The FEBS Journal

FEBS PRESS science publishing by scientists

Genetic and functional diversity of the multiple lungfish myoglobins

Julia Lüdemann[1], Angela Fago[2], Sven Falke[3], Michelle Wisniewsky[1], Igor Schneider[4], Andrej Fabrizius[1] and Thorsten Burmester[1]

1 Institute of Zoology, Department of Biology, University of Hamburg, Germany
2 Department of Bioscience, Aarhus University, Denmark
3 Institute for Biochemistry and Molecular Biology, Department of Chemistry, University of Hamburg, Germany
4 Instituto de Ciências Biológicas, Universidade Federal do Pará, Belém, Brazil

Keywords
binding kinetics; lungfish; myoglobin; oxygen; subfunctionalisation

Correspondence
T. Burmester, Institute of Zoology, University of Hamburg, Martin-Luther-King-Platz 3, D-20146 Hamburg, Germany
Tel: +49(0)40-42838-3913
E-mail: thorsten.burmester@uni-hamburg.de

(Received 21 November 2018, revised 21 March 2019, accepted 11 October 2019)

doi:10.1111/febs.15094

It is known that the West African lungfish (*Protopterus annectens*) harbours multiple myoglobin (Mb) genes that are differentially expressed in various tissues and that the Mbs differ in their abilities to confer tolerance towards hypoxia. Here, we show that other lungfish species (*Protopterus dolloi*, *Protopterus aethiopicus* and *Lepidosiren paradoxa*) display a similar diversity of Mb genes and have orthologous Mb genes. To investigate the functional diversification of these genes, we studied the structures, O_2 binding properties and nitrite reductase enzymatic activities of recombinantly expressed *P. annectens* Mbs (PanMbs). CD spectroscopy and small-angle X-ray scattering revealed the typical globin-fold in all investigated recombinant Mbs, indicating a conserved structure. The highest O_2 affinity was measured for PanMb2 (P_{50} = 0.88 Torr at 20 °C), which is mainly expressed in the brain, whereas the muscle-specific PanMb1 has the lowest O_2 affinity (P_{50} = 3.78 Torr at 20 °C), suggesting that tissue-specific O_2 requirements have resulted in the emergence of distinct Mb types. Two of the mainly neuronally expressed Mbs (PanMb3 and PanMb4b) have the highest nitrite reductase rates. These data show different O_2 binding and enzymatic properties of lungfish Mbs, reflecting multiple subfunctionalisation and neofunctionalisation events that occurred early in the evolution of lungfish. Some Mbs may have also taken over the functions of neuroglobin and cytoglobin, which are widely expressed in vertebrates but appear to be missing in lungfish.

Introduction

Lungfish (Dipnoi) are of particular scientific interest because they are considered as 'living fossils' and the closest living relatives of tetrapods [1–3]. Moreover, lungfish harbour various specific morphological, physiological and molecular adaptations that allow them to tolerate hypoxic environments [4]. Recent studies identified eight distinct myoglobin (*Mb*) genes in the West African lungfish (*Protopterus annectens*) [5]. This diversity is unusual because most vertebrates harbour a single *Mb* in the haploid genomes [6].

Myoglobin is a member of the superfamily of globins, which is a classic model system for the study of the evolution and function of proteins and genes [7–10]. Globins are small, globular proteins that are mainly responsible for the binding of oxygen (O_2), but also may have other functions. Globins have a Fe^{2+}

Abbreviations
Cygb, cytoglobin; GbE, globin E; Hb, haemoglobin; Mb, myoglobin; Ngb, neuroglobin; PAM, point accepted mutation; RNA-Seq, RNA sequencing; ROS, reactive oxygen species; RPKM, reads per kilobase exon model per million reads.

ion in the centre of their porphyrin prosthetic group and share the unique structure that typically consists of eight alpha helices. In addition to the well-known Mb and haemoglobin (Hb), vertebrates harbour six other globin types, which commenced to diversify even before the radiation of vertebrates [6]. The globin types are as follows: neuroglobin (Ngb) [11], cytoglobin (Cygb) [12–14], globin X, globin Y, globin E (GbE) [15,16] and androglobin [17]. However, the functions of these recently identified globins are mostly not well understood [6,18,19].

Myoglobin is a monomeric protein consisting of ~ 150 amino acids (~ 17 kDa). It was the first protein with a resolved protein structure at an atomic scale [20,21]. The main function of Mb is the transport and storage of O_2 in striated (skeletal and heart) muscles [22–24]. For the effective extraction of oxygen from the blood, the typical Mb has a higher O_2 affinity, that is lower P_{50} value (partial pressure at which the protein is 50% saturated with O_2), than Hb. Most mammalian and avian Mbs have a P_{50} of ~ 1 Torr at 20 °C [25–27]. In contrast, Mbs of fish have P_{50} values between 1 Torr (e.g. bluefin tuna, blue marlin and common carp) [28,29] and 4.9 Torr (rainbow trout) [30,31]. This variation can be related to structural differences: fish Mbs are usually shorter (~ 146 aa) than mammalian Mbs and lack the D-helix [24]. Remarkably, Mb O_2 affinity, unlike Hb O_2 affinity, does not correlate with the size of the individual [32].

In the past 20 years, it has become evident that Mb also exhibits enzymatic functions. For example, oxygenated Mb rapidly converts the bioactive nitric oxide (NO) signalling molecule into nitrate (NO_3^-) under normoxic conditions [33]. The removal of NO enhances mitochondrial respiration in normoxia as NO acts as a reversible inhibitor of cytochrome oxidase [34]. While reducing NO, oxy-Mb is oxidised to met-Mb that needs to be reduced by the cellular Mb reductase [35–37]. Furthermore, under hypoxia, Mb has the opposite effect whereby deoxy-Mb functions as nitrite reductase producing NO from nitrite (NO_2^-) [38], thereby limiting O_2 consumption rates [24,33,39]. It has also been suggested that Mb functions in oxidative defence by removing reactive oxygen species (ROS) [35–37]. Additional functions of Mb may involve an interaction with fatty acids [40].

Most vertebrates harbour a single *Mb* gene in their haploid genomes [6]. Known exceptions are some cyprinid fish, which harbour two distinct *Mb* genes [41,42]. While Mb1 is widely expressed in various tissues, the expression of Mb2 is restricted to the brain [41,43]. These two Mbs apparently have distinct primary functions: *in vitro* studies have suggested that Mb1 mainly supplies O_2 and produces NO, whereas Mb2 efficiently eliminates H2 O_2 in the brain [28]. Some vertebrate species lack an *Mb* gene, such as frogs, some ice-fishes, the stickleback and the opossum *Monodelphis domestica* [44–47]. These species must have specific mechanisms to compensate for the lack of Mb. Especially, in teleost fish the loss of cardiac Mb evolved repeatedly during their evolution. The Mb deficit is in connection with a pale heart colour and affects water- and air-breathing fish species from all salinities and habitats [48].

Using quantitative real-time reverse transcription PCR and RNA sequencing (RNA-Seq), we found that the multiple Mb genes of the West African lungfish (PanMbs) are differentially expressed in tissues [5]. Surprisingly, most of the PanMbs have the highest expression levels in the brain. Further, we found that striated muscles express different sets of PanMbs: PanMb1 was highest in the skeletal muscle and PanMb4a in the heart. The unexpected diversity of the Mbs in the West African lungfish hinted to a functional diversification by subfunctionalisation and/or neofunctionalisation. Some PanMbs may have taken over the functions of Ngb and Cygb, which could not be detected in the lungfish transcriptomes. To better understand the evolution of *Mb* genes, we have studied their diversification in four lungfish species. Further, the functional differentiation was accessed by oxygen-binding, enzymatic and structural studies.

Material and methods

Identification and analyses of lungfish *Mb* cDNA sequences

The transcriptomes of the West African lungfish *P. annectens*, the slender lungfish *Protopterus dolloi*, the marbled lungfish *Protopterus aethiopicus* and the South American lungfish *Lepidosiren paradoxa* were retrieved from the public SRA database at GenBank (for accession numbers, see Table S1). The transcriptomes of each lungfish species were assembled using the CLC Genomics Workbench, version 11.0.1 (Qiagen, Hilden, Germany) using default parameters. *Mb* cDNA sequences were identified employing BLAST searches, using the Mb cDNA sequences from *P. annectens* [5] as queries. When required, the putative Mb sequences were re-assembled from the Illumina reads. RNA-Seq analyses were performed with the CLC Genomics Workbench (mismatch cost: 2, insertion cost: 3, deletion cost: 3, length fraction: 0.95, similarity fraction: 0.95). The mRNA levels of the Mbs were then calculated as reads per kilobase million (RPKM).

The Mb amino acid sequences were included in an alignment of vertebrate Mbs used before [5] (Table S2). A multiple sequence alignment was obtained with MAFFT using the L-INS-i method [49,50]. Phylogenetic analysis was performed with MRBAYES 3.2.3 [51,52] using the JTT model [53], which was selected by a PROTTEST analysis [50,54]. MRBAYES was run for 5 000 000 generations with two independent runs and four simultaneous chains. The trees were sampled every 1000th generation, and the posterior probabilities were estimated after discarding the initial 25% of the trees as burn-in. The maximum likelihood analysis was carried out using the IQ-TREE web server with the JTT substitution model [55–57]. One thousand bootstrap alignments were run with a maximum of 1000 iterations and a minimum correlation coefficient of 0.99. The perturbation strength was 0.5 and the IQ-TREE stopping rule 100. The values written on the branches show the ultrafast bootstrap support (Fig. 1).

Recombinant expression and purification of *P. annectens* Mbs

Tissue preparation, RNA extraction and cDNA cloning of *P. annectens* Mbs were performed as described before [5]. The cDNAs were cloned into pET16b using the restriction sites for correct sequence orientation (Novagen-Merck Biosciences, Darmstadt, Germany). The vectors were then transfected into *Escherichia coli* BL21(DE3) pLysS for recombinant expression. *Escherichia coli* were grown overnight at 37 °C in a 5 mL L-medium (1% bactotryptone, 0.5% yeast extract, 0.5% NaCl, pH 7.5), supplemented with 10 $\mu g \cdot mL^{-1}$ ampicillin and 34 $\mu g \cdot mL^{-1}$ chloramphenicol. Subsequently, the culture was transferred to 1 L L-medium supplemented with 1 mL of 1 M 5-aminolevulinic acid and cells were grown at 37 °C until an OD600 between 0.4 and 0.8 was reached. The expression was then induced by 0.4 mM isopropyl-1-thio-β-D-galactopyranoside, and the culture was incubated for an additional 3 h at 37 °C. The cell suspension was centrifuged at 4000 ***g*** and 4 °C for 45 min, and the pellet with the bacterial cells was dissolved in protease inhibitor buffer [50 mM Tris/HCl, 1 mM dithiothreitol, 1 mM Pefabloc, 1 mM $MgCl_2$, 1× Complete™ protease inhibitor mix (Roche Applied Science, Penzberg, Germany), 80 µg DNase, RNase, pH 8.0]. After exposing the cells to three freeze–thaw steps in liquid nitrogen and disruption by sonification (10 × 30 s), the cells were incubated at 37 °C for 2 h to digest the DNA and RNA and centrifuged for 1 h at 4500 ***g*** and 4 °C. The Mb proteins were purified from the supernatant on a His60-Ni-column (Qiagen) according to the manufacturer's instructions. The histidine-tag was removed by incubation with the Factor Xa protease (20 $\mu g \cdot mL^{-1}$) for 6 h at 37 °C. After inactivation of the protease with 2 µM dansyl-Glu-Gly-Arg-chloromethyl-ketone, the recombinant Mb proteins were transferred to 10 mM HEPES, pH 7.8, 0.5 mM EDTA and stored at 4 °C until further analysis.

O_2 equilibrium measurements

A modified diffusion chamber technique was used to determine the O_2 equilibrium curves as described earlier [29,58–60]. Samples of 5 µL (~ 0.2 mM haem) in 0.1 M HEPES, 0.5 mM EDTA, pH 7.2 at 20 °C were measured in at least duplicates. Before measuring, the ferric lungfish Mbs were reduced to the ferrous form under N_2 for 10 min at room temperature using a met-Hb reductase system [61,62] consisting of glucose 6-phosphate (677 µM), glucose 6-phosphate dehydrogenase (0.33 $\mu g \cdot mL^{-1}$), NADP (45 µM), ferredoxin (0.17 µM), ferredoxin-NADP reductase (0.77 µM) and catalase (7 nM). All chemicals were supplied by Sigma-Aldrich (Brøndbyvester, Denmark).

The GMS 500 gas mixing system (Loligo System, Viborg, Denmark) generated a water-saturated gas mixture of ultrapure (> 99.998%) N_2 and O_2, which was further used to equilibrate the thin smear of Mb solution by stepwise increasing of the O_2 tension (PO_2). A photomultiplier (model RCA 931-A) and an Eppendorf model 1100 M photometer recorded the absorbance changes continuously at 436 nm to derive saturation at each PO_2 step. The absorption signal was recorded and analysed using the in-house developed software (Specktrosampler, K. Beedholm, Aarhus Univ. Denmark) [63]. At the beginning and the end of each experiment, 0% and 100% O_2 saturation levels were obtained as reference. P_{50} and cooperativity values were calculated from the zero intercept ($\log(Y/(1-Y)) = 0$) and slope of Hill plots, respectively: $\log(Y/(1-Y))$ versus $\log PO_2$, where Y is the fractional saturation. Four to six saturation steps were taken for each curve.

Gel filtration experiments

The apparent molecular weight of PanMbs was determined by gel filtration using an Äkta Pure chromatographic system (GE Healthcare, Freiburg, Germany) together with a Superdex 75 10/300 (GE Healthcare) column. A flow rate of 1 $mL \cdot min^{-1}$ and an elution buffer containing 50 mM potassium phosphate, 0.5 mM EDTA, pH 7.0, 0.15 M NaCl were applied. The absorbance was measured at 280 nm (proteins) and 415 nm (haem-group). Sample solutions (100 µL each) with a concentration of 0.2 mM haem were loaded on the system; human Hb and horse Mb were used as standard at the same concentration and volume as the samples.

Nitrite reductase activity

After complete deoxygenation of the PanMbs (10 µM haem in deoxygenated 50 mM HEPES, pH 7.13) by anaerobically titrating with 1 mM of sodium dithionite in a 1 cm sealed cuvette, sodium nitrite (0.1 $mM \cdot cuvette^{-1}$) of freshly made stock solutions was added anaerobically with a Hamilton

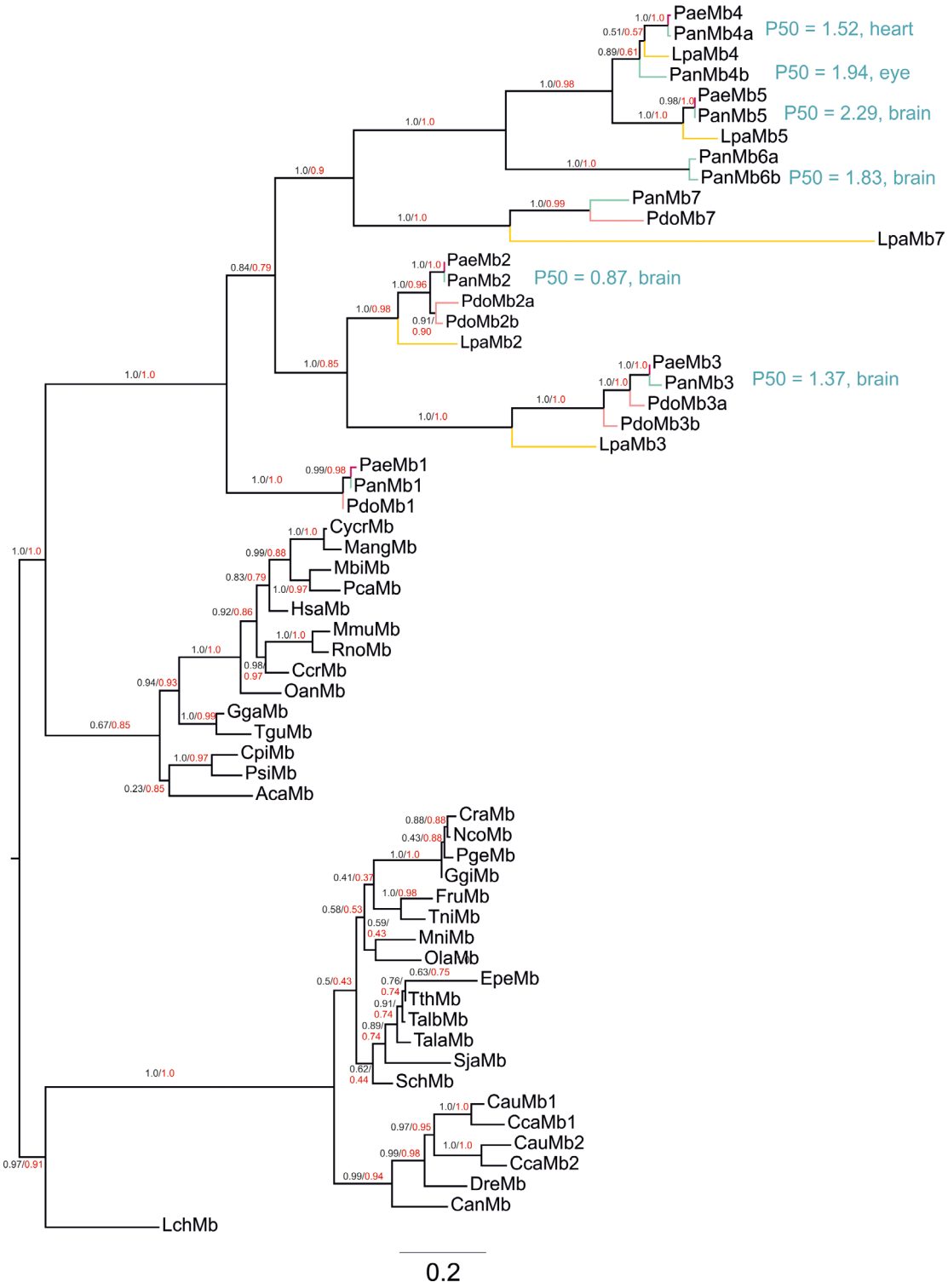

Fig. 1. Bayesian phylogenetic tree of vertebrate Mbs. Tree reconstruction was carried out with the amino sequences assuming the JTT model of protein evolution [60]. The bar represents 0.2 substitutions per site. The numbers at the nodes are the Bayesian posterior probabilities (black) and the ultrafast bootstrap support (red). For the abbreviations and accession numbers, see Table S1. For the PanMbs, the P_{50} values (in Torr) and the main expression sites are given. (blue: *Protopterus annectens* Mbs; pink: *Protopterus aethiopicus* Mbs; light red: *Protopterus dolloi* Mbs; yellow: *Lepidosiren paradoxa* Mbs).

gastight syringe (Bonaduz, Switzerland) [64]. The reaction was carried out at 20 °C and under pseudo-first-order conditions, that is with substrate nitrite in excess of protein. The measurement of the reaction kinetics was started immediately, and absorbance was recorded at different time intervals at 435 nm and 20 °C.

Dynamic light scattering

Dynamic light scattering (DLS) measurements were performed to verify the oligomeric state and solution homogeneity of the PanMbs. Sample solutions (15 μL) were exposed to a 660 nm diode laser in a quartz glass cuvette inside a Spectrolight 300 instrument (XtalConcepts, Hamburg, Germany) at 20 °C for 100 s. The autocorrelation function of scattered light was evaluated using the CONTIN algorithm to determine diffusion constants and finally the hydrodynamic radius (Rh) distribution of the protein solutions.

CD spectroscopy

The secondary structure was investigated by comparing the far-UV CD spectra of the respective PanMbs. Protein samples were prepared in 10 mM Tris/HCl inside a 1 mm path length glass cuvette, which was positioned inside a Jasco J-815 spectropolarimeter (Jasco, Easton, MD, USA). For each sample, 10 spectra were recorded and averaged with the subsequent subtraction of the averaged ellipticity of the buffer solution. Experiments were performed at 19 °C with an increment of 0.1 nm per data point and a scanning speed of 100 nm·min^{-1}. A far-UV CD spectrum of horse Mb (ID: CD0000047000) was obtained from the Protein CD Data Bank [65].

Small-angle X-ray scattering

Homogeneous monodisperse solutions of histidine-tagged PanMbs were applied to small-angle X-ray scattering (SAXS) experiments to analyse the tertiary structure envelope of selected PanMbs in solution. The experiments were performed at the EMBL beamline P12 [66] of the PETRA III synchrotron radiation source (DESY, Hamburg, Germany) applying a wavelength of 0.124 nm at a sample to detector distance of 3.1 m. The data were recorded in a momentum transfer range of 0.03–4.80 nm^{-1} with 20 exposures for 0.05 s per sample and at a temperature of 14 °C. The protein concentration range was 0.4–3.5 mg·mL^{-1}. The SASFLOW data processing pipeline [67], as well as PRIMUS[68], was applied for primary data quality verification and reduction, to exclude radiation damage as well as averaging of the scattering amplitudes of the individual X-ray exposures. Subsequently, the averaged scattering intensity of the buffer solution, which was exposed before and after each protein, was subtracted, and samples were scaled relative to the protein concentration. Radii of gyration (Rg) were derived from both the Guinier plot and the pair-distance distribution function using the PRIMUS user interface including AUTOGNOM. The pair-distance distribution function was used as an input for further *ab initio* modelling. Herein GASBOR [69] was applied to calculate chain-like assemblies of dummy amino acid residues with P1 symmetry to visualise the protein shape.

Results

Identification of myoglobin cDNA in four lungfish species

Assemblies of the transcriptomes of *P. dolloi*, *P. aethiopicus*, *P. annectens* and *L. paradoxa* (Table S1) revealed multiple Mb copies in all four lungfish species (Table S2). Eight Mb cDNA sequences had been identified before in *P. annectens* [5], and a ninth PanMb cDNA sequence was found in the transcriptomes that have recently become available (Table S1). The Mb genes were named according to their orthology and phylogenetic position in the tree (Fig. 1), which required renumbering of two previously described *P. annectens* Mb genes (PanMb4 ⇒ PanMb4a and PanMb7 ⇒ PanMb4b). The newly identified PanMb cDNA sequence was named PanMb7. The main expression site of the Mbs in *P. annectens* was the brain, which agrees with the previous results [5] (Fig. 2A; Table S3).

The transcriptomes of the slender lungfish *P. dolloi* harbour six Mbs (PdoMbs), which were named by orthology to the PanMbs PdoMb1, PdoMb2a, PdoMb2b PdoMb3a, PdoMb3b and PdoMb7. The PdoMbs share between 54.6% and 94% sequence identity on nucleotide level and 36.9–93.3% identity on the protein level. The available *P. dolloi* transcriptomes derive from pelvic and pectoral fins, and nasal lymphoid aggregates. The two fin tissues, which presumably mainly consist of bones and muscles, have very similar patterns and show that PdoMb7 is the strongly expressed Mb in these samples (Fig. 2B; Table S4). In the marbled lungfish *P. aethiopicus*, five Mbs were identified (PaeMb1–5) that share 44.9–61% of the nucleotides and 29.7–66.7% of the amino acids. The expression of the PaeMbs was evaluated by RNA-Seq, which showed the highest mRNA levels for PaeMb2 in the developing jaw (Fig. 2C; Table S5). The transcriptomes of *L. paradoxa* revealed five distinct Mb sequences, named LpaMb2, LpaMb3, LpaMb4, LpaMb5 and LpaMb7, respectively. They have a sequence identity of 46.2–74.2% on the nucleotide and 28.2–72.7% on the amino acid level. Like in *P. annectens*, the highest Mb expression levels were observed in the brain. Also, like PanMb2, LpaMb2

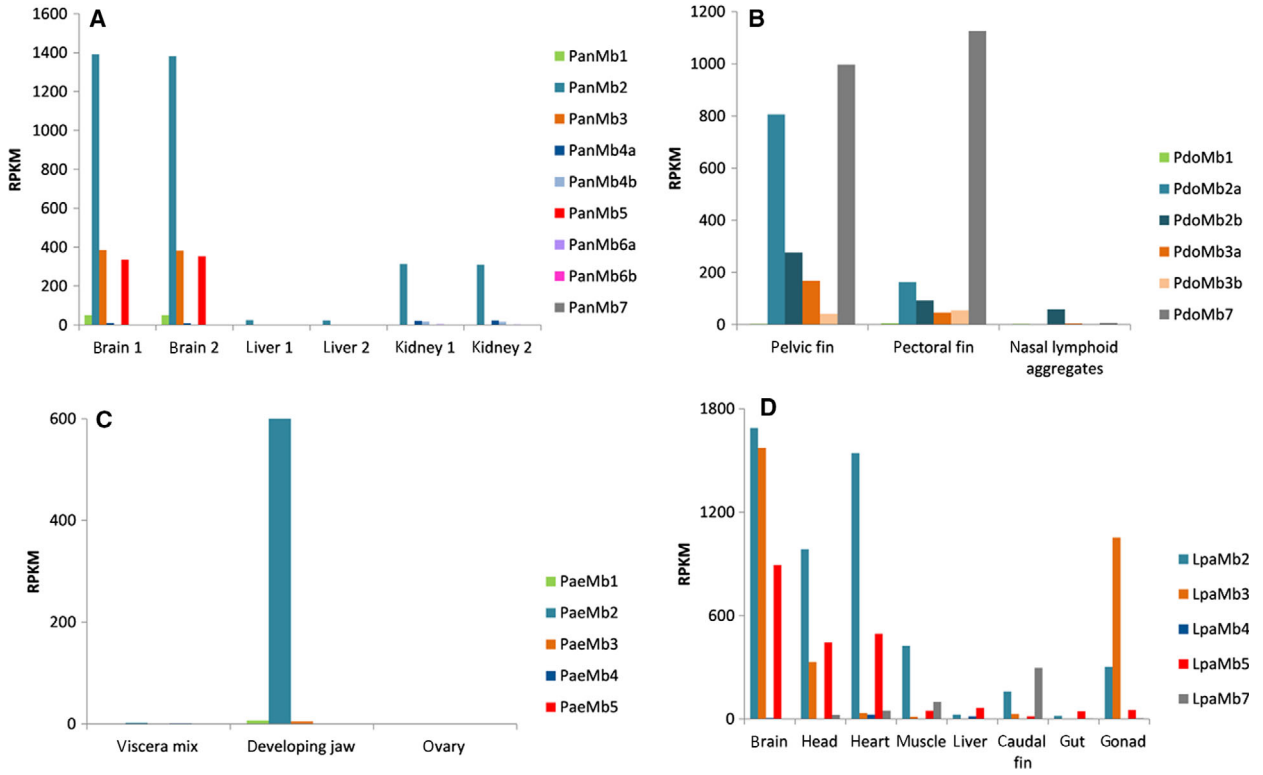

Fig. 2. Expression of the Mb genes in selected *Protopterus annectens* (A), *Protopterus dolloi* (B), *Protopterus aethiopicus* (C) and *Lepidosiren paradoxa* (D) tissues. The mRNA levels were estimated by RNA-Seq and are displayed as RPKM values. Transcriptome accession numbers are given in Table S1. The copy numbers are given in Tables S3–S6.

had the highest mRNA level (Fig. 2D; Table S6). In none of the lungfish transcriptomes, cDNA sequences that represent Ngb or Cygb were detected. Comparisons between orthologs showed mostly high conservation (e.g. 80–88% identity between *P. annectens* and *L. paradoxa* Mb orthologs), with the exception of the Mb7 proteins (28% identity of PanMb7 and LpaMb7). Phylogenetic analysis shows that all lungfish Mbs form a single clade, which is in sister group position to the tetrapod Mbs [supported with 1.00 Bayesian posterior probability (BPP); Fig. 1]. This clade is the sister group of the coelacanth Mb, and the monophyly of the sarcopterygian Mbs is supported with 1.0 BPP. Within the lungfish, distinct Mbs from different species form common clades, giving rise to orthologous Mb genes in the different species. Some Mbs are apparently absent in certain species (such as Mb1 in *L. paradoxa*), others as far as the transcriptome data indicate are in-paralog that have duplicated within a species (thus named, e.g. PanMb6a and 6b, or PdoMb3a and 3b). PanMb4b may be an out-paralog that has been lost in other lungfish species, but the support values within this clade are comparably low.

O_2 binding equilibria of *P. annectens* myoglobins

The Mbs of *P. annectens*, PanMb1, 2, 3, 4a, 4b, 5 and 6b, were successfully recombinantly expressed applying the pET vector system. PanMb6a could not be cloned, possibly due to a low expression level and/or high similarity to PanMb6b. PanMb7 could also not be cloned.

The O_2 equilibrium curves of all seven PanMbs were hyperbolic and showed that all PanMbs reversibly bind oxygen (Fig. 3). However, the O_2 affinities (P_{50}) and cooperativity values (n) were quite distinct (Table 1). Within the given parameters, PanMb2 had the highest O_2 affinity ($P_{50} = 0.88 \pm 0.13$ Torr), followed by PanMb3 ($P_{50} = 1.37 \pm 0.04$ Torr), PanMb4a ($P_{50} = 1.52 \pm 0.10$ Torr), PanMb6b ($P_{50} = 1.83 \pm 0.07$ Torr), PanMb4b ($P_{50} = 1.94 \pm 0.30$ Torr), PanMb5 ($P_{50} = 2.29 \pm 0.08$ Torr) and PanMb1 ($P_{50} = 3.78 \pm 0.01$ Torr) with the lowest affinity (Fig. 3). Hill plots show slight cooperativity ($n > 1$) for all PanMbs, suggesting an oligomeric assembly (Fig. 3B): PanMb1 ($n = 1.63 \pm 0.15$), PanMb2 ($n = 1.35 \pm 0.18$), PanMb3 ($n = 1.57 \pm 0.27$), PanMb4a ($n = 1.59 \pm 0.12$), PanMb4b ($n = 1.64 \pm 0.07$), PanMb5 ($n = 1.63 \pm 0.06$) and PanMb6 ($n = 1.33 \pm 0.17$).

Enzymatic activities of *P. annectens* myoglobins

The nitrite reductase activity of the deoxy PanMbs was measured after addition of 0.1 mM nitrite. The reactions took place at different observed rates (Fig. 4, Table 1, Table S7), with different derived second-order rate constants (expressed as $M^{-1}{\cdot}s^{-1}$). For PanMb3 (42.13 ± 0.0002 $M^{-1}{\cdot}s^{-1}$) and PanMb4b (33.63 ± 0.0025 $M^{-1}{\cdot}s^{-1}$), the nitrite reductase activity was high, and the nitrite reductase rates of the other Mbs were lower: PanMb1 (26.95 ± 0.005 $M^{-1}{\cdot}s^{-1}$), PanMb2 (13.78 ± 0.0007 $M^{-1}{\cdot}s^{-1}$), PanMb4a (22.91 ± 0.0018 $M^{-1}{\cdot}s^{-1}$) and PanMb5 (10.26 ± 0.0032 $M^{-1}{\cdot}s^{-1}$). PanMb6b showed no apparent nitrite reductase activity.

Structural analyses of *P. annectens* myoglobins

Gel filtration of the oxy-PanMbs showed approximately the same molecular mass for all PanMbs (Fig. 5). These results were confirmed by SDS/PAGE (Fig. 5). The absorbance peaks of the respective chromatograms also indicate that all PanMbs elute mainly as monomers. An additional small peak in PanMb1, PanMb2 and PanMb6b suggests minor dimeric fractions in these recombinant proteins.

Dynamic light scattering experiments showed that all seven analysed PanMbs have approximately the same Rh (Table 2), which is comparable to Rh of human Mb [70]. SAXS measurements revealed very similar Rg and the typical rigid and compact Mb fold for the investigated Mbs with only minor deviation. *Ab initio* models of the PanMbs based on solution scattering are well superimposable with the respective *in silico* models calculated via SWISS-MODEL [71,72] using default modelling parameters (Fig. 6). Further, the respective CD spectra were highly similar and indicated a conserved secondary structure composition compared to each other and to an Mb reference spectrum (around 70% of α-helical structure; Fig. 7). The secondary structure composition was also not significantly affected by the purification tag.

Discussion

Emergence of multiple myoglobins as a possible adaptation of lungfish

Lungfish (Dipnoi) share with the land-living vertebrates (tetrapods) the ability to breathe air. They first appear in the fossil record in the Devonian ~ 400 million years ago [73]. There are only six extant lungfish species that live in rivers and (seasonal) freshwater lakes in the tropics [4]. Four species of the genus *Protopterus* live in Africa (*P. aethiopicus*, *P. amphibius*, *P. annectens* and *P. dolloi*), *L. paradoxa* in South America and *Neoceratodus forsteri* in Australia. Although lungfish are often considered to be 'living fossils' [74], they also show some specific adaptations. For example, lungfish can withstand hypoxia and can survive certain periods outside the water. African lungfishes may also aestivate during the dry season in a mud cocoon [75]. During this period, lungfishes reduce their energy metabolism, increase urea production to reduce ammonia toxicity and turn on the expression of cytoprotective genes [76,77]. Lungfish also have the largest animal genomes, with C-values up to 132 pg per haploid genome [78].

As none of the other globins (with the exception of GbE; see [79]) or, to the best of our knowledge, any other gene family show signs of gene amplification, it is unlikely that the multiplication of the Mb genes is caused by hypothetic whole-genome amplifications putatively

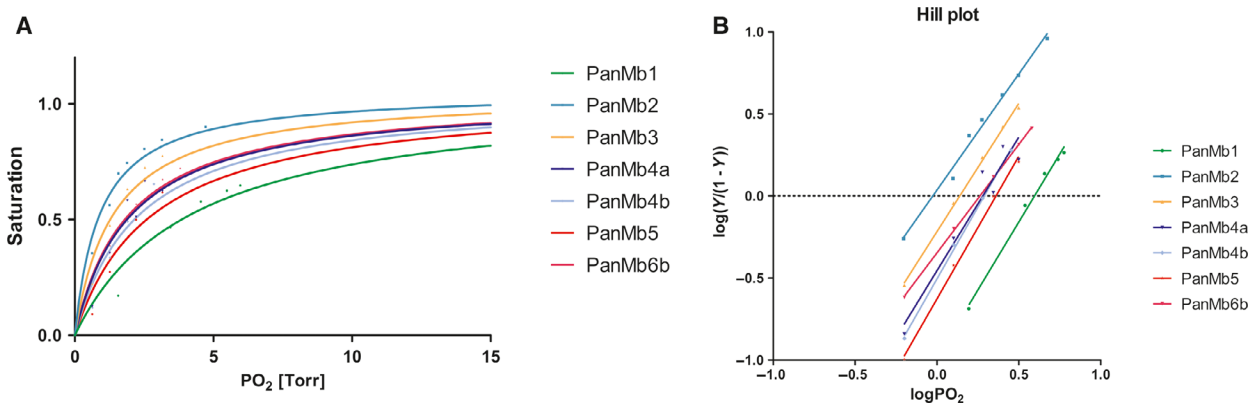

Fig. 3. Oxygen equilibrium curves of recombinant PanMbs, measured at pH 7.2, 20 °C. The sigmoidal curves (A) and linear form of the Hill plot (B) are depicted. Slight cooperativity is evidenced by the slope ($n > 1$) of the Hill plot.

Table 1. Summary of the results of the PanMbs.

	Main expression site	Coordination	P_{50}(Torr)	Hill Coeff	Nitrite reductase rate($s^{-1}\cdot M^{-1}$)
PanMb1	Muscle	Penta	3.78 ± 0.01	1.63 ± 0.15	26.95 ± 0.005
PanMb2	Brain/eye	Penta	0.88 ± 0.13	1.35 ± 0.18	13.78 ± 0.0007
PanMb3	Brain/eye	Penta	1.37 ± 0.04	1.57 ± 0.27	42.13 ± 0.0002
PanMb4a	Heart	penta	1.52 ± 0.10	1.59 ± 0.12	22.91 ± 0.0018
PanMb4b	Brain/eye	Penta	1.94 ± 0.30	1.64	33.63 ± 0.0025
PanMb5	Brain/eye	Penta	2.29 ± 0.08	1.63 ± 0.06	10.26 ± 0.0032
PanMb6b	Brain/eye	Penta	1.83 ± 0.07	1.33 ± 0.17	0.00

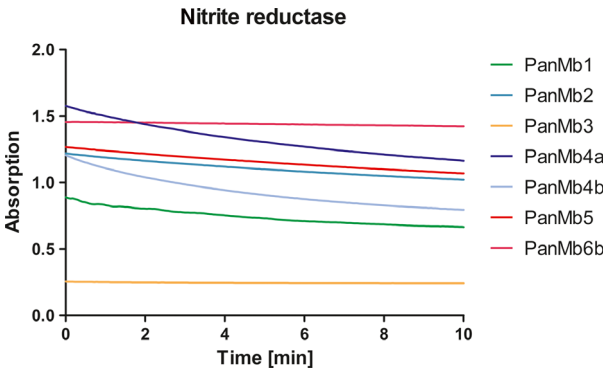

Fig. 4. Nitrite reductase activity of recombinant PanMbs, measured at pH 7.2, 20 °C, shows the decrease of the absorbance over time of the deoxy Soret peak.

linked to the large lungfish genomes. Duplicated and even triplicated *Mb* genes have been known from some teleost species [41–43,48,80]. Lungfish are the current record holders with up to nine distinct *Mbs*. While the duplicated and triplicated *Mb* genes of teleosts are relatively recent innovations in evolution, the amplification of Mb genes in lungfish must have occurred before the genera Lepidosiren and Protopterus split more than 100 million years ago [81]. In the future, the availability of genomic sequences of the Australian lungfish *N. forsteri* may push this date further back to the divergence time of Lepidosireniformes and Ceratodontiformes > 200 million years ago. In general, genomic data would help to confirm some of our assumptions. With the help of genomic analyses, we could explicitly attest the losses of Cygb and Ngb in lungfish.

The amplification of the different lungfish Mbs occurred in the lungfish stem lineage (Fig. 1), which might have been a response to a lifestyle with low and changing O_2 conditions [5] or other physiological challenges. It is noteworthy that most Mb orthologs have been maintained for > 100 million years, suggesting that the lungfish requires specific *Mbs* rather than just multiple copies of the gene. Only in four cases, we found additional paralogs (PdoMb2a and 2b, PdoMb3a and 3b, PanMb4a and 4b, and PanMb6a and 6b). The *Mb7* gene may have been lost in *P. aethiopicus*, but additional transcriptome or genome sequencing is required for verification. Our results strongly hint to subfunctionalisation or neofunctionalisation of the *Mb* genes during lungfish evolution, a hypothesis that is supported by their tissue-specific expression patterns [5], as well as the different kinetic and enzymatic properties of the recombinant proteins (see below).

Conserved structure of *P. annectens* myoglobins

Despite notable differences of the amino acid sequences (up to 71.1% difference), the secondary structures of the PanMbs are highly similar and show essentially the same high content of α-helices. The minor differences most likely result from, for example the N terminally shortened sequence of PanMb6b or slightly different surface charge distribution. As seen in the DLS measurements, all PanMbs have almost identical hydrodynamic radii, which fit a nearly globular shape of monomeric Mbs in solution (Table 2). This observation was verified by the SAXS data, which shows similarities to the tertiary structure of sperm whale Mb originally determined by Kendrew *et al.* [82]. Size-exclusion chromatography experiments confirmed these results and found mainly monomeric PanMb proteins, although for some a minor dimeric fraction was observed, which may explain the slightly cooperative O_2-binding kinetics. However, the SAXS data did not indicate a significant concentration-dependent oligomerisation of the applied monomeric proteins.

Functional differentiation of *P. annectens* myoglobins

The O_2 affinities of the PanMbs range between P_{50} values from 0.88 to 3.8 Torr. PanMb1 has the lowest O_2 affinity with P_{50} = 3.8 Torr, similar to the low affinity

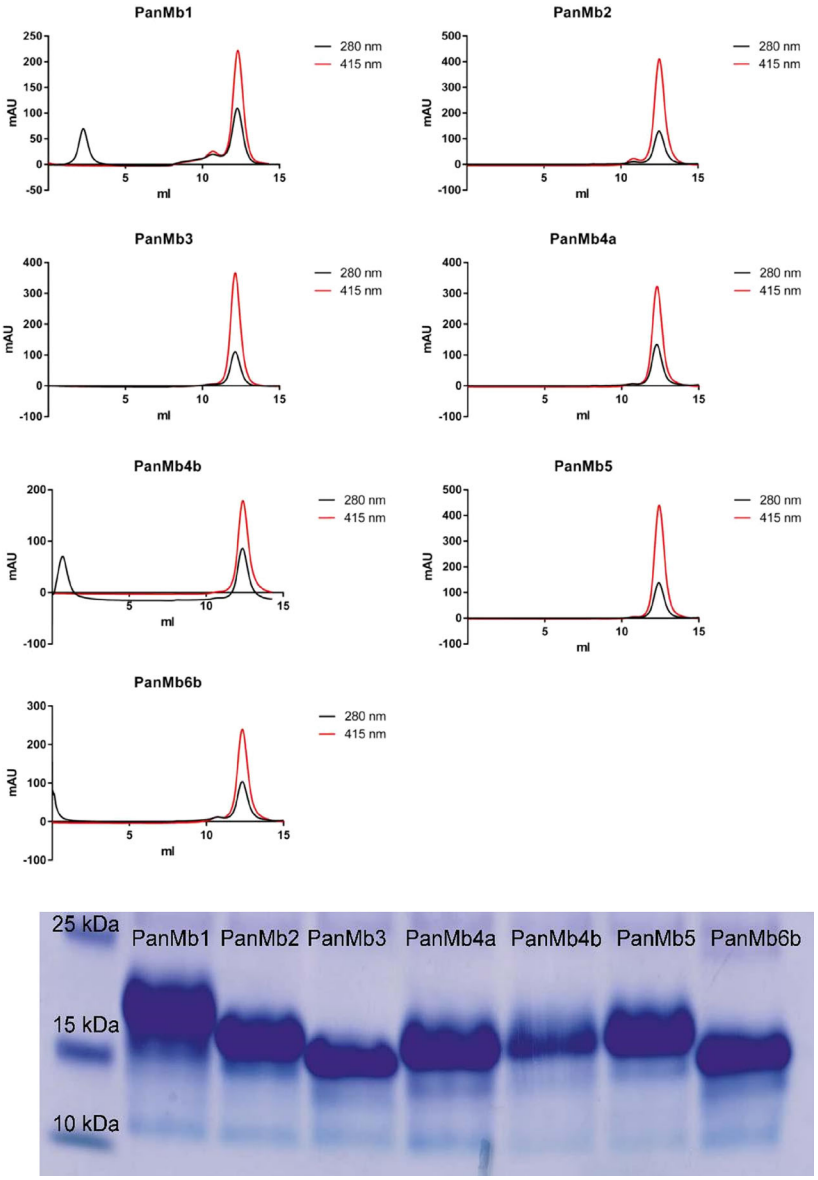

Fig. 5. FPLC diagrams of recombinant PanMbs. Protein absorbance was measured at 280 nm, haem-specific absorbance at 415 nm. SDS/PAGE gel electrophoresis of recombinantly expressed *Protopterus annectens* Mb1–6 proteins.

Table 2. Characteristic size parameters of PanMbs, as obtained by DLS and SAXS, respectively.

Protein	Rh (nm)	Rg (nm)[a]	D_{max} (nm)	Oligomeric state
PanMb1	2.2 ± 0.3	1.99 ± 0.04 (1.97 ± 0.01)	6.2	Monomer
PanMb2	2.1 ± 0.3	1.83 ± 0.08 (1.85 ± 0.02)	6.4	Monomer
PanMb4a	1.9 ± 0.3	2.13 ± 0.16 (2.13 ± 0.01)	6.7	Monomer
PanMb5	2.0 ± 0.3	1.86 ± 0.10 (1.85 ± 0.02)	5.6	Monomer
PanMb6b	1.9 ± 0.2	2.04 ± 0.20 (2.03 ± 0.01)	6.3	Monomer

[a] Calculated according to the Guinier approximation; the numbers in parentheses show Rg values according to the pair-distance distribution function.

of some fish Mb, such as mackerel Mb with P_{50} = 3.7 Torr at 25 °C [29] and trout Mb with P_{50} = 3.4–4.9 Torr at 25 °C [30,31]. PanMb1 is mainly expressed in the skeletal muscle and is also by far the most highly expressed Mb of this tissue [5]. The comparably low affinity suggests that PanMb1 is suitable

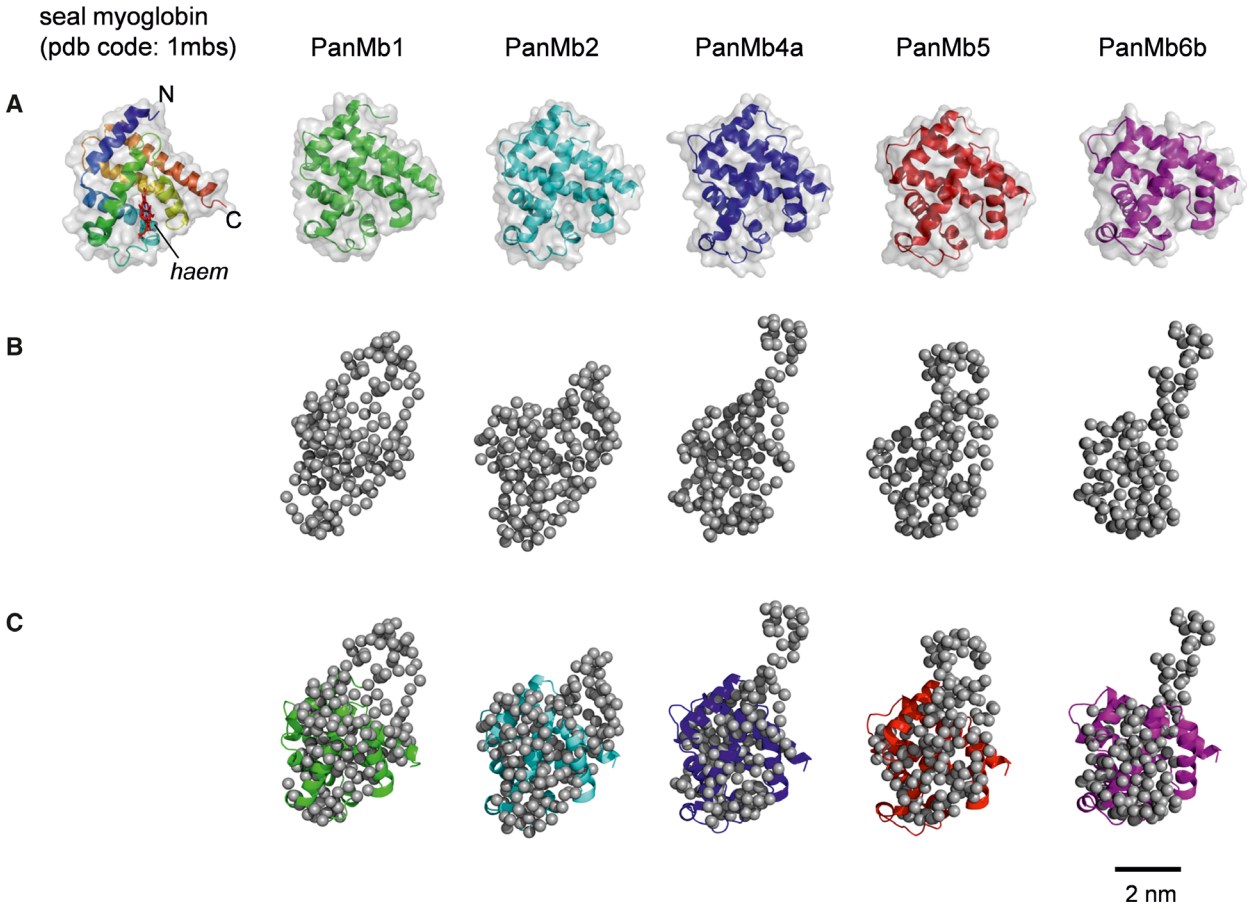

Fig. 6. Superimposition of *in silico* models calculated by SWISS-MODEL and Gasbor *ab initio* models. (A) Cartoon plot of PanMb1–6 *in silico* models in comparison with a Mb high-resolution X-ray structure possessing the conserved Mb fold. The surface representation is coloured light grey. (B) *Ab initio* models of C terminally tagged PanMbs calculated by GASBOR based on the respective P(R) function. A chain-like set of dummy spheres was used to fit the X-ray scattering intensity distribution. The GASBOR chi-square values of the fit function with the experimental data are 1.11, 1.07, 1.34, 0.99 and 1.03 in the sequence; the models are displayed from left to right. A globular core structure of the PanMbs in solution is conserved with a slightly more distinct structure of the respective termini. (C) Superimposition of both types of models displayed above. The scale bar is 2 nm. The Mb of the harbour seal (pdb code: 1mbs) was used as reference.

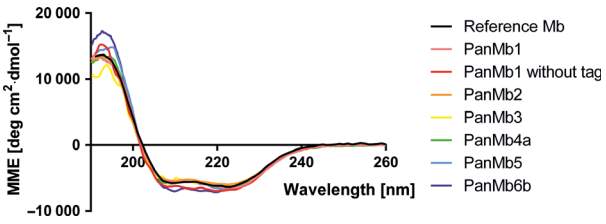

Fig. 7. CD spectra of recombinant PanMbs compared to a horse Mb as reference.

for efficient delivery of O_2 to the mitochondria. This finding correlates with the high efficiency of PanMb1 to enhance the activity of the mitochondrial dehydrogenases under hypoxia in cell culture [5]. Higher O_2 affinities were observed for the other PanMbs, with the brain-specific PanMb2 having the highest O_2 affinity (Table 1). The specific O_2 affinities of the different PanMbs probably reflect the individual requirements of the respective tissue. For example, the neurons may require an Mb with higher O_2 affinity than the skeletal muscle.

Deoxygenated globin proteins can reduce nitrite to NO [39,83]. All PanMbs except PanMb6b display noticeable nitrite reductase activity (Table 1). It is, therefore, conceivable that the high expression levels of Mb help to protect hypoxic tissues by increasing vasodilation and blood supply. Also, NO may inhibit mitochondrial respiration and limit the generation of ROS [24,38].

The lungfish Mbs support the idea that distinct evolutionary mechanisms can control the haem redox and the

Table 3. Amino acid structure of PanMbs at conserved globin positions.

	CD1	CD3	E7	E10	E11	F8
PanMb1	Phe	Lys	His	Thr	Val	His
PanMb2	Phe	Lys	His	Thr	Val	His
PanMb3	Phe	Lys	His	Thr	Val	His
PanMb4a	Phe	Lys	His	Val	Val	His
PanMb4b	Phe	Lys	His	Val	Val	His
PanMb5	Phe	Lys	His	Val	Val	His
PanMb6a	Phe	Lys	Gln	Val	Val	His
PanMb6b	Phe	Lys	Gln	Val	Val	His
PanMb7	Phe	Lys	His	Leu	Val	His

nonredox (O_2 binding) activities separately, depending on the need of the respective tissue or cell [84]. PanMb6a and PanMb6b are the only Mbs of the lungfish with an E7 glutamine substitution in the sequence (Table 3). The elephant Mb has the same substitution [85] and reacts with NO 500–1000 times faster than Mbs without substitution. Also, PanMb6b is the only one with no apparent nitrite reductase activity. Taken together, the E7 glutamine substitution is unlikely to be the only reason to explain these differences, but in case of the lungfish Mbs, it seems to be responsible for the negligible nitrite reductase enzymatic activity of deoxy PanMb6b.

Functional differentiation of lungfish myoglobin in evolution

Our results strongly suggest that the various lungfish Mbs carry out distinct biological functions. It is noteworthy that the muscle-specific Mb1 is in a sister group position to all other lungfish Mbs (Fig. 1). Thus, the original function of lungfish Mb probably was the supply of O_2 to the muscle tissue, the classic Mb-function found for most vertebrate Mbs. The other lungfish Mbs differentiated from that role after multiple gene duplications. Because the protein sequences of the other lungfish Mbs are highly similar to the sequences of the PanMbs, we suppose them having similar functions in the same tissues. On the one hand, an adaptation to a temporarily hypoxic environment may have led to the broad expression pattern of Mb in various tissues, thereby enhancing O_2 supply or reducing respiration rates via nitrite reduction to NO. All Mbs, in particular, PanMb3 and PanMb4, are highly efficient as nitrate reductases (see Table 1).

In comparison, vertebrate Mbs have a much lower nitrite reductase activity [24]. In addition, it is conceivable that the derived loss of Ngb and Cygb provided the impetus for the functional differentiation of lungfish Mbs, of which some might have taken over the yet still unclear roles of Ngb and Cygb. This hypothesis would also explain the preferential expression of some lungfish Mbs in the brain, which is also the main expression site of Ngb and Cygb [6,18]

Acknowledgements

The authors would like to thank Alexey Kikhney (EMBL Hamburg) for the assistance in operating the SAXS beamline P12 at the PETRA III storage ring and Elin E. Petersen (Aarhus) for skilled assistance in the laboratory. This work is supported by the Deutsche Forschungsgemeinschaft (BU956/19-1), CNPq Universal Program Grant 403248/2016-7 to I.S and Natur og Univers, Det Frie Forskningsråd (grant 4181-00094 to AF).

Conflict of interest

The authors declare no conflict of interest.

Author contributions

TB planned experiments; JL, AF, SF and MW performed experiments; JL, SF, IS, AF and TB analysed data; JL and TB wrote the paper.

References

1 Brinkmann H, Venkatesh B, Brenner S & Meyer A (2004) Nuclear protein-coding genes support lungfish and not the coelacanth as the closest living relatives of land vertebrates. *Proc Natl Acad Sci USA* **101**, 4900–4905.

2 Amemiya CT, Alfoldi J, Lee AP, Fan S, Philippe H, Maccallum I, Braasch I, Manousaki T, Schneider I, Rohner N *et al.* (2013) The African coelacanth genome provides insights into tetrapod evolution. *Nature* **496**, 311–316.

3 Biscotti MA, Gerdol M, Canapa A, Forconi M, Olmo E, Pallavicini A, Barucca M & Schartl M (2016) The lungfish transcriptome: a glimpse into molecular evolution events at the transition from water to land. *Sci Rep.* **6**, 21571.

4 Jørgensen JM & Joss J (2010) The Biology of Lungfishes. Science Publishers, Enfield, CT.

5 Koch J, Ludemann J, Spies R, Last M, Amemiya CT & Burmester T (2016) Unusual diversity of myoglobin genes in the lungfish. *Mol Biol Evol* **33**, 3033–3041.

6 Burmester T & Hankeln T (2014) Function and evolution of vertebrate globins. *Acta Physiol (Oxf)* **211**, 501–514.

7 Hardison RC (1996) A brief history of hemoglobins: plant, animal, protist, and bacteria. *Proc Natl Acad Sci USA* **93**, 5675–5679.

8 Gillemans N, McMorrow T, Tewari R, Wai AW, Burgtorf C, Drabek D, Ventress N, Langeveld A, Higgs D, Tan-Un K *et al.* (2003) Functional and comparative analysis of globin loci in pufferfish and humans. *Blood* **101**, 2842–2849.

9 Graur D & Li WH (2000) Fundamentals of Molecular Evolution. Sinauer Associates, Incorporated, Sunderland, MA.

10 Vinogradov SN, Hoogewijs D, Bailly X, Mizuguchi K, Dewilde S, Moens L & Vanfleteren JR (2007) A model of globin evolution. *Gene* **398**, 132–142.

11 Burmester T, Weich B, Reinhardt S & Hankeln T (2000) A vertebrate globin expressed in the brain. *Nature* **407**, 520–523.

12 Kawada N, Kristensen DB, 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. *J Biol Chem* **276**, 25318–25323.

13 Burmester T, Ebner B, Weich B & Hankeln T (2002) Cytoglobin: a novel globin type ubiquitously expressed in vertebrate tissues. *Mol Biol Evol* **19**, 416–421.

14 Trent JT 3rd & Hargrove MS (2002) A ubiquitously expressed human hexacoordinate hemoglobin. *J Biol Chem* **277**, 19538–19545.

15 Kugelstadt D, Haberkamp M, Hankeln T & Burmester T (2004) Neuroglobin, cytoglobin, and a novel, eye-specific globin from chicken. *Biochem Biophys Res Commun* **325**, 719–725.

16 Blank M, Kiger L, Thielebein A, Gerlach F, Hankeln T, Marden MC & Burmester T (2011) Oxygen supply from the bird's eye perspective: globin E is a respiratory protein in the chicken retina. *J Biol Chem*. **286**, 26507–26515.

17 Hoogewijs D, Ebner B, Germani F, Hoffmann FG, Fabrizius A, Moens L, Burmester T, Dewilde S, Storz JF, Vinogradov SN *et al.* (2012) Androglobin: a chimeric globin in metazoans that is preferentially expressed in Mammalian testes. *Mol Biol Evol* **29**, 1105–1114.

18 Hankeln T, Ebner B, Fuchs C, Gerlach F, Haberkamp M, Laufs TL, Roesner A, Schmidt M, Weich B, Wystub S *et al.* (2005) Neuroglobin and cytoglobin in search of their role in the vertebrate globin family. *J Inorg Biochem* **99**, 110–119.

19 Burmester T & Hankeln T (2009) What is the function of neuroglobin? *J Exp Biol* **212**, 1423–1428.

20 Kendrew JC, Bodo G, Dintzis HM, Parrish RG, Wyckoff H & Phillips DC (1958) A three-dimensional model of the myoglobin molecule obtained by X-ray analysis. *Nature* **181**, 662–666.

21 Kendrew JC (1963) Myoglobin and the structure of proteins. *Science* **139**, 1259–1266.

22 Merx MW, Flogel U, Stumpe T, Godecke A, Decking UK & Schrader J (2001) Myoglobin facilitates oxygen diffusion. *FASEB J* **15**, 1077–1079.

23 Wittenberg JB & Wittenberg BA (2003) Myoglobin function reassessed. *J Exp Biol* **206**, 2011–2020.

24 Helbo S, Weber RE & Fago A (2013) Expression patterns and adaptive functional diversity of vertebrate myoglobins. *Biochim Biophys Acta* **1834**, 1832–1839.

25 Antonini E & Brunori M (1971) Hemoglobin and Myoglobin in Their Reactions with Ligands. Frontiers of Biology, North-Holland, Amsterdam.

26 Enoki Y, Matsumura K, Ohga Y, Kohzuki H & Hattori M (1995) Oxygen affinities (P_{50}) of myoglobins from four vertebrate species (Canis familiaris, Rattus norvegicus, Mus musculus and Gallus domesticus) as determined by a kinetic and an equilibrium method. *Comp Biochem Physiol B Biochem Mol Biol* **110**, 193–199.

27 Enoki Y, Ohga Y, Ishidate H & Morimoto T (2008) Primary structure of myoglobins from 31 species of birds. *Comp Biochem Physiol B Biochem Mol Biol* **149**, 11–21.

28 Helbo S, Dewilde S, Williams DR, Berghmans H, Berenbrink M, Cossins AR & Fago A (2012) Functional differentiation of myoglobin isoforms in hypoxia-tolerant carp indicates tissue-specific protective roles. *Am J Physiol Regul Integr Comp Physiol* **302**, R693–R701.

29 Madden PW, Babcock MJ, Vayda ME & Cashon RE (2004) Structural and kinetic characterization of myoglobins from eurythermal and stenothermal fish species. *Comp Biochem Physiol B Biochem Mol Biol* **137**, 341–350.

30 Helbo S & Fago A (2011) Allosteric modulation by S-nitrosation in the low-O(2) affinity myoglobin from rainbow trout. *Am J Physiol Regul Integr Comp Physiol* **300**, R101–R108.

31 Pedersen CL, Faggiano S, Helbo S, Gesser H & Fago A (2010) Roles of nitric oxide, nitrite and myoglobin on myocardial efficiency in trout (*Oncorhynchus mykiss*) and goldfish (Carassius auratus): implications for hypoxia tolerance. *J Exp Biol* **213**, 2755–2762.

32 Schmidt-Nielsen K (1990) Animal Physiology: Adaptation and Environment, 4th edn. Cambridge University Press, Cambridge.

33 Flögel U, Fago A & Rassaf T (2010) Keeping the heart in balance: the functional interactions of myoglobin with nitrogen oxides. *J Exp Biol* **213**, 2726–2733.

34 Moncada S & Erusalimsky JD (2002) Does nitric oxide modulate mitochondrial energy generation and apoptosis? *Nat Rev Mol Cell Biol* **3**, 214–220.

35 George P & Irvine DH (1951) Reaction of methmyoglobin with hydrogen peroxide. *Nature* **168**, 164–165.

36 Osawa Y & Korzekwa K (1991) Oxidative modification by low levels of HOOH can transform myoglobin to an oxidase. *Proc Natl Acad Sci USA* **88**, 7081–7085.

37 Flögel U, Gödecke A, Klotz LO & Schrader J (2004) Role of myoglobin in the antioxidant defense of the heart. *FASEB J* **18**, 1156–1158.

38 Hendgen-Cotta UB, Merx MW, Shiva S, Schmitz J, Becher S, Klare JP, Steinhoff HJ, Goedecke A, Schrader J, Gladwin MT *et al.* (2008) Nitrite reductase activity of myoglobin regulates respiration and cellular viability in myocardial ischemia-reperfusion injury. *Proc Natl Acad Sci USA* **105**, 10256–10261.
39 Fago A (2017) Functional roles of globin proteins in hypoxia-tolerant ectothermic vertebrates. *J Appl Physiol (1985)* **123**, 926–934.
40 Shih L, Chung Y, Sriram R & Jue T (2014) Palmitate interaction with physiological states of myoglobin. *Biochim Biophys Acta* **1840**, 656–666.
41 Fraser J, de Mello LV, Ward D, Rees HH, Williams DR, Fang Y, Brass A, Gracey AY & Cossins AR (2006) Hypoxia-inducible myoglobin expression in nonmuscle tissues. *Proc Natl Acad Sci USA* **103**, 2977–2981.
42 Roesner A, Mitz SA, Hankeln T & Burmester T (2008) Globins and hypoxia adaptation in the goldfish, *Carassius auratus*. *FEBS J* **275**, 3633–3643.
43 Cossins AR, Williams DR, Foulkes NS, Berenbrink M & Kipar A (2009) Diverse cell-specific expression of myoglobin isoforms in brain, kidney, gill and liver of the hypoxia-tolerant carp and zebrafish. *J Exp Biol* **212**, 627–638.
44 Maeda N & Fitch WM (1982) Isolation and amino acid sequence of a monomeric hemoglobin in heart muscle of the bullfrog, *Rana catesbeiana*. *J Biol Chem* **257**, 2806–2815.
45 Fuchs C, Burmester T & Hankeln T (2006) The amphibian globin gene repertoire as revealed by the Xenopus genome. *Cytogenet Genome Res* **112**, 296–306.
46 Sidell BD & O'Brien KM (2006) When bad things happen to good fish: the loss of hemoglobin and myoglobin expression in Antarctic icefishes. *J Exp Biol* **209**, 1791–1802.
47 Hoffmann FG, Opazo JC & Storz JF (2011) Differential loss and retention of cytoglobin, myoglobin, and globin-E during the radiation of vertebrates. *Genome Biol Evol* **3**, 588–600.
48 Macqueen DJ, Garcia de la serrana D & Johnston IA (2014) Cardiac myoglobin deficit has evolved repeatedly in teleost fishes. *Biol Lett* **10**, 20140225.
49 Katoh K & Toh H (2008) Recent developments in the MAFFT multiple sequence alignment program. *Brief Bioinform* **9**, 286–298.
50 Katoh K, Asimenos G & Toh H (2009) Multiple alignment of DNA sequences with MAFFT. *Methods Mol Biol* **537**, 39–64.
51 Ayres DL, Darling A, Zwickl DJ, Beerli P, Holder MT, Lewis PO, Huelsenbeck JP, Ronquist F, Swofford DL, Cummings MP *et al.* (2012) BEAGLE: an application programming interface and high-performance computing library for statistical phylogenetics. *Syst Biol* **61**, 170–173.
52 Huelsenbeck JP & Ronquist F (2001) MRBAYES: Bayesian inference of phylogenetic trees. *Bioinformatics* **17**, 754–755.
53 Jones DT, Taylor WR & Thornton JM (1992) The rapid generation of mutation data matrices from protein sequences. *Comput Appl Biosci* **8**, 275–282.
54 Katoh K, Kuma K, Toh H & Miyata T (2005) MAFFT version 5: improvement in accuracy of multiple sequence alignment. *Nucleic Acids Res* **33**, 511–518.
55 Nguyen LT, Schmidt HA, von Haeseler A & Minh BQ (2015) IQ-TREE: a fast and effective stochastic algorithm for estimating maximum-likelihood phylogenies. *Mol Biol Evol* **32**, 268–274.
56 Hoang DT, Chernomor O, von Haeseler A, Minh BQ & Vinh LS (2018) Modeling site heterogeneity with posterior mean site frequency profiles accelerates accurate phylogenomic estimation. *Syst Biol* **67**, 216–235.
57 Trifinopoulos J, Nguyen LT, von Haeseler A & Minh BQ (2016) W-IQ-TREE: a fast online phylogenetic tool for maximum likelihood analysis. *Nucleic Acids Res* **44**, W232–W235.
58 Weber RE (1981) Cationic control of O_2 affinity in lugworm erythrocruorin. *Nature* **292**, 386.
59 Weber RE (1992) Use of ionic and zwitterionic (Tris/BisTris and HEPES) buffers in studies on hemoglobin function. *J Appl Physiol 1985* **72**, 1611–1615.
60 Sick H & Gersonde K (1969) Method for continuous registration of O_2-binding curves of hemoproteins by means of a diffusion chamber. *Anal Biochem* **32**, 362–376.
61 Hayashi A, Suzuki T & Shin M (1973) An enzymic reduction system for metmyoglobin and methemoglobin, and its application to functional studies of oxygen carriers. *Biochim Biophys Acta* **310**, 309–316.
62 Fago A, Hundahl C, Dewilde S, Gilany K, Moens L & Weber RE (2004) Allosteric regulation and temperature dependence of oxygen binding in human neuroglobin and cytoglobin. Molecular mechanisms and physiological significance. *J Biol Chem* **279**, 44417–44426.
63 Jendroszek A, Malte H, Overgaard CB, Beedholm K, Natarajan C, Weber RE, Storz JF & Fago A (2018) Allosteric mechanisms underlying the adaptive increase in hemoglobin-oxygen affinity of the bar-headed goose. *J Exp Biol* **221**(Pt 18).
64 Fago A, Parraga DG, Petersen EE, Kristensen N, Giouri L & Jensen FB (2017) A comparison of blood nitric oxide metabolites and hemoglobin functional properties among diving mammals. *Comp Biochem Physiol A Mol Integr Physiol* **205**, 35–40.
65 Whitmore L, Miles AJ, Mavridis L, Janes RW & Wallace BA (2017) PCDDB: new developments at the

protein circular dichroism data bank. *Nucleic Acids Res* **45**, D303–D307.
66 Blanchet CE, Spilotros A, Schwemmer F, Graewert MA, Kikhney A, Jeffries CM, Franke D, Mark D, Zengerle R, Cipriani F *et al*. (2015) Versatile sample environments and automation for biological solution X-ray scattering experiments at the P12 beamline (PETRA III, DESY). *J Appl Crystallogr* **48**, 431–443.
67 Franke D, Kikhney AG & Svergun DI (2012) Automated acquisition and analysis of small angle X-ray scattering data. *Nucl Instrum Meth A* **689**, 52–59.
68 Petoukhov MV, Franke D, Shkumatov AV, Tria G, Kikhney AG, Gajda M, Gorba C, Mertens HDT, Konarev PV & Svergun DI (2012) New developments in the ATSAS program package for small-angle scattering data analysis. *J Appl Cryst* **45**, 342–350. https://doi.org/10.1107/S0021889812007662.
69 Svergun DI, Petoukhov MV & Koch MH (2001) Determination of domain structure of proteins from X-ray solution scattering. *Biophys J* **80**, 2946–2953.
70 Fedorov BA & Denesyuk AI (1978) Sperm whale myoglobin structure in solution differs from its structure in crystal by a shift of the 'hairpin' GH. *FEBS Lett* **88**, 114–117.
71 Biasini M, Bienert S, Waterhouse A, Arnold K, Studer G, Schmidt T, Kiefer F, Gallo Cassarino T, Bertoni M, Bordoli L *et al*. (2014) SWISS-MODEL: modelling protein tertiary and quaternary structure using evolutionary information. *Nucleic Acids Res* **42**, W252–W258.
72 Bienert S, Waterhouse A, de Beer TA, Tauriello G, Studer G, Bordoli L & Schwede T (2017) The SWISS-MODEL Repository-new features and functionality. *Nucleic Acids Res* **45**, D313–D319.
73 Clack JA (2011) The fossil record of lungfishes. In The Biology of Lungfishes (Jørgensen JM & Joss J, eds), pp. 1–42. Science Publishers, Enfield, CT.
74 Lee J, Alrubaian J & Dores RM (2006) Are lungfish living fossils? Observation on the evolution of the opioid/orphanin gene family. *Gen Comp Endocrinol* **148**, 306–314.
75 Ballantyne JS & Frick NT (2010) Lungfish metabolism. In The Biology of Lungfishes (Jørgensen JM & Joss J, eds), pp. 305–340. Science Publishers, New Ipswich, NH.
76 Chew SF, Chan NK, Loong AM, Hiong KC, Tam WL & Ip YK (2004) Nitrogen metabolism in the African lungfish (*Protopterus dolloi*) aestivating in a mucus cocoon on land. *J Exp Biol* **207**, 777–786.
77 Hiong KC, Ip YK, Wong WP & Chew SF (2013) Differential gene expression in the brain of the African lungfish, *Protopterus annectens*, after six days or six months of aestivation in air. *PLoS ONE* **8**, e71205.
78 Gregory TR. (2017) Animal Genome Size Database. http://www.genomesize.com
79 Lüdemann J, Verissimo KM, Dreger K, Fago A, Schneider I & Burmester T (2019) Globin E is a myoglobin-related, respiratory protein highly expressed in lungfish oocytes. *Sci Rep* **9**, 280.
80 Gallagher MD & Macqueen DJ (2017) Evolution and expression of tissue globins in ray-finned fishes. *Genome Biol Evol* **9**, 32–47.
81 Betancur RR, Orti G & Pyron RA (2015) Fossil-based comparative analyses reveal ancient marine ancestry erased by extinction in ray-finned fishes. *Ecol Lett* **18**, 441–450.
82 Kendrew JC, Dickerson RE, Strandberg BE, Hart RG, Davies DR, Phillips DC & Shore VC (1960) Structure of myoglobin: a three-dimensional Fourier synthesis at 2 A. resolution. *Nature* **185**, 422–427.
83 Tejero J & Gladwin MT (2014) The globin superfamily: functions in nitric oxide formation and decay. *Biol Chem* **395**, 631–639.
84 Fago A, Rohlfing K, Petersen EE, Jendroszek A & Burmester T (2018) Functional diversification of sea lamprey globins in evolution and development. *Biochim Biophys Acta* **1866**, 283–291.
85 Romero-Herrera AE, Goodman M, Dene H, Bartnicki DE & Mizukami H (1981) An exceptional amino acid replacement on the distal side of the iron atom in proboscidean myoglobin. *J Mol Evol* **17**, 140–147.

Supporting information

Additional supporting information may be found online in the Supporting Information section at the end of the article.

Table S1. List of lungfish SRA data sets used in this study.
Table S2. List of Mb sequences used in this study.
Table S3. RPKM of *Mb* genes in the transcriptomes of *Protopterus annectens* tissues used for expression estimation by RNA-Seq.
Table S4. RPKM of *Mb* genes in the transcriptomes of *Protopterus dolloi* tissues used for expression estimation by RNA-Seq.
Table S5. RPKM of *Mb* genes in the transcriptomes of *Protopterus aethiopicuss* tissues used for expression estimation by RNA-Seq.
Table S6. RPKM of *Mb* genes in the transcriptomes of *Lepidosiren paradoxa* tissues used for expression estimation by RNA-Seq.
Table S7. Overview of nitrite reductase observed rates (min-1 and s-1) and of second order rate constants ($M^{-1} \cdot s^{-1}$) of PanMbs. Conditions: 0.1 mM nitrite, pH 7.2, 20 °C.