

Evolution / Évolution

Structure and function evolution in the superfamily of globins

Henri Wajcman^{a,*}, Laurent Kiger^b, Michael C. Marden^b

^a INSERM U841, Équipe 11, CHU Henri-Mondor, 94010 Créteil, France

^b INSERM U779, Hôpital de Bicêtre, 94275 Le Kremlin-Bicêtre cedex, France

Accepted after revision 30 July 2008

Available online 28 November 2008

Presented by Claude Combes

Abstract

The superfamily of globins has emerged some 4000 Myr from a common ancestor, which was among the basic protein components required for life. Globins are present in the three kingdoms of life. From a structure point of view, these molecules are defined by the presence of a characteristic protein fold, rich in α -helix, surrounding a heme group. Depending on the species or organs, they may be physiologically active as monomers, tetramers or large size polymers. Their function varies from the classical reversible binding of oxygen for transport and storage to cytoprotection against reactive oxygen species, NO scavenging, signaling in oxygen dependent metabolic pathways, or possibly other specific properties involving ligand or electron transfer. All these aspects are discussed in this review. **To cite this article:** *H. Wajcman et al., C. R. Biologies 332 (2009).*

© 2008 Académie des sciences. Published by Elsevier Masson SAS. All rights reserved.

Résumé

Évolution de la structure et de la fonction dans la superfamille des globines. La superfamille des globines est apparue il y a quelques 4 milliards d'années à partir d'un ancêtre commun qui se trouvait parmi les protéines indispensables à la vie. Les globines sont présentes dans les trois règnes du monde vivant. Ces molécules se définissent par une structure commune constituée par un repliement caractéristique de six à huit hélices α autour d'une molécule d'hème. Selon les espèces ou les organes, ces molécules peuvent être physiologiquement actives sous forme de monomères, de tétramères voire de polymères de grande taille. Leur fonction varie d'une liaison réversible à l'oxygène permettant le classique transport ou stockage de ce gaz à celui d'une protection de la cellule contre les agressions oxydantes. Il peut également s'agir d'un piégeage du monoxyde d'azote, ou encore d'un rôle de signalisation dans le contrôle de métabolisme ou de fonctions oxygène dépendantes. Des rôles nouveaux dans des processus de transports d'électrons ou de fixation de ligands sont régulièrement décrits. Tous ces aspects sont discutés dans cette revue. **Pour citer cet article :** *H. Wajcman et al., C. R. Biologies 332 (2009).*

© 2008 Académie des sciences. Published by Elsevier Masson SAS. All rights reserved.

Keywords: Hemoglobin; Myoglobin; Flavohemoglobin; Leghemoglobin

Mots-clés : Hémoglobine ; myoglobine ; Flavohémoglobine ; Leghémoglobine

In this review we will discuss the structure and function of proteins belonging to the globin superfamily, and their diversity considering the various species. We will first recall the common structural features shared by all

* Corresponding author.

E-mail address: Henri.Wajcman@inserm.fr (H. Wajcman).

these molecules, which is a typical three-dimensional folding of a small number of α helices – named the globin fold – protecting a non-covalently bound heme group and allowing a reversible oxygen binding. These molecules appeared about 4000 Myr and, depending on the species, developed functional properties varying from scavenging against oxidant species to oxygen transport and storage, and in some cases, some original unexpected properties.

1. Definition and variety of globins

A few years ago, the definition of a globin was simply a protein resembling myoglobin (Mb) or the alpha and beta chains of hemoglobin (Hb). Based on DNA analysis of human and many other organisms, the globin family has expanded. As more proteins are discovered and the details of their function emerge, it is more difficult to make a simple test of a globin. The exact sequence is highly variable and therefore not a clear test. We do not intend to include the known cytochromes as they are clearly separate in structure and function. Our globins bind oxygen reversibly and have structures nearly superimposable, but, as we discuss below, this still leaves room for some new molecular mechanisms.

Heme is a chemically highly active group that has been involved in biological processes as soon as life appeared. Heme participates in many biochemical functions among which are the binding and transport of gaseous ligands (O_2 , NO, CO), scavenging of free oxidant species, oxido-reduction, or oxygen-sensing. The specificity of these functions is directed by the structure of the protein to which it is associated. To remain soluble and active it needs to be surrounded by a hydrophobic environment that is obtained by a few structural 3D protein arrangements, the most common being the globin fold.

The classical globin molecule is characterized by the canonical 3/3 α -helical myoglobin fold, which has been described fifty years ago by Kendrew [1]. This molecule, ca 150 residues long, is characterized by a heme group surrounded by 8 helices designated A through H from the N to the C terminal. Helices A, B, C and E are on distal side of the heme and helices F, G and H on the proximal side (Fig. 1A). To this structure, a pattern of 37 hydrophobic residues at conserved, solvent-inaccessible positions has to be added [2]. These residues are summarized in Table 1. In human hemoglobin, the replacement of any of these residues leads to a hemoglobin molecule which is unstable or has altered oxygen binding properties. The main function of vertebrate hemoglobins and myoglobin is to bind oxy-

Table 1

Conserved residues of the globin structure.

| | |
|--|---|
| Hydrophobic residues at contacts between helices | A8, A11, A12, A15, B6, B9, B10, B13, B14, C4, E4, E7, E8, E11, E12, E15, E18, E19, F1, F4, G5, G8, G11, G12, G13, G15, G16, H7, H8, H11, H12, H15, and H19, |
| Inter-helical residues | CD1, CD4 and FG4 |
| Binding to heme iron | His at F8 |

gen reversibly for transport and storage, but additional roles, such as cytoprotection against reactive oxygen species and NO scavenging, are now well documented.

Initially, the description of globins was limited to myoglobins and to α - and β -subunits of hemoglobins in vertebrates. Hemoglobins were further described in some insects and as symbiotic hemoglobins (leghemoglobins) in the nodules of legume plants. Non-symbiotic hemoglobins were further found in a large variety of plants other than legumes [3]. The emergence of high throughput genome sequencing facilities and powerful high performance bioinformatic tools revealed an unexpected wide occurrence of putative globin sequences in the three kingdoms of life leading to a rapid growth of a “globin superfamily”. In all these molecules a similar 3D structure, characterized by the typical globin fold, has been found despite a large variety of functions and linear amino-acid sequences.

