



UNIVERSITY OF LEEDS

This is a repository copy of *Antarctic fish versus human cytoglobins – The same but yet so different*.

White Rose Research Online URL for this paper:  
<http://eprints.whiterose.ac.uk/120901/>

Version: Accepted Version

---

**Article:**

Cuypers, B, Vermeylen, S, Hammerschmid, D et al. (11 more authors) (2017) Antarctic fish versus human cytoglobins – The same but yet so different. *Journal of Inorganic Biochemistry*, 173. C. pp. 66-78. ISSN 0162-0134

<https://doi.org/10.1016/j.jinorgbio.2017.04.025>

---

© 2017 Elsevier Inc. This manuscript version is made available under the CC-BY-NC-ND 4.0 license <http://creativecommons.org/licenses/by-nc-nd/4.0/>

**Reuse**

Items deposited in White Rose Research Online are protected by copyright, with all rights reserved unless indicated otherwise. They may be downloaded and/or printed for private study, or other acts as permitted by national copyright laws. The publisher or other rights holders may allow further reproduction and re-use of the full text version. This is indicated by the licence information on the White Rose Research Online record for the item.

**Takedown**

If you consider content in White Rose Research Online to be in breach of UK law, please notify us by emailing [eprints@whiterose.ac.uk](mailto:eprints@whiterose.ac.uk) including the URL of the record and the reason for the withdrawal request.



[eprints@whiterose.ac.uk](mailto:eprints@whiterose.ac.uk)  
<https://eprints.whiterose.ac.uk/>

# Antarctic Fish versus Human Cytooglobins – The Same but yet so Different

Bert Cuypers<sup>a†</sup>, Stijn Vermeulen<sup>b†</sup>, Dietmar Hammerschmid<sup>b,c</sup>, Stanislav Trashin<sup>d</sup>, Vanoushe Rahemi<sup>d</sup>, Albert Konijnenberg<sup>d</sup>, Amy De Schutter<sup>a</sup>, C-H Christina Cheng<sup>e</sup>, Daniela Giordano<sup>f</sup>, Cinzia Verde<sup>f,g</sup>, Karolien De Wael<sup>d</sup>, Frank Sobott<sup>c</sup>, Sylvia Dewilde<sup>b</sup>, Sabine Van Doorslaer<sup>a\*</sup>

<sup>a</sup> BIMEF Laboratory, Department of Physics, University of Antwerp, Belgium

<sup>b</sup> PPES Laboratory, Department of Biomedical Sciences, University of Antwerp, Belgium

<sup>c</sup> BAMS Laboratory, Department of Chemistry, University of Antwerp, Belgium

<sup>d</sup> AXES Laboratory, Department of Chemistry, University of Antwerp, Belgium

<sup>e</sup> Department of Animal Biology, University of Illinois, United States of America

<sup>f</sup> Institute of Biosciences and BioResources, CNR, Naples, Italy

<sup>g</sup> Department of Biology, University Roma 3, Rome, Italy

<sup>†</sup> These authors contributed equally to this work

\* Corresponding author, e-mail: [sabine.vandoorslaer@uantwerpen.be](mailto:sabine.vandoorslaer@uantwerpen.be)

## ABSTRACT

The cytoglobins of the Antarctic fish *Chaenocephalus aceratus* and *Dissostichus mawsoni* have many features in common with human cytoglobin. These cytoglobins are heme proteins in which the ferric and ferrous forms have a characteristic hexacoordination of the heme iron, i.e. axial ligation of two endogenous histidine residues, as confirmed by electron paramagnetic resonance, resonance Raman and optical absorption spectroscopy. The combined spectroscopic analysis revealed only small variations in the heme-pocket structure, in line with the small variations observed for the redox potential. Nevertheless, some striking differences were also discovered. Resonance Raman showed that the stabilization of an exogenous heme ligand, such as CO, occurs differently in human cytoglobin than in the Antarctic fish cytoglobins. Furthermore, while it has been extensively reported that human cytoglobin is essentially monomeric and can form an intramolecular disulfide bridge that can influence the ligand binding kinetics, the 3D modeling of Antarctic fish cytoglobins indicates that the cysteine residues are too far apart to form such an intramolecular bridge. Moreover, gel filtration and mass spectrometry reveal the occurrence of covalent multimers (up to pentamers) in the Antarctic fish cytoglobins that are formed also at low concentrations. Stabilization of these oligomers by disulfide-bridge formation is possible, but not essential. We cannot exclude that if intermolecular disulfide bridges are formed, they can influence the heme-pocket structure.

## KEYWORDS

Cytoglobin; protein expression; electron paramagnetic resonance; resonance Raman spectroscopy; mass spectrometry; redox.

## 1 Introduction

Cytoglobin (Cygb) is a vertebrate globin with a currently unknown function that is ubiquitously expressed in different tissues in relatively low concentrations [1]. It is expressed in both the nucleus and the cytoplasm depending on the cell type [2]. Possible functions of Cygb include O<sub>2</sub> transport to the mitochondrial respiratory chain [1,3], O<sub>2</sub> storage [4], nitrite reductase activity [5], peroxidase activity [3], and a function as tumor suppressor gene [6]. Cygb has been reported to contribute to hypoxia adaptation in the subterranean mole rat *Spalax* [7]. Moreover, it has been hypothesized that Cygb protects cells against oxidative stress [8] and plays a role in fibrotic organ disorders [9-14]. The possibility of using Cygb as a biomarker has also recently been evaluated [15].

Human cytoglobin (CYGB) is expressed as a homodimer in *Escherichia coli*, which is stabilized by electrostatic interactions and hydrogen bonds [16-19]. Note that we here use the generally accepted abbreviations in globin research, whereby the human variant of globins is indicated in uppercase (e.g. CYGB), while the other variants are given in lower case (e.g. Cygb). It consists of two monomers with a sequence length of 190 amino acids. This sequence shows few similarities with vertebrate hemoglobin (Hb) and myoglobin (Mb) (< 30% identity) [1]. The tertiary structure of CYGB contains eight  $\alpha$ -helices, numbered A to H, that are ordered in a ‘three-over-three’  $\alpha$ -helix configuration characteristic for many globins [17]. In contrast to other globins, CYGB has extensions at the N- and C-termini, which seem to be disordered in the crystal structures [17,19,20] and might be involved in lipid binding [21]. Note that in one of the reported crystal structures, an additional helix is found in the N-terminus prior to the A helix [20], indicating that the termini may acquire order under certain circumstances. Similar to neuroglobin (Ngb) [22], the heme iron in Cygbs is hexacoordinated in the absence of exogenous ligands [17], being ligated to both the distal histidine in position 7 of the E-helix (HisE7) and proximal histidine (HisF8) in the ferrous (Fe<sup>2+</sup>) and ferric (Fe<sup>3+</sup>) states. In this way, there is a competition between the endogenous HisE7 and an external ligand, such as O<sub>2</sub>, to bind the central heme iron, and cleavage of the distal HisE7–Fe bond is necessary for binding of exogenous ligands [23,24].

```

C. aceratus Cygb-1 -----MERMQGEAE-GDHLERPSPLTDKEKVMIQDSWAKVYENCDDTGV 43
D. mawsoni Cygb-1 -----MERMQGEAE-GDHLERPSPLTDKEKVMIQDSWAKVYENCDDTGV 43
T. nigroviridis Cygb-1 -----MERMQRDGE-VDHVEQPGPLTEKEKVMIQDSWAKVFQSCDDAGV 43
T. nigroviridis Cygb-2 MSHREPPPPP-----QLAVQRRDVGQDGPERRAEPISDTEREMIRDAWGHVYKNCEDVGV 54
O. latipes Cygb-1 -----MERKQ--GE-VDHLERSRPLTDKERVMIQDSWAKVYQNCDDAGV 41
O. latipes Cygb-2 MSCRESPPPPSPPPQMLGVQRGECE--DRPERAEPISDAEMEIIQHTWGHVYKNCEDVGV 58
D. rerio Cygb-2 -----MEKEREDDE-TEGRERPEPLTDVERGI IKDTWARVYASCEDVGV 43
D. rerio Cygb-1 -----MEGDGG-VQLTQSPDSLTEEDVCVIQDTWKPVYAERDNAGV 40
CYGB -----MEKVPGEML-IERRERSEELSEAERKAVQAMWARLYANCEDVGV 43
                                     A B

C. aceratus Cygb-1 AILVRLFVGFPPSSRQYFSQFKHIEEPEELERSAQLRKHANRVMNGLNTLVESLDNSEKVA 103
D. mawsoni Cygb-1 AILVRLFVNFPPSSRQYFSQFKHIEEPEELERSAQLRKHANRVMNGLNTLVESLDNSEKVA 103
T. nigroviridis Cygb-1 AILVRFVNFPPSSKQYFKDFKHMEEPPEMQQSVQLRKHAHRVMTALNTLVESLDNADRA 103
T. nigroviridis Cygb-2 SILIRFFVNFPPSAKQYFSQFQDMEEPEEMERSSQLRHHACRVMNALNTVVENLHDPEKVS 114
O. latipes Cygb-1 AILVRLFVNFPPSSKQYFSQFKHIEDAEELKSSQLRKHARRVMNAINTLVESLDNSDKVS 101
O. latipes Cygb-2 SVLIRFFVNFPPSAKQYFSQFQDMQDPEEMEKSSQLRQHARRVMNAINTVVENLQDPEKVS 118
D. rerio Cygb-2 TILIRFFVNFPPSAKQYFSQFQDMEDPEEMEKSSQLRKHARRVMNAINTVVENLHDPEKVS 103
D. rerio Cygb-1 AVLVRFVNFPPSAKQYFEHRELQDPAEMQNAQLKKHGQRVLNALNTLVENLRDADKLN 100
CYGB AILVRFVNFPPSAKQYFSQFKHMEDEPEEMERSQLRKHACRVMGALNTVVENLHDPDKVS 103
                                     B C D E

C. aceratus Cygb-1 SVLKLLGKAHALRHKVEPVYFKILSGVILEVLGEAFSEVVT-PEVAAAATKLLATLYCGI 162
D. mawsoni Cygb-1 SVLKLLGKAHALRHKVEPVYFKILSGVILEVLGEAFSEVVT-PEVAAAATKLLATLYCGI 162
T. nigroviridis Cygb-1 SVLKSVGRAHALRHNVDPKYFKILSGVILEVLGEAFTEIIT-AEVASAWTKLLANMCCGI 162
T. nigroviridis Cygb-2 SVLAVVGRAHAVKHKVEPMYFKILSGVILEVLCEDFPEFFT-ADVQLVWSKLMATVYWHV 173
O. latipes Cygb-1 SVLNAVGKAHAIRHKVDPVYFKILSGVILEVLGEAYPQVMT-AEVASAWTNLLAILCCSI 160
O. latipes Cygb-2 SVLALVGKAHAVKHKVEPIYFKIXSGVMSVLSDFPEFFT-AEVQLVWTKLMAAVYWHV 177
D. rerio Cygb-2 SVLVLVGKAHAFKYKVEPVYFKILSGVILEILAEFEGECFT-PEVQTSWTKLMAALYWHI 162
D. rerio Cygb-1 TIFNQMGKSHALRHKVDVYFKILAGVILEVLVEAFPOCFSPAEVQSSWSKLMGILYWQM 160
CYGB SVLALVGKAHALKHKVEPVYFKILSGVILEVVAEEFASDFP-PETQRAWAKLRGLIYSHV 162
                                     F G H

C. aceratus Cygb-1 NAIYEEVGVGSKHSