolecules, 240*, 124471. https://doi.org/10.1016/j.ijbiomac.2023.124471.

McClure, R. S., Overall, C. C., McDermott, J. E., Hill, E. A., Markillie, L. M., McCue, L. A., ... Beliaev, A. S. (2016). Network analysis of transcriptomics expands regulatory landscapes in *Synechococcus* sp. PCC 7002. *Nucleic Acids Research, 44*(18), 8810–8825. https://doi.org/10.1093/nar/gkw737.

Merchant, S. S., Prochnik, S. E., Vallon, O., Harris, E. H., Karpowicz, S. J., Witman, G. B., ... Grossman, A. R. (2007). The Chlamydomonas genome reveals the evolution of key animal and plant functions. *Science (New York, N. Y.), 318*(5848), 245–250. https://doi.org/10.1126/science.1143609.

Milani, M., Pesce, A., Nardini, M., Ouellet, H., Ouellet, Y., Dewilde, S., ... Bolognesi, M. (2005). Structural bases for heme binding and diatomic ligand recognition in truncated hemoglobins. *Journal of Inorganic Biochemistry, 99*(1), 97–109. https://doi.org/10.1016/j.jinorgbio.2004.10.035.

Milani, M., Pesce, A., Ouellet, Y., Dewilde, S., Friedman, J., Ascenzi, P., ... Bolognesi, M. (2004). Heme-ligand tunneling in group I truncated hemoglobins. *Journal of Biological Chemistry, 279*(20), 21520–21525. https://doi.org/10.1074/jbc.M401320200.

Milani, M., Savard, P. Y., Ouellet, H., Ascenzi, P., Guertin, M., & Bolognesi, M. (2003). A TyrCD1/TrpG8 hydrogen bond network and a TyrB10-TyrCD1 covalent link shape the heme distal site of *Mycobacterium tuberculosis* hemoglobin O. *Proceedings of the National Academy of Sciences of the United States of America, 100*(10), 5766–5771. https://doi.org/10.1073/pnas.1037676100.

Minaeva, E., Zalutskaya, Z., Filina, V., & Ermilova, E. (2017). Truncated hemoglobin 1 is a new player in *Chlamydomonas reinhardtii* acclimation to sulfur deprivation. *PLoS One, 12*(10), e0186851. https://doi.org/10.1371/journal.pone.0186851.

Miranda, J. L., Maillett, D. H., Soman, J., & Olson, J. S. (2005). Thermoglobin, oxygen-avid hemoglobin in a bacterial hyperthermophile. *Journal of Biological Chemistry, 280*(44), 36754–36761. https://doi.org/10.1074/jbc.M505918200.

Mühlhaus, T., Weiss, J., Hemme, D., Sommer, F., & Schroda, M. (2011). Quantitative shotgun proteomics using a uniform $^{15}$N-labeled standard to monitor proteome dynamics in time course experiments reveals new insights into the heat stress response of *Chlamydomonas reinhardtii*. *Molecular & Cellular Proteomics, 10*(9), https://doi.org/10.1074/mcp.M110.004739.

Muraki, N., Takeda, K., Nam, D., Muraki, M., & Aono, S. (2021). Structural characterization of Y29F mutant of thermoglobin from a hyperthermophilic cacterium *Aquifex aeolicus*. *Chemistry Letters, 50*(4), 603–606. https://doi.org/10.1246/cl.200879.

Murray, J. W., Delumeau, O., & Lewis, R. J. (2005). Structure of a nonheme globin in environmental stress signaling. *Proceedings of the National Academy of Sciences of the United States of America, 102*(48), 17320–17325. https://doi.org/10.1073/pnas.0506599102.

Nardini, M., Pesce, A., & Bolognesi, M. (2022). Truncated (2/2) hemoglobin: Unconventional structures and functional roles in vivo and in human pathogenesis. *Molecular Aspects of Medicine, 84*, 101049. https://doi.org/10.1016/j.mam.2021.101049.

Nardini, M., Pesce, A., Labarre, M., Richard, C., Bolli, A., Ascenzi, P., ... Bolognesi, M. (2006). Structural determinants in the group III truncated hemoglobin from *Campylobacter jejuni*. *Journal of Biological Chemistry, 281*(49), 37803–37812. https://doi.org/10.1074/jbc.M607254200.

Nardini, M., Pesce, A., Thijs, L., Saito, J. A., Dewilde, S., Alam, M., ... Bolognesi, M. (2008). Archaeal protoglobin structure indicates new ligand diffusion paths and modulation of haem-reactivity. *EMBO Reports, 9*(2), 157–163. https://doi.org/10.1038/sj.embor.7401153.

Newcombe, E. A., Delaforge, E., Hartmann-Petersen, R., Skriver, K., & Kragelund, B. B. (2022). How phosphorylation impacts intrinsically disordered proteins and their function. *Essays in Biochemistry, 66*(7), 901–913. https://doi.org/10.1042/EBC20220060.

Nye, D. B., & Lecomte, J. T. J. (2018). Replacement of the distal histidine reveals a noncanonical heme binding site in a 2-on-2 hemoglobin. *Biochemistry, 57*(40), 5785–5796. https://doi.org/10.1021/acs.biochem.8b00752.

Nye, D. B., Preimesberger, M. R., Majumdar, A., & Lecomte, J. T. J. (2018). Histidine–lysine axial ligand switching in a hemoglobin: A role for heme propionates. *Biochemistry, 57*, 631–644. https://doi.org/10.1021/acs.biochem.7b01155.

O'Neill, E. (2020). Mining natural product biosynthesis in eukaryotic algae. *Marine Drugs, 18*(2), https://doi.org/10.3390/md18020090.

Ouellet, H., Milani, M., LaBarre, M., Bolognesi, M., Couture, M., & Guertin, M. (2007). The roles of Tyr(CD1) and Trp(G8) in *Mycobacterium tuberculosis* truncated hemoglobin O in ligand binding and on the heme distal site architecture. *Biochemistry, 46*(41), 11440–11450. https://doi.org/10.1021/bi7010288.

Ouellet, H., Ranguelova, K., Labarre, M., Wittenberg, J. B., Wittenberg, B. A., Magliozzo, R. S., & Guertin, M. (2007). Reaction of *Mycobacterium tuberculosis* truncated hemoglobin O with hydrogen peroxide: Evidence for peroxidatic activity and formation of protein-based radicals. *Journal of Biological Chemistry, 282*(10), 7491–7503. https://doi.org/10.1074/jbc.M609155200.

Padyana, A. K., Bhat, V. B., Madyastha, K. M., Rajashankar, K. R., & Ramakumar, S. (2001). Crystal structure of a light-harvesting protein C-phycocyanin from *Spirulina platensis*. *Biochemical and Biophysical Research Communications, 282*(4), 893–898. https://doi.org/10.1006/bbrc.2001.4663.

Pastore, A., & Lesk, A. M. (1990). Comparison of the structures of globins and phycocyanins: Evidence for evolutionary relationship. *Proteins, 8*, 133–155. https://doi.org/10.1002/prot.340080204.

Perutz, M. F. (1979). Regulation of oxygen affinity of hemoglobin: Influence of structure of the globin on the heme iron. *Annual Review of Biochemistry, 48*, 327–386. https://doi.org/10.1146/annurev.bi.48.070179.001551.

Perutz, M. F., & Mathews, F. S. (1966). An X-ray study of azide methaemoglobin. *Journal of Molecular Biology, 21*(1), 199–202. https://doi.org/10.1016/0022-2836(66)90088-x.

Pesce, A., Bolognesi, M., & Nardini, M. (2013). Protoglobin: Structure and ligand-binding properties. In R. K. Poole (Vol. Ed.), *Microbial globins – Status and opportunities: Vol. 63*, (pp. 79–96). Academic Press. https://doi.org/10.1016/B978-0-12-407693-8.00003-0.

Pesce, A., Couture, M., Dewilde, S., Guertin, M., Yamauchi, K., Ascenzi, P., ... Bolognesi, M. (2000). A novel two-over-two a-helical sandwich fold is characteristic of the truncated hemoglobin family. *EMBO Journal, 19*(11), 2424–2434. https://doi.org/10.1093/emboj/19.11.2424.

Pesce, A., Thijs, L., Nardini, M., Desmet, F., Sisinni, L., Gourlay, L., ... Dewilde, S. (2009). HisE11 and HisF8 provide bis-histidyl heme hexa-coordination in the globin domain of *Geobacter sulfurreducens* globin-coupled sensor. *Journal of Molecular Biology, 386*(1), 246–260. https://doi.org/10.1016/j.jmb.2008.12.023.

Phillips, S. E. (1980). Structure and refinement of oxymyoglobin at 1.6 Å resolution. *Journal of Molecular Biology, 142*, 531–554. https://doi.org/10.1016/0022-2836(80)90262-4.

Plouviez, M., Wheeler, D., Shilton, A., Packer, M. A., McLenachan, P. A., Sanz-Luque, E., ... Guieysse, B. (2017). The biosynthesis of nitrous oxide in the green alga *Chlamydomonas reinhardtii*. *Plant Journal, 91*(1), 45–56. https://doi.org/10.1111/tpj.13544.

Poole, R. K. (2020). Flavohaemoglobin: The pre-eminent nitric oxide–detoxifying machine of microorganisms. *F1000Research, 9*. https://doi.org/10.12688/f1000research.20563.1.

Porter, N. J., Danelius, E., Gonen, T., & Arnold, F. H. (2022). Biocatalytic carbene transfer using diazirines. *Journal of the American Chemical Society, 144*(20), 8892–8896. https://doi.org/10.1021/jacs.2c02723.

Potts, M., Angeloni, S. V., Ebel, R. E., & Bassam, D. (1992). Myoglobin in a cyanobacterium. *Science (New York, N. Y.), 256*(5064), 1690–1691. https://doi.org/10.1126/science.256.5064.1690.

Preimesberger, M. R., Johnson, E. A., Nye, D. B., & Lecomte, J. T. J. (2017). Covalent attachment of the heme to *Synechococcus* hemoglobin alters its reactivity toward nitric oxide. *Journal of Inorganic Biochemistry, 177*, 171–182. https://doi.org/10.1016/j.jinorgbio.2017.09.018.

Reeder, B. J. (2017). Redox and peroxidase activities of the hemoglobin superfamily: Relevance to health and disease. *Antioxidants & Redox Signaling, 26*(14), 763–776. https://doi.org/10.1089/ars.2016.6803.

Reeder, B. J., & Hough, M. A. (2014). The structure of a class 3 nonsymbiotic plant haemoglobin from *Arabidopsis thaliana* reveals a novel N-terminal helical extension. *Acta Crystallographica D Biological Crystallography, 70*(Pt 5), 1411–1418. https://doi.org/10.1107/S1399004714004878.

Reedy, C. J., Elvekrog, M. M., & Gibney, B. R. (2008). Development of a heme protein structure-electrochemical function database. *Nucleic Acids Research, 36*, D307–D313. https://doi.org/10.1093/nar/gkm814.

Rhéault, J.-F., Gagné, È., Guertin, M., Lamoureux, G., Auger, M., & Lagüe, P. (2015). Molecular model of hemoglobin N from *Mycobacterium tuberculosis* bound to lipid bilayers: A combined spectroscopic and computational study. *Biochemistry, 54*(11), 2073–2084. https://doi.org/10.1021/bi5010624.

Rice, S. L., Boucher, L. E., Schlessman, J. L., Preimesberger, M. R., Bosch, J., & Lecomte, J. T. J. (2015). Structure of *Chlamydomonas reinhardtii* THB1, a group 1 truncated hemoglobin with a rare histidine-lysine heme ligation. *Acta Crystallographica F Structural Biology Communications, 71*(Pt 6), 718–725. https://doi.org/10.1107/S2053230X15006949.

Rice, S. L., Preimesberger, M. R., Johnson, E. A., & Lecomte, J. T. J. (2014). Introduction of a covalent histidine-heme linkage in a hemoglobin: A promising tool for heme protein engineering. *Journal of Inorganic Biochemistry, 141*, 198–207. https://doi.org/10.1016/j.jinorgbio.2014.09.009.

Ritter, S., Hiller, R. G., Wrench, P. M., Welte, W., & Diederichs, K. (1999). Crystal structure of a phycourobilin-containing phycoerythrin at 1.90-Å resolution. *Journal of Structural Biology, 126*(2), 86–97. https://doi.org/10.1006/jsbi.1999.4106.

Rockwell, N. C., Martin, S. S., & Lagarias, J. C. (2023). Elucidating the origins of phycocyanobilin biosynthesis and phycobiliproteins. *Proceedings of the National Academy of Sciences of the United States of America, 120*(17), e2300770120. https://doi.org/10.1073/pnas.2300770120.

Rother, M., Oelgeschläger, E., & Metcalf, W. W (2007). Genetic and proteomic analyses of CO utilization by *Methanosarcina acetivorans*. *Archives of Microbiology, 188*(5), 463–472. https://doi.org/10.1007/s00203-007-0266-1.

Royer, W. E., Knapp, J. E., Strand, K., & Heaslet, H. A. (2001). Cooperative hemoglobins: Conserved fold, diverse quaternary assemblies and allosteric mechanisms. *Trends in Biochemical Sciences, 26*(5), 297–304. https://doi.org/10.1016/S0968-0004(01)01811-4.

Samuel, P. P., Smith, L. P., Phillips, G. N., Jr., & Olson, J. S. (2015). Apoglobin stability is the major factor governing both cell-free and in vivo expression of holomyoglobin. *Journal of Biological Chemistry, 290*(39), 23479–23495. https://doi.org/10.1074/jbc.M115.672204.

Samuel, P. P., White, M. A., Ou, W. C., Case, D. A., Phillips, G. N., Jr., & Olson, J. S. (2020). The interplay between molten globules and heme disassociation defines human hemoglobin disassembly. *Biophysical Journal, 118*, 1381–1400. https://doi.org/10.1016/j.bpj.2020.01.031.

Sanz-Luque, E., Chamizo-Ampudia, A., Llamas, A., Galvan, A., & Fernandez, E. (2015). Understanding nitrate assimilation and its regulation in microalgae. *Frontiers in Plant Science, 6*. https://doi.org/10.3389/fpls.2015.00899.

Sanz-Luque, E., Ocaña-Calahorro, F., de Montaigu, A., Chamizo-Ampudia, A., Llamas, A., Galvan, A., & Fernández, E. (2015). THB1, a truncated hemoglobin, modulates nitric oxide levels and nitrate reductase activity. *Plant Journal, 81*(3), 467–479. https://doi.org/10.1111/tpj.12744.

Sanz-Luque, E., Ocaña-Calahorro, F., Galván, A., & Fernández, E. (2015). THB1 regulates nitrate reductase activity and THB1 and THB2 transcription differentially respond to NO and the nitrate/ammonium balance in *Chlamydomonas*. *Plant Signaling & Behavior, 10*(8), e1042638. https://doi.org/10.1080/15592324.2015.1042638.

Sasso, S., Stibor, H., Mittag, M., & Grossman, A. R. (2018). The natural history of model organisms: From molecular manipulation of domesticated *Chlamydomonas reinhardtii* to survival in nature. *eLife, 7*, e39233. https://doi.org/10.7554/eLife.39233.

Schaus, L., Das, A., Knight, A. M., Jimenez-Osés, G., Houk, K. N., Garcia-Borràs, M., … Huang, X. (2023). Protoglobin-catalyzed formation of *cis*-trifluoromethyl-substituted cyclopropanes by carbene transfer. *Angewandte Chemie International Edition, 62*(4), e202208936. https://doi.org/10.1002/anie.202208936.

Schirmer, T., Bode, W., Huber, R., Sidler, W., & Zuber, H. (1985). X-ray crystallographic structure of the light-harvesting biliprotein C-phycocyanin from the thermophilic cyanobacterium *Mastigocladus laminosus* and its resemblance to globin structures. *Journal of Molecular Biology, 184*(2), 257–277. https://doi.org/10.1016/0022-2836(85)90379-1.

Schirmer, T., Huber, R., Schneider, M., Bode, W., Miller, M., & Hackert, M. L. (1986). Crystal structure analysis and refinement at 2.5 Å of hexameric C-phycocyanin from the cyanobacterium *Agmenellum quadruplicatum*: The molecular model and its implications for light-harvesting. *Journal of Molecular Biology, 188*(4), 651–676. https://doi.org/10.1016/S0022-2836(86)80013-4.

Schirrmeister, B. E., Gugger, M., & Donoghue, P. C. J. (2015). Cyanobacteria and the great oxidation event: Evidence from genes and fossils. *Palaeontology, 58*(5), 769–785. https://doi.org/10.1111/pala.12178.

Schmidt, M., Krasselt, A., & Reuter, W. (2006). Local protein flexibility as a prerequisite for reversible chromophore isomerization in α-phycoerythrocyanin. *Biochimica et Biophysica Acta (BBA) – Proteins and Proteomics, 1764*(1), 55–62. https://doi.org/10.1016/j.bbapap.2005.10.022.

Schneider, T., Tan, Y., Li, H., Fisher, J. S., & Zhang, D. (2022). Photoglobin, a distinct family of non-heme binding globins, defines a potential photosensor in prokaryotic signal transduction systems. *Computational and Structural Biotechnology Journal, 20*, 261–273. https://doi.org/10.1016/j.csbj.2021.12.022.

Schuster, C. D., Salvatore, F., Moens, L., & Martí, M. A. (2024). Globin phylogeny, evolution and function, the newest update. *Proteins: Structure, Function, and Bioinformatics, 2024*, 1–15. https://doi.org/10.1002/prot.26659.

Scott, N. L., & Lecomte, J. T. J. (2000). Cloning, expression, purification, and preliminary characterization of a putative hemoglobin from the cyanobacterium *Synechocystis* sp. PCC 6803. *Protein Science: A Publication of the Protein Society, 9*, 587–597.

Scott, N. L., Xu, Y., Shen, G., Vuletich, D. A., Falzone, C. J., Li, Z., … Lecomte, J. T. J. (2010). Functional and structural characterization of the 2/2 hemoglobin from *Synechococcus* sp. PCC 7002. *Biochemistry, 49*, 7000–7011. https://doi.org/10.1021/bi100463d.

Sgammato, R., Brempt, N. V., Aerts, R., Doorslaer, S. V., Dewilde, S., Herrebout, W., & Johannessen, C. (2023). Interaction of nitrite with ferric protoglobin from *Methanosarcina acetivorans* – An interesting model for spectroscopic studies of the haem–ligand interaction. *Dalton Transactions, 52*(10), 2976–2987. https://doi.org/10.1039/D2DT03252J.

Shandilya, M., Kumar, G., Gomkale, R., Singh, S., Khan, M. A., Kateriya, S., & Kundu, S. (2021). Multiple putative methemoglobin reductases in *C. reinhardtii* may support enzymatic functions for its multiple hemoglobins. *International Journal of Biological Macromolecules, 171*, 465–479. https://doi.org/10.1016/j.ijbiomac.2021.01.023.

Shibayama, N. (2020). Allosteric transitions in hemoglobin revisited. *Biochimica et Biophysica Acta (BBA) – General Subjects, 1864*(2), 129335. https://doi.org/10.1016/j.bbagen.2019.03.021.

Sievers, F., & Higgins, D. G. (2021). The clustal omega multiple alignment package. In K. Katoh (Ed.). *Multiple sequence alignment: Methods and protocols* (pp. 3–16). Springer US. https://doi.org/10.1007/978-1-0716-1036-7_1.

Smagghe, B. J., Sarath, G., Ross, E., Hilbert, J. L., & Hargrove, M. S. (2006). Slow ligand binding kinetics dominate ferrous hexacoordinate hemoglobin reactivities and reveal differences between plants and other species. *Biochemistry, 45*(2), 561–570. https://doi.org/10.1021/bi051902l.

Smagghe, B. J., Trent, J. T., 3rd, & Hargrove, M. S. (2008). NO dioxygenase activity in hemoglobins is ubiquitous in vitro, but limited by reduction in vivo. *PLoS One, 3*(4), e2039. https://doi.org/10.1371/journal.pone.0002039.

Stone, J. R., Sands, R. H., Dunham, W. R., & Marletta, M. A. (1996). Spectral and ligand-binding properties of an unusual hemoprotein, the ferric form of soluble guanylate cyclase. *Biochemistry, 35*(10), 3258–3262. https://doi.org/10.1021/bi952386.

Stranzl, G. R., Santelli, E., Bankston, L. A., La Clair, C., Bobkov, A., Schwarzenbacher, R., ... Liddington, R. C. (2011). Structural insights into inhibition of *Bacillus anthracis* sporulation by a novel class of non-heme globin sensor domains. *Journal of Biological Chemistry, 286*(10), 8448–8458. https://doi.org/10.1074/jbc.M110.207126.

Sturms, R., DiSpirito, A. A., Fulton, D. B., & Hargrove, M. S. (2011). Hydroxylamine reduction to ammonium by plant and cyanobacterial hemoglobins. *Biochemistry, 50*(50), 10829–10835. https://doi.org/10.1021/bi201425f.

Sturms, R., DiSpirito, A. A., & Hargrove, M. S. (2011). Plant and cyanobacterial hemoglobins reduce nitrite to nitric oxide under anoxic conditions. *Biochemistry, 50*(19), 3873–3878. https://doi.org/10.1021/bi2004312.

Terashima, M., Specht, M., Naumann, B., & Hippler, M. (2010). Characterizing the anaerobic response of *Chlamydomonas reinhardtii* by quantitative proteomics. *Molecular & Cellular Proteomics, 9*(7), 1514–1532. https://doi.org/10.1074/mcp.M900421-MCP200.

Thapa, N., Chaudhari, M., Iannetta, A. A., White, C., Roy, K., Newman, R. H., ... KC, D. B. (2021). A deep learning based approach for prediction of *Chlamydomonas reinhardtii* phosphorylation sites. *Scientific Reports, 11*(1), 12550. https://doi.org/10.1038/s41598-021-91840-w.

Thorsteinsson, M. V., Bevan, D. R., Ebel, R. E., Weber, R. E., & Potts, M. (1996). Spectroscopical and functional characterization of the hemoglobin of *Nostoc commune* UTEX 584 (Cyanobacteria). *Biochimica et Biophysica Acta – Protein Structure and Molecular Enzymology, 1292*(1), 133–139. https://doi.org/10.1016/0167-4838(95)00178-6.

Tiso, M., Tejero, J., Kenney, C., Frizzell, S., & Gladwin, M. T. (2012). Nitrite reductase activity of nonsymbiotic hemoglobins from *Arabidopsis thaliana*. *Biochemistry, 51*(26), 5285–5292. https://doi.org/10.1021/bi300570v.

Trent, J. T., 3rd, Kundu, S., Hoy, J. A., & Hargrove, M. S. (2004). Crystallographic analysis of *Synechocystis* cyanoglobin reveals the structural changes accompanying ligand binding in a hexacoordinate hemoglobin. *Journal of Biological Chemistry, 341*(4), 1097–1108. https://doi.org/10.1016/j.jmb.2004.05.070.

Uppal, S., Khan, M. A., & Kundu, S. (2020). Identification and characterization of a recombinant cognate hemoglobin reductase from *Synechocystis* sp. PCC 6803. *International Journal of Biological Macromolecules, 162*, 1054–1063. https://doi.org/10.1016/j.ijbiomac.2020.06.214.

Vávra, J., Sergunin, A., Jeřábek, P., Shimizu, T., & Martínková, M. (2022). Signal transduction mechanisms in heme-based globin-coupled oxygen sensors with a focus on a histidine kinase (AfGcHK) and a diguanylate cyclase (YddV or EcDosC). *Biological Chemistry, 403*(11–12), 1031–1042. https://doi.org/10.1515/hsz-2022-0185.

Vinogradov, S. N., Bailly, X., Smith, D. R., Tinajero-Trejo, M., Poole, R. K., & Hoogewijs, D. (2013). Microbial eukaryote globins. *Advances in Microbial Physiology, 63*, 391–446. https://doi.org/10.1016/B978-0-12-407693-8.00009-1.

Vinogradov, S. N., Fernandez, I., Hoogewijs, D., & Arredondo-Peter, R. (2011). Phylogenetic relationships of 3/3 and 2/2 hemoglobins in Archaeplastida genomes to bacterial and other eukaryote hemoglobins. *Molecular Plant, 4*(1), 42–58. https://doi.org/10.1093/mp/ssq040.

Vinogradov, S. N., Hoogewijs, D., Bailly, X., Arredondo-Peter, R., Guertin, M., Gough, J., … Vanfleteren, J. R. (2005). Three globin lineages belonging to two structural classes in genomes from the three kingdoms of life. *Proceedings of the National Academy of Sciences of the United States of America, 102*(32), 11385–11389. https://doi.org/10.1073/pnas.0502103102.

Vinogradov, S. N., Hoogewijs, D., Bailly, X., Mizuguchi, K., Dewilde, S., Moens, L., & Vanfleteren, J. R. (2007). A model of globin evolution. *Gene, 398*(1–2), 132–142. https://doi.org/10.1016/j.gene.2007.02.041.

Vinogradov, S. N., Tinajero-Trejo, M., Poole, R. K., & Hoogewijs, D. (2013). Bacterial and archaeal globins – A revised perspective. *Biochimica et Biophysica Acta (BBA) – Proteins and Proteomics, 1834*, 1789–1800. https://doi.org/10.1016/j.bbapap.2013.03.021.

Vu, B. C., Jones, A. D., & Lecomte, J. T. J. (2002). Novel histidine-heme covalent linkage in a hemoglobin. *Journal of the American Chemical Society, 124*, 8544–8545. https://doi.org/10.1021/ja026569c.

Vuletich, D. A., & Lecomte, J. T. J. (2006). A phylogenetic and structural analysis of truncated hemoglobins. *Journal of Molecular Evolution, 62*, 196–210. https://doi.org/10.1007/s00239-005-0077-4.

Wagner, U. G., Muller, N., Schmitzberger, W., Falk, H., & Kratky, C. (1995). Structure determination of the biliverdin apomyoglobin complex: Crystal structure analysis of two crystal forms at 1.4 and 1.5 Å resolution. *Journal of Molecular Biology, 247*(2), 326–337. https://doi.org/10.1006/jmbi.1994.0142.

Wang, H., Gau, B., Slade, W. O., Juergens, M., Li, P., & Hicks, L. M. (2014). The global phosphoproteome of *Chlamydomonas reinhardtii* reveals complex organellar phosphorylation in the flagella and thylakoid membrane. *Molecular & Cellular Proteomics, 13*(9), 2337–2353. https://doi.org/10.1074/mcp.M114.038281.

Watts, R. A., Hunt, P. W., Hvitved, A. N., Hargrove, M. S., Peacock, W. J., & Dennis, E. S. (2001). A hemoglobin from plants homologous to truncated hemoglobins of microorganisms. *Proceedings of the National Academy of Sciences of the United States of America, 98*(18), 10119–10124. https://doi.org/10.1073/pnas.191349198.

Wenke, B. B., Lecomte, J. T. J., Héroux, A., & Schlessman, J. L. (2014). The 2/2 hemoglobin from the cyanobacterium *Synechococcus* sp. PCC 7002 with covalently attached heme: Comparison of X-ray and NMR structures. *Proteins, 82*(3), 528–534. https://doi.org/10.1002/prot.24409.

Wittenberg, J. B., Bolognesi, M., Wittenberg, B. A., & Guertin, M. (2002). Truncated hemoglobins: A new family of hemoglobins widely distributed in bacteria, unicellular eukaryotes and plants. *Journal of Biological Chemistry, 277*, 871–874. https://doi.org/10.1074/jbc.R100058200.

Zhang, R., & Kennedy, M. A. (2021). Current understanding of the structure and function of pentapeptide repeat proteins. *Biomolecules, 11*(5), https://doi.org/10.3390/biom11050638.

Zhang, W., & Phillips, G. N. (2003). Structure of the oxygen sensor in *Bacillus subtilis*: Signal transduction of chemotaxis by control of symmetry. *Structure (London, England: 1993), 11*(9), 1097–1110. https://doi.org/10.1016/S0969-2126(03)00169-2.

Zhao, X., Zhou, J., Du, G., & Chen, J. (2021). Recent advances in the microbial synthesis of hemoglobin. *Trends in Biotechnology, 39*(3), 286–297. https://doi.org/10.1016/j.tibtech.2020.08.004.

Zones, J. M., Blaby, I. K., Merchant, S. S., & Umen, J. G. (2015). High-resolution profiling of a synchronized diurnal transcriptome from *Chlamydomonas reinhardtii* reveals continuous cell and metabolic differentiation. *The Plant Cell, 27*(10), 2743–2769. https://doi.org/10.1105/tpc.15.00498.