Within this family, a large group is made from globins, with sequences shorter than normal by 20 to 40 residues and named “truncated globins” (Fig. 1B). They were found in unicellular eukaryotes, [4], cyanobacteria [5], a nemertean [6], and a large number of bacteria [7,8]. The multiple amino acid deletions, insertions, and replacements resulted in these molecules in a distinctive alteration of the globin fold. These globins are single domain molecules whose crystal structures exhibit an abbreviated 2/2 α -helical fold characterized by a very short or absent A helix, a decreased CE inter-helical region and most of the F helix occurring as a loop, with only helices BC, E, G and H left to surround the heme group. Globins longer than normal (> 160 aa) were observed in a green alga [9] and in plants [10] with similar structural features. Since these molecules result from many short deletions and insertions spread in different regions all along the molecule, the term “2/2 Hb” has been coined to refer to this group. Amino-acid sequence similarity between a mammalian globin and a bacteria globin may be less than 10% but the three-dimensional pattern defining the globin fold is well conserved.

A question frequently debated is whether the 3/3 and 2/2 globins evolved independently from a cytochrome

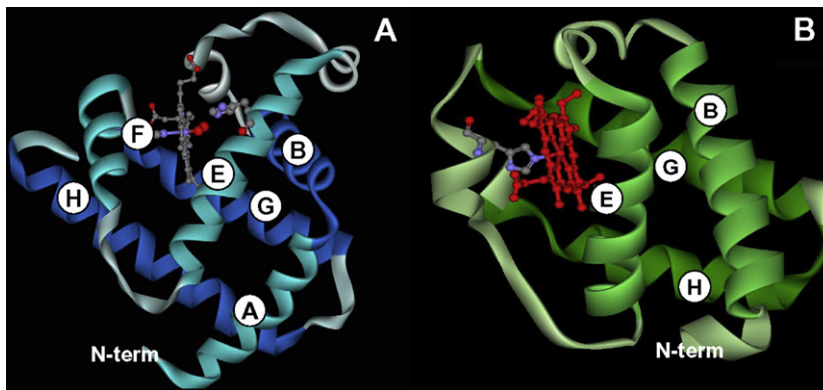


Fig. 1. Comparison between the three-dimensional structure of a 3/3 and a 2/2 globin. (A) Sperm whale myoglobin is the typical example of a 3/3 globin, the heme group is surrounded by 3 helices on its proximal site (F, G, and H) and 3 helices on its distal site (A, B and E). (B) *chlamydomonas eugametos* hemoglobin is an example of a 2/2 Hb. Here the heme group is sandwiched between helices B and E on the distal site and helices G and H on the proximal site, helix F is reduced to a single turn in which a His group – still designated as F8 in the 3D nomenclature – is coordinated to the heme iron.

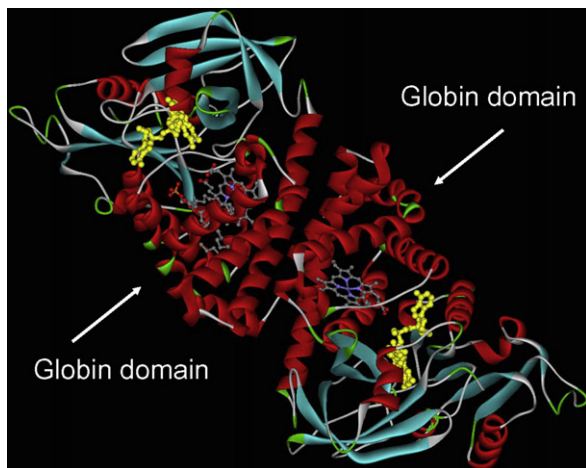


Fig. 2. Flavohemoglobin from *Alcaligenes eutropus*. This molecule has been crystallized as a dimer. The helical part of the molecule, which corresponds mostly to the globin domain surrounding the heme group, is shown in red.

b-like or other precursor(s) and which one of them represents the ancestral globin structure.

Another important lineage within the globin superfamily is made from chimeric proteins such as the flavohemoglobins which present a C-terminal FAD-binding domain and a N-terminal 3/3 Hb domain (Fig. 2). Globin coupled sensors with various C-terminal domains and having a 3/3 or a 2/2 Hb domain forms a second group [11].

Through the large accumulation of genomic information obtained in the last decade, additional globin families were recognized. This list includes the neuroglobins [12,13] and cytoglobins [14,15], which seem to be present in all vertebrates from humans to birds

and fish, GbE, the eye-specific globin in the domestic chicken *Gallus gallus*, related to cytoglobin [16], and GbX, a new and fifth type of globin gene in fish and amphibians, thought to have been lost in the higher vertebrates. Protoglobins, which are ca. 195 residues long, found in archaea and in several bacteria [17] are proposed to represent the true “ancestral” globin.

2. The 2/2 hemoglobins

The 2/2 hemoglobins, whose main structural characteristics have been described above, can be classified in three groups, according to some structural features. They display classically a primary structure 20 to 40 residues shorter than that of mammalian Hbs. The three groups of 2/2 Hbs are named I, II and III according to phylogenetic analyses based on amino acid sequences. Members of these groups are designated by an N, O, or P suffix, respectively [18]. Group III displays a high level of overall sequence conservation whereas group I and group II can be further separated into two and four subgroups, respectively. The sequence similarity between the proteins belonging to the different groups is low, with frequently less than 20% overall identity. In contrast, it may be higher than 80% within a group. Phylogenetic analyses suggest an evolutionary scenario in which the group II HbO gene is the ancestral one, and group I and group III genes result from duplication and transfer events. 2/2 Hbs from more than one group may coexist in a same bacteria, which suggest that they may play a large variety of functions. Among the functions that these molecules may provide are long-term ligand or substrate storage, nitric oxide (NO) detoxification,

O₂/NO sensing, redox reactions, and O₂ delivery under hypoxic conditions [19,20].

3. Chimeric proteins

3.1. Flavohemoglobins

Prevalent in many bacterial and fungal pathogens, flavohemoglobins [21] are also known as a “fungal defense” enzymes (Fig. 2). They have been shown to be closely tied to virulence. Upon infectious invasion, macrophages engulf pathogens and produce oxidants like nitric oxide to protect the organism. De Jesus-Berrios et al. [22] have shown that inactivation of flavohemoglobin led to a significant loss of virulence both in vitro and in vivo. Amino acid sequence analyses indicate that bacterial flavohemoglobins have 3/3 Hb fold and are closer to metazoan globins than most of the bacteria hemoglobins, which display a 2/2 Hb fold.

3.2. Globin coupled sensors

Globin coupled sensors (GCSs) form a large group of chimeric proteins having a N-terminal globin domain acting as a sensor to stimulate a signal controlled pathway [23]. They are one of the four known groups of heme sensors protein, among which one of the most studied was the kinase FixL, a PAS group protein [24]. In these molecules, the heme group, through a conformation change, induced by ligand binding, switch on or off an oxygen dependent metabolic pathway activating some genes. The first GCSs described were regulators of the aerotactic responses in *Bacillus subtilis* (HemAT-*Bs*) and *Halobacterium salinarum* (HemAT-*Hs*) [25]. Today, more than 30 GCSs are known and the evolution tree of their globin part suggests that they derived from a common ancestor, the “protoglobin”. Such protoglobins have indeed been discovered in two Archaea and two Bacteria [17]. Protoglobin from the strictly anaerobic methanogenic Archaea (*Methanosarcina acetivorans*) has been recently crystallized and revealed an almost completely buried heme group being only accessible to ligand through apolar tunnels [26]. GCSs could be classified in several categories: the first is made by aerotactics and contains the HemATs. Another group is formed by “second messengers” coupled, for example, to a guanylyl cyclase catalytic domain [23]. Two GCSs were found to be involved in protein-protein interactions through a STAS domain (sulfate transporter and antisigma factor antagonist). Finally, some GCSs were found to be tightly bound to the membrane through a domain which spans the membrane several times.

The evolutionary pathways of GCS and protoglobin are not yet clear. The initial ancestor could be a multidomain protein which split into components that further evolved on their own. Conversely, another hypothesis is that evolution may have started from individual domains forming a multidomain protein in which the two partners evolved together.

4. Phylogenic origin of globins

Globin, under the form of protoglobin, is likely to have been present in the first organisms living on Earth. It was perhaps included in the basic protein equipment of the Last Universal Common Ancestor (LUCA), the source of all life on Earth some 4000 Myr. The question debated now is whether the 3/3 and 2/2 globins evolved independently from a cytochrome *b*-like or other precursor(s) and which one of them represents the ancestral globin structure. During evolution, the problem of sequestering the water-insoluble heme in a hydrophobic protein cavity was solved in several ways such as the cytochrome fold or the catalase, and PAS domains, but the similarity of the globin fold in the 2/2 and 3/3 Hbs is a strong argument in favor of a unique globin origin. Although atmospheric O₂ levels were very low in the early history of the Earth [<0.0008 atm], local concentrations could have been high enough to promote the evolution of the globin ancestor [27].

Up to now there is no answer as to whether the 3/3 or the 2/2 fold is ancestral; both structures are present in the three kingdoms of life. In addition, this debate could not be clarified by the presence or the absence of introns in the genes encoding globins. Introns are present in all the 3/3 Hb genes and in many 2/2 Hbs, and when present they are always localized at the same position. It is therefore not yet possible to decide between the two scenarios.

The evolutionary tree of globins is also complicated by evidence for horizontal gene transfer (HGT) in some 3/3 Hbs from bacteria and in some flavohemoglobins [28].

5. Evolution and adaptation of globin function

5.1. General considerations

The presence in all three kingdoms of life of a large variety of globin molecules belonging to the different lineages raises the question of their functions. It is assumed that the globin ancestor was able to react with diatomic gaseous ligands inducing a change configuration that could be used in many tasks. Hemoglobin is

therefore able to detect, scavenge, and detoxify O₂ and O₂-derived species (e.g., NO and CO). The principal role of existent 2/2 Hbs is considered to be NO detoxification [29]. It is also the case for flavohemoglobins acting aerobically as NO oxygenases and anaerobically as NO reductases [30]. This NO detoxification is linked to bacterial resistance: it has been shown in pathogenic microorganisms that flavohemoglobins protect the bacteria from human macrophage NO-mediated killing and promotes the virulence of bacteria.

In plants three types of Hbs have been identified: symbiotic, non-symbiotic and 2/2 Hbs. Two classes of Hbs evolved prior to the divergence of monocots and dicots. Leghemoglobins, which are pentacoordinate O₂ transporters, emerged from the first class, whereas hexacoordinate scavenger/sensor Hbs emerged from both classes [31]. The function of the symbiotic hemoglobins is to establish a low oxygen concentration for the effective functioning of the bacterial nitrogenase necessary for nitrogen fixation. Although the role of 2/2 Hbs in plants is completely unknown, the non-symbiotic hemoglobins, which biosynthesis is induced in plant cells upon exposure to low oxygen concentrations, are thought to play a role in a metabolic pathway which also involves nitric oxide and which provides an alternative type of respiration to the mitochondrial electron transport under hypoxic conditions [32,33].

In the animal kingdom, the 3/3 Hbs and 2/2 Hbs are usually involved in oxygen transport and storage and NO detoxification, respectively. Nevertheless, in some species, hemoglobin molecules developed special evolution pattern allowing their carriers to occupy new ecological niches [29].

In some invertebrates, hemoglobins are found within the circulating cells of the coelomic fluid and/or within muscular type cells or cells of organ that need to be protected from anoxia or oxidant stress. These molecules are usually monomers or made from a small number of subunits.

In other species, such as insects, worms or mollusks, Hbs are freely dissolved in vascular, coelomic, or perienteric body fluids. They are circulating as giant polymers [34,35]. This is observed in the erythrocruorins from several oligochaetes (earthworm), achaetes (leech), vestimentiferans (hydrothermal vent tube worm) and polychaetes. *Lumbricus* erythrocruorin, shown here as an example, assembles from 144 oxygen-binding hemoglobin subunits and 36 structural “linker” subunits. Four distinct hemoglobin subunits are arranged into 12 dodecamers, which assemble onto a central core formed from linker subunits (Fig. 3) [36]. This supramolecular assembly involving several disul-

fide bonds raises many questions about the synchronization of the gene expression for the different structural elements and the subsequent folding and assembly mechanism. Such large complexes offer a number of important advantages as oxygen transport vehicles: erythrocruorins are retained in the vascular system, each complex has large oxygen binding capacity and subunits can be arranged to permit cooperative oxygen binding and regulatory features. *Lumbricus terrestris* erythrocruorin is probably the best studied giant Hb; it has been proposed as a potential model for a blood substitute [37].

5.2. High oxygen affinity: Structure function relationships

The globins have often served as a model for studying structure-function relations in proteins. However, they are not necessarily an easy system to describe for ligand binding, since the crystal structure did not reveal a ligand binding pathway. Rather than a lock and key type system involving a large surface area, ligand binding apparently requires protein fluctuations. The static structure therefore is not sufficient to provide information such as trajectories and affinities [1].

The globins of very high oxygen affinity led to a new description of the critical structural elements. It was originally thought that it was essentially an affair of iron-oxygen binding, in which the heme pocket provides a minor modulation, especially by the distal (E7) residue. It has now been shown that both the E7 and B10 position can participate in H-bonding directly to the oxygen molecule. This places the center of interest on the oxygen ligand itself, capable of forming up to three liaisons. The E7 and B10 positions will provide the bonds for major changes in affinity, which now approach a variation of nearly a factor of a million (see Fig. 4) [38,39].

Globins displaying these extremely high affinities will necessarily have specific residues in these positions capable of forming the H-bonds. Note that the specific residue in the B10 position may be a Tyr or a Phe, depending on the distance requirements. In the E7 position, the classical His residue may form a bond, but the very high affinities seem to employ a Gln residue.

5.3. Hexacoordination

The discovery of neuroglobin has led to the formulation of an alternate mechanism of oxygen binding. The external ligands (such as oxygen, CO, or NO) may still bind to the iron atom, but in competition with the

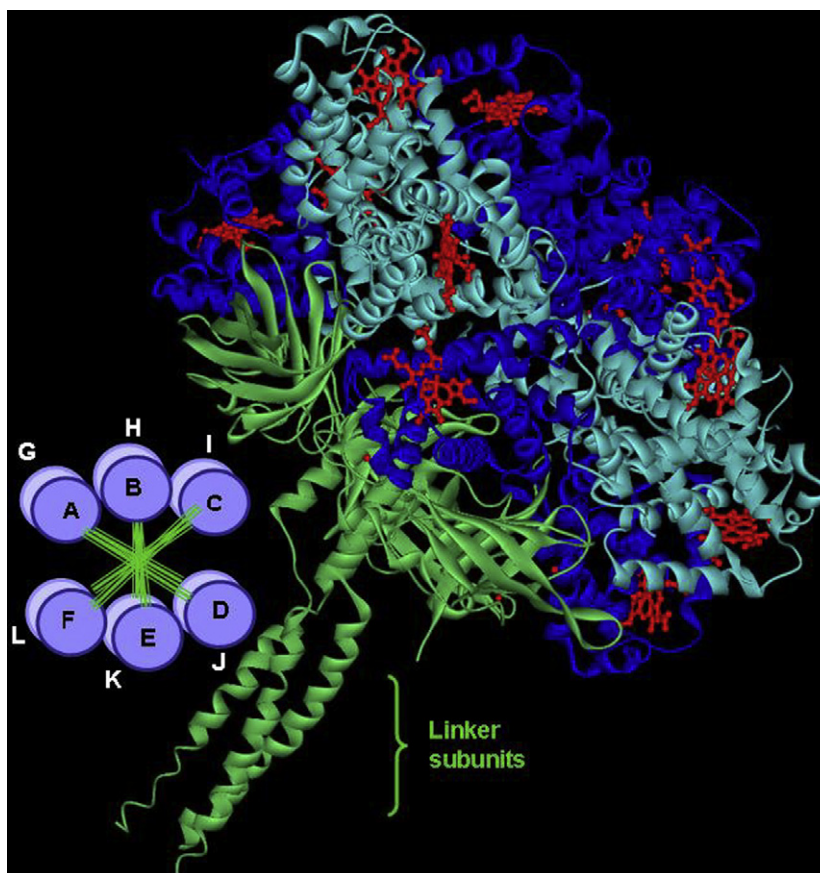


Fig. 3. General structure of *Lumbricus terrestris* erythrocrucorin showing the assembly of 12 dodecamers (A to L) in two parallel layers (bottom left). The three-dimensional structure of a single dodecamer is shown with a resolution of 3.5 Å, 12 such units bind through their linker subunits, which form a central core. In total a complete erythrocrucorin molecule carries 144 heme groups, each surrounded by a globin fold.

distal E7 His residue [40,41]. In the absence of external ligands, the ferrous iron may not be in a penta-coordinated state, since the distal His will bind to form a hexa-coordinated system [40]. Other globins such as cytoglobin, and new species of globin also display this mechanism of competitive binding; they are now collectively classed as hexa-coordinated globins to stress the fact that the penta-coordinated state is never more than a transient state.

Note that this mechanism has some interesting physiological consequences. The overall (observed) oxygen affinity is no longer the simple ratio of oxygen binding rates, but rather will be a ratio of the intrinsic oxygen and His binding affinities.

$$K_{O_2, \text{obs}} = K_{O_2, \text{penta}} / (1 + K_{\text{His}}) \quad (1)$$

In the case of neuroglobin, the His affinity ($K_{\text{His}} = k_{\text{on, His}} / k_{\text{off, His}}$) is over 100; in order to obtain a typical physiological (Mb-like) oxygen affinity, the intrinsic oxygen binding parameters must provide this extra fac-

tor of 100; this is accomplished by the extra H-bonds (via the E7 and B10 positions) as described in the section above [38–40]. This results in very slow oxygen dissociation, requiring, nearly 1 second. The compensating His affinity leads to the “normal” oxygen affinity of about 1 mm Hg, but results in a rather sluggish globin in terms of response to external ligand levels, since re-equilibration will require several seconds.

Another direct consequence of the competition is a reduced effect of external factors, such as temperature and pH. While the individual rate coefficients may be sensitive to temperature, the overall affinity is the ratio of affinities. If both the internal (His) and external (O_2) ligands have a similar response to temperature, the overall observed oxygen affinity may be nearly independent of temperature. This has been observed for oxygen binding to various Hex-globins, and even a reverse temperature effect for CO binding [41].

In species where Hbs play only a local role or a function in a signal pathway, oxygen transport is usually

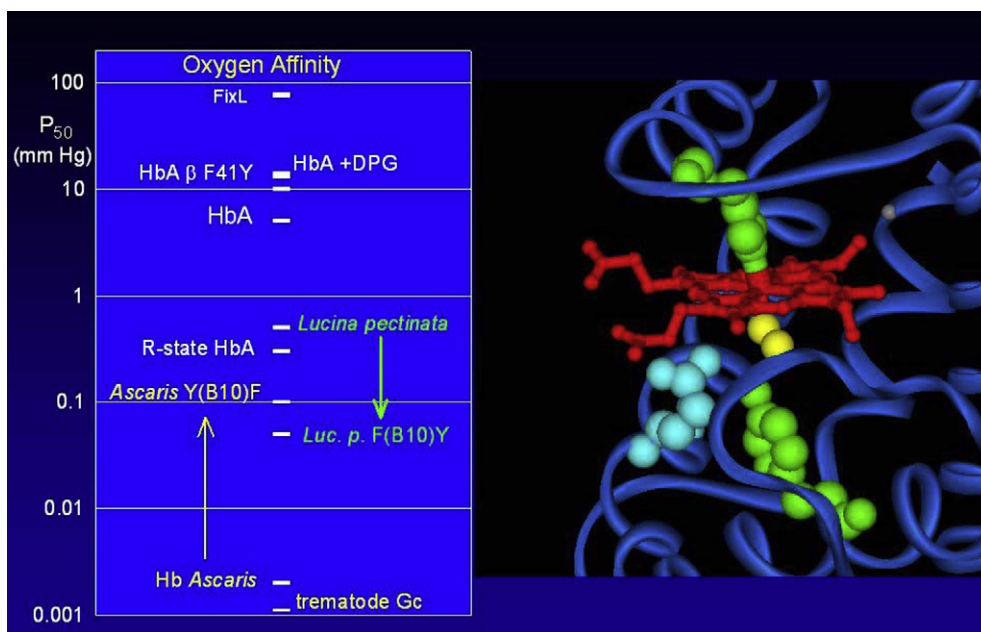


Fig. 4. The range of observed oxygen affinities covers nearly a factor of a million. The oxygen affinity (left panel) is the partial pressure of oxygen for 50% saturation of the hemes, presented here on a logarithmic scale to cover the large variation. Typical oxygen affinities are in the range of 0.1 to 10 mm Hg. The reference R and T states of tetrameric human Hb differ by about a factor of 100; the observed affinity for HbA can be modulated by the effector DPG; this effect can be mimicked by the introduction of mutations, such as replacing the phenylalanine at position 41 of the beta chain by a tyrosine. At the lower limit of oxygen affinity (high P_{50}) is FixL, which is not a globin but a heme based oxygen sensor from *rhizobium*. Very high oxygen affinities have been observed for the globins of certain parasites, such as *Ascaris* or the trematodes (*Gastrothylax crumenifer* shown here). The major contributions to the very high oxygen affinities are the formation of H-bonds directly to the Fe-bound oxygen molecule (yellow on the molecular image on the right hand panel) via critical protein residues at positions E7 and B10 (blue and green respectively) below the heme group (red). Modifications at these sites can provoke changes of over an order of magnitude in the oxygen affinity. A change of the B10 position residue of Hb *Ascaris* leads to a shift towards the normal range, while introduction of mutations at this site may shift a normal globin towards the very high affinity, as shown here for the globin from the clam *Lucina pectinata*.

assumed by other respiratory pigments such as hemocyanin.

5.4. A few examples

Hemoglobin from *Ascaris* works probably as an oxygen scavenger. Two Hbs, that have been thoroughly studied, are present in this species. The first one is an extracellular one that circulates as a 328 kDa octamer. Each of the 8 subunits is made from a 39.5 kDa polypeptide chain formed by two covalently bound globin domains (D1 and D2). In the C-terminal part of D2, the sequence “His-Lys-Glu-Glu” is repeated 4 times and is likely to play the role of a zipper for the association into octamer [42]. Electron microscopy pictures show that the octamers are made by two layers formed each by 4 subunits. This Hb displays one of the highest known oxygen affinities: its P_{50} is of about 0.004 mm Hg, which is 25 000 times higher than that of mammalian Hbs [39,43,44]. Since association rates are limited by diffusion, these very high affinities necessarily

display a considerably low value of the dissociation rate, making difficult the release of the bound oxygen. This property will hardly make this molecule suitable for oxygen transport requiring a prompt uptake and release. At the limit it could be a way to store oxygen in order to assume a minimal oxygen supply in case of extreme anaerobic conditions. Conversely, it has also been proposed that this Hb could be an oxygen scavenger allowing its destruction through a pathway involving NO [45]. In fact, nothing precise is known about the physiological role of this Hb, which remains a challenge. In the case of this Hb the only strong data are the molecular mechanism leading to this considerable oxygen affinity which consists in a network of hydrogen bonds extending in the heme pocket between a Gln at position E7 and the Tyr B10 stabilizing the $\text{Fe}(\delta^+) - \text{O}_2(\delta^-)$ dipole binding [39].

In the same species another type of Hb is present in the wall cells of the animal central cavity. It is a homodimer of 40.6 kDa stabilized by SH bridges. Its sequence is 38% similar to that of domain D1 and its P_{50} is of

about 0.1 mm Hg, which could allow a role for oxygen storage at the proximity of muscular cells. In the animal kingdoms, Hbs of nematodes, like neuroglobins and cytoglobins, are the only globins encoded by a gene containing 3 intervening sequences.

In some marine invertebrates, Hb is involved in sulfur transport. For example, the giant extracellular hexagonal bilayer hemoglobin (HBL-Hb) of the deep-sea hydrothermal vent tube worm *Riftia pachyptila* transports simultaneously O₂ and H₂S in the blood from the gills to a specific organ, the trophosome, that harbors sulfide-oxidizing endosymbionts. Within this organ, symbiotic bacteria use the sulfur as final electron acceptor in a metabolic pathway transforming the dissolved CO₂ into sugars [46]. Three Hbs are present in these animals: the vascular blood contains Hbs V1 and V2, and the coelomic fluid comprises Hb C1. These Hbs are able to bind oxygen and sulfide simultaneously and reversibly at two different sites. The amount of H₂S that they transport is high, varying from 1–10 mmol/L, depending upon its concentration in the surrounding medium. Hb of *Riftia pachyptila* is a giant polymer made from a basic unit formed by 4 chains with 1 free thiol group, connected by linker polypeptides. The MW of HbV1 is of about 3500 kDa, contains about 150 globin chains and some 40 linkers [47]. Hb V2 and C1 are smaller, with a mw of about 400 kDa, and contains 24 globin chains, many free SH groups are present.

Another example of Hb with high affinity for thiol group is observed in *Lucina pectinata*. This tropical clam contains two kinds of Hbs: the first one (HbI) is a monomer, 143 residues long, which serves to transport H₂S to autotrophic bacteria whereas two other globins (HbII and Hb III), respectively dimeric and tetrameric in the μ M range of [heme], are used for oxygen transport [48,49]. Site directed mutagenesis introducing a Gln residue at position E7 and a Phe at B10 and E11 in sperm whale myoglobin led to a Hb molecule with a high affinity for H₂S but nevertheless with one order of magnitude less than in *Lucina pectinata*. Even though further modifications of the heme pocket are thus required for the fine tuning mechanism of ligand binding this underlines the presence of key residues located at the distal side of the heme pocket for switching from O₂ to H₂S transport. Same conclusion holds inside the different *Lucina pectinata* globins. Indeed for Hb II and Hb III while the E7 and E11 residues do not change compared to Hb I the B10 Phe is replaced by a Tyr leading almost to three orders of magnitude decrease of the O₂ dissociation rate. Thus the addition of a single hydroxyl group is able to stabilize the dipolar Fe–O₂ bond through a favorable hydrogen bond, perhaps in associ-

ation with a network of electrostatic interactions such as those in the *Ascaris* globin explains the difference of O₂ off-rates between these globins. Note that the B10 residue is also involved in the large decrease of the O₂ off-rates for Hb II and Hb III compared to Hb I and more generally in the ligand accessibility from the solvent to the heme.

5.5. Neuroglobin: an allosteric protein?

As can be expected from the competitive ligand binding, a modification of the His affinity will directly influence the observed oxygen affinity. The effect is exploited by neuroglobin via an internal disulfide bond: when the S–S bond is broken, the His affinity increases by an order of magnitude, thereby decreasing the oxygen affinity [40]. One could therefore imagine a partner molecule which reduces the cysteines of neuroglobin and thereby provokes the release of oxygen. This would lead to a model of Ngb as an allosteric protein providing oxygen delivery on demand of a specific partner.

Neuroglobin and cytoglobin have shed light on an alternate globin mechanism for modulating the ligand binding. These globins belong to the class of hexacoordinated globins: in the absence of external ligands, the distal E7 His forms a stable bond with the heme iron. This bond is reversible so that the external ligand binding requires the replacement of the constitutive residue. The globins are then able to use a mechanism competitive ligand binding to the heme iron. The fine tuning mechanism appears more in the stress induced at a distance by an intra-disulfide bond; note that the location of this disulfide bond is different for Ngb (CD loop) vs Cygb (spanning the E helix with AB turn). Breaking the disulfide bond allows a shift in the orientation of the E-helix, resulting in an enhanced affinity of the distal E7 His for the heme and thereby a release of O₂. The redox state of the cell could act on the state of the cysteines of the neuroglobin and further influence other biochemical reactions involving an electron transfer to the molecular oxygen or another electron acceptor molecule and the NO catabolism.

5.6. An optical filter function

Hb from the nematode *Mermis nigrescens* (Mn-GLB-E), a parasite of grasshoppers and other Orthopteran insects, has an optical, light shadowing function [50]. The protein accumulates to high concentration as intracellular crystals in the ocellus of mature phototactic adult females while also being expressed at low concentration in other tissues. From eggs ingested by

the host, larvae grow to the adult length inside the body cavity. The last larvae stage break out of the host and burrow into the soil where the final molt takes place and the adult nematode matures. The gravid adult females emerge 1–2 years later from the soil and exhibit a positive phototaxis during the search for suitable egg-laying sites in grass. This positive phototaxis results from a pigmentation of the ocellus by formation of a thick Hb crystal. A sensory nerve ending lies anterior to this area and will point the head away from a light source as long as the ocellus would not be shadowed.

5.7. The smallest hemoglobin

The neural and body wall tissue of a nemertean worm, *Cerebratulus lacteus*, express a major Hb components for which the amino acid sequences have been deduced from cDNA and genomic DNA [51]. This Hb is 109 residues long and forms the smallest stable Hb known. The globin genes have three exons and two introns with splice sites in the highly conserved positions of globin genes. Alignment of the sequences with those of other globins indicates that helix A is almost missing, helix H is very short and GH corner elongated. Nevertheless the general 3D structure is that of a globin fold. It has been hypothesized that this Hb could be an ancestor of neuroHbs.

6. Conclusions

A large variety of Hbs with different structures and oxygen binding properties have been reported among invertebrates. The sequence alone is not sufficient to classify the protein as a globin; the large heterogeneity in the expanding database includes globins with differences above 90%. The protein structure and function need to be included in the overall consideration to further develop this superfamily. It is out of the scope of this mini-review to discuss all of them and we refer the reader to a few well documented reviews on this topic [52,53].

References

- [1] J.C. Kendrew, G. Bodo, H.M. Dintzis, R.G. Parrish, H. Wyckoff, D.C. Phillips, A three-dimensional model of the myoglobin molecule obtained by X-ray analysis, *Nature* 181 (1958) 662–666.
- [2] S.N. Vinogradov, D. Hoogewijs, X. Bailly, K. Mizuguchi, S. Dewilde, L. Moens, J.R. Vanfleteren, A model of globin evolution, *Gene* 398 (2007) 132–142.
- [3] S.N. Vinogradov, D. Hoogewijs, X. Bailly, R. Arredondo-Peter, J. Gough, S. Dewilde, L. Moens, J.R. Vanfleteren, A phylogenomic profile of globins, *BMC Evol. Biol.* 7 (2006) 31.
- [4] T. Takagi, Hemoglobins from single-celled organisms, *Curr. Opin. Struct. Biol.* 3 (1993) 413–418.
- [5] M. Potts, S. Angeloni, R. Ebel, D. Bassam, Myoglobin in a cyanobacterium, *Science* 256 (1992) 1690–1692.
- [6] T. Vandergon, C. Riggs, T. Gorr, J. Colacino, A. Riggs, The mini-hemoglobins in neural and body wall tissue of the nemertean worm *Cerebratulus lacteus*, *J. Biol. Chem.* 273 (1998) 16998–17011.
- [7] M. Couture, S. Yeh, B.A. Wittenberg, J.B. Wittenberg, Y. Ouellet, D. Rousseau, M. Guertin, A cooperative oxygen-binding hemoglobin from *Mycobacterium tuberculosis*, *Proc. Natl. Acad. Sci. USA* 96 (1999) 11223–11228.
- [8] J.B. Wittenberg, M. Bolognesi, B.A. Wittenberg, M. Guertin, Truncated hemoglobins: a new family of hemoglobins widely distributed in bacteria, unicellular eukaryotes, and plants, *J. Biol. Chem.* 277 (2002) 871–874.
- [9] M. Couture, H. Chamberland, B. St-Pierre, J. Lafontaine, M. Guertin, Nuclear genes encoding chloroplast hemoglobins in the unicellular green alga *Chlamydomonas eugametos*, *Mol. Gen. Genet.* 243 (1994) 185–197.
- [10] R. Watts, P. Hunt, A. Hvitved, M. Hargrove, W. Peacock, E. Dennis, A hemoglobin from plants homologous to truncated hemoglobins of microorganisms, *Proc. Natl. Acad. Sci. USA* 98 (2001) 10119–10124.
- [11] S.N. Vinogradov, D. Hoogewijs, X. Bailly, R. Arredondo-Peter, M. Guertin, J. Gough, S. Dewilde, L. Moens, J.R. Vanfleteren, Three globin lineages belonging to two structural classes in genomes from the three kingdoms of life, *Proc. Natl. Acad. Sci. USA* 102 (2005) 11385–11389.
- [12] T. Burmester, B. Weich, S. Reinhardt, T. Hankeln, A vertebrate globin expressed in the brain, *Nature* 407 (2000) 520–523.
- [13] J. Trent, R. Watts, M. Hargrove, Human neuroglobin, a hexacoordinate hemoglobin the reversibly binds oxygen, *J. Biol. Chem.* 276 (2001) 30106–30110.
- [14] T. Burmester, B. Ebner, B. Weich, T. Hankeln, Cytoglobin: a novel globin type ubiquitously expressed in vertebrate tissues, *Mol. Biol. Evol.* 19 (2002) 416–421.
- [15] T. Hankeln, B. Ebner, C. Fuchs, F. Gerlach, M. Haberkamp, T.L. Laufs, A. Roesner, M. Schmidt, B. Weich, S. Wystub, S. Saaler-Reinhardt, S. Reuss, M. Bolognesi, D. De Sanctis, M.C. Marden, L. Kiger, L. Moens, S. Dewilde, E. Nevo, A. Avivi, R.E. Weber, A. Fago, T. Burmester, Neuroglobin and cytoglobin in search of their role in the vertebrate globin family, *J. Inorg. Biochem.* 99 (2005) 110–119.
- [16] D. Kugelstadt, M. Haberkamp, T. Hankeln, T. Burmester, Neuroglobin, cytoglobin, and a novel, eye-specific globin from chicken, *Biochem. Biophys. Res. Commun.* 325 (2004) 719–725.
- [17] T. Freitas, S. Hou, E. Dioum, J. Saito, J. Newhouse, G. Gonzalez, M.A. Gilles-Gonzalez, M. Alam, Ancestral hemoglobins in Archaea, *Proc. Natl. Acad. Sci. USA* 101 (2004) 6675–6680.
- [18] M. Nardini, A. Pesce, M. Milani, M. Bolognesi, Protein fold and structure in the truncated (2/2) globin family, *Gene* 398 (2007) 2–11.
- [19] J.B. Wittenberg, M. Bolognesi, B.A. Wittenberg, M. Guertin, Truncated hemoglobins: a new family of hemoglobins widely distributed in bacteria, unicellular eukaryotes, and plants, *J. Biol. Chem.* 277 (2002) 871–874.
- [20] D.A. Vuletich, J.T. Lecomte, A phylogenetic and structural analysis of truncated hemoglobins, *J. Mol. Evol.* 62 (2006) 196–210.
- [21] H. Zhu, A. Riggs, Yeast flavohemoglobin is an ancient protein related to globins and a reductase family, *Proc. Natl. Acad. Sci. USA* 89 (1992) 5015–5019.

- [22] M. de Jesús-Berríos, L. Liu, J.C. Nussbaum, G.M. Cox, J.S. Stamler, J. Heitman, Enzymes that counteract nitrosative stress promote fungal virulence, *Curr. Biol.* 13 (2003) 1963–1968.
- [23] M.A. Gilles-Gonzalez, G. Gonzalez, Heme-based sensors: defining characteristics, recent developments, and regulatory hypotheses, *J. Inorg. Biochem.* 99 (2005) 1–22.
- [24] M.A. Gilles-Gonzalez, G. Gonzalez, M.F. Perutz, L. Kiger, M.C. Marden, C. Poyart, Heme-based sensors, exemplified by the kinase FixL, are a new class of heme protein with distinctive ligand binding and autoxidation, *Biochemistry* 33 (1994) 8067–8073.
- [25] S. Hou, R.W. Larsen, D. Boudko, C.W. Riley, E. Karatan, M. Zimmer, G.W. Ordal, M. Alam, Myoglobin-like aerotaxis transducers in Archaea and Bacteria, *Nature* 403 (2000) 540–544.
- [26] M. Nardini, A. Pesce, L. Thijs, J.A. Saito, S. Dewilde, M. Alam, P. Ascenzi, M. Coletta, C. Ciaccio, L. Moens, M. Bolognesi, Archaeal protoglobin structure indicates new ligand diffusion paths and modulation of haem-reactivity, *EMBO Rep.* 9 (2008) 157–163.
- [27] T.A. Freitas, J.A. Saito, S. Hou, M. Alam, Globin-coupled sensors, protoglobins, and the last universal common ancestor, *J. Inorg. Biochem.* 99 (2005) 23–33.
- [28] L. Moens, J. Vanfleteren, Y. Van de Peer, K. Peeters, O. Kapp, J. Czeluzniak, M. Goodman, M. Blaxter, S.N. Vinogradov, Globins in nonvertebrate species: dispersal by horizontal gene transfer and evolution of the structure-function relationships, *Mol. Biol. Evol.* 13 (1996) 324–333.
- [29] H. Wajcman, L. Kiger, Hemoglobin, from microorganisms to man: a single structural motif, multiple functions, *C. R. Biol.* 325 (2002) 1159–1174.
- [30] A.M. Gardner, P.R. Gardner, Flavohemoglobin detoxifies nitric oxide in aerobic, but not anaerobic, *Escherichia coli*. Evidence for a novel inducible anaerobic nitric oxide-scavenging activity, *J. Biol. Chem.* 277 (2002) 8166–8171.
- [31] J.A. Hoy, H. Robinson, J.T. Trent, S. Kakar, B.J. Smagghe, M.S. Hargrove, Plant hemoglobins: a molecular fossil record for the evolution of oxygen transport, *J. Mol. Biol.* 371 (2007) 168–179.
- [32] V. Garrocho-Villegas, S.K. Gopalasubramaniam, R. Arredondo-Peter, Plant hemoglobins: what we know six decades after their discovery, *Gene* 398 (2007) 78–85.
- [33] K.H. Hebelstrup, A.U. Igamberdiev, R.D. Hill, Metabolic effects of hemoglobin gene expression in plants, *Gene* 398 (2007) 86–93.
- [34] B.N. Green, T. Gotoh, T. Suzuki, F. Zal, F.H. Lallier, A. Toulmond, S.N. Vinogradov, Observation of large, non-covalent globin subassemblies in the approximately 3600 kDa hexagonal bilayer hemoglobins by electrospray ionization time-of-flight mass spectrometry, *J. Mol. Biol.* 309 (2001) 553–560.
- [35] T. Suzuki, S.N. Vinogradov, Globin and linker sequences of the giant extracellular hemoglobin from the leech *Macrobdella decora*, *J. Protein. Chem.* 22 (2003) 231–242.
- [36] W.E. Royer, H. Sharma, K. Strand, J.E. Knapp, B. Bhyravhatla, Lumbricus erythrocyruorin at 3.5 Å resolution: architecture of a megadalton respiratory complex, *Structure* 14 (2006) 1167–1177.
- [37] R.E. Hirsch, L.A. Jelicks, B.A. Wittenberg, D.K. Kaul, H.L. Shear, J.P. Harrington, A first evaluation of the natural high molecular weight polymeric *Lumbricus terrestris* hemoglobin as an oxygen carrier, *Artif. Cells Blood Substit. Immobil. Biotechnol.* 25 (1997) 429–444.
- [38] T.E. Carver, R.E. Brantley, E.W. Singleton, R.M. Arduini, M.L. Quillin, G.N. Phillips, J.S. Olson, A novel site-directed mutant of myoglobin with an unusually high O₂ affinity and low autooxidation rate, *J. Biol. Chem.* 267 (1992) 14443–14450.
- [39] I. De Baere, M.F. Perutz, L. Kiger, M.C. Marden, C. Poyart, Formation of two hydrogen bonds from the globin to the heme-linked oxygen molecule in *Ascaris* hemoglobin, *Proc. Natl. Acad. Sci. USA* 91 (1994) 1594–1597.
- [40] D. Hamdane, L. Kiger, S. Dewilde, B.N. Green, A. Pesce, J. Uzan, T. Burmester, T. Hankeln, M. Bolognesi, L. Moens, M.C. Marden, The redox state of the cell regulates the ligand binding affinity of human neuroglobin and cytoglobin, *J. Biol. Chem.* 278 (2003) 51713–51721.
- [41] J. Uzan, S. Dewilde, T. Burmester, T. Hankeln, L. Moens, D. Hamdane, M.C. Marden, L. Kiger, Neuroglobin and other hexacoordinated hemoglobins show a weak temperature dependence of oxygen binding, *Biophys. J.* 87 (2004) 1196–1204.
- [42] I. De Baere, L. Liu, L. Moens, J. Van Beeumen, C. Gielen, J. Richelle, C. Trotman, J. Finch, M. Gerstein, M. Perutz, Polar zipper sequence in the high-affinity hemoglobin of *Ascaris suum*: amino acid sequence and structural interpretation, *Proc. Natl. Acad. Sci. USA* 89 (1992) 4638–4642.
- [43] J. Yang, A.P. Kloek, D.E. Goldberg, F.S. Mathews, The structure of *Ascaris* hemoglobin domain I at 2.2 Å resolution: molecular features of oxygen avidity, *Proc. Natl. Acad. Sci. USA* 92 (1995) 4224–4228.
- [44] L. Kiger, A.K. Rashid, N. Griffon, M. Haque, L. Moens, Q.H. Gibson, C. Poyart, M.C. Marden, Trematode hemoglobins show exceptionally high oxygen affinity, *Biophys. J.* 75 (1998) 990–998.
- [45] D.M. Minning, A.J. Gow, J. Bonaventura, R. Braun, M. Dewhurst, D.E. Goldberg, J.S. Stamler, *Ascaris* haemoglobin is a nitric oxide-activated ‘deoxygenase’, *Nature* 401 (1999) 497–502.
- [46] X. Bailly, D. Jollivet, S. Vanin, J. Deutsch, F. Zal, F. Lallier, A. Toulmond, Evolution of the sulfide-binding function within the globin multigenic family of the deep-sea hydrothermal vent tubeworm *Riftia pachyptila*, *Mol. Biol. Evol.* 19 (2002) 1421–1433.
- [47] C. Chabasse, X. Bailly, S. Sanchez, M. Rousselot, F. Zal, Gene structure and molecular phylogeny of the linker chains from the giant annelid hexagonal bilayer hemoglobins, *J. Mol. Evol.* 63 (2006) 365–374.
- [48] D.W. Kraus, J.B. Wittenberg, Hemoglobins of the *Lucina pectinata*/bacteria symbiosis. I. Molecular properties, kinetics and equilibria of reactions with ligands, *J. Biol. Chem.* 265 (1990) 16043–16053.
- [49] M. Rizzi, J.B. Wittenberg, A. Coda, M. Fasano, P. Ascenzi, M. Bolognesi, Structure of the sulfide-reactive hemoglobin from the clam *Lucina pectinata*. Crystallographic analysis at 1.5 Å resolution, *J. Mol. Biol.* 244 (1994) 86–99.
- [50] A.H.J. Burr, P. Hunt, D.R. Wagar, S. Dewilde, M.L. Blaxter, J.R. Vanfleteren, L. Moens, A hemoglobin with an optical function, *J. Biol. Chem.* 275 (2000) 4810–4815.
- [51] A. Pesce, M. Nardini, S. Dewilde, E. Geuens, K. Yamauchi, P. Ascenzi, A.F. Riggs, L. Moens, M. Bolognesi, The 109 residue nerve tissue minihemoglobin from *Cerebratulus lacteus* highlights striking structural plasticity of the alpha-helical globin fold, *Structure* 10 (2002) 725–735.
- [52] M. Bolognesi, D. Bordo, M. Rizzi, C. Tarricone, P. Ascenzi, Nonvertebrate hemoglobins: structural bases for reactivity, *Prog. Biophys. Mol. Biol.* 68 (1997) 29–68.
- [53] R.E. Weber, S.N. Vinogradov, Nonvertebrate hemoglobins: functions and molecular adaptations, *Physiol. Rev.* 81 (2001) 569–628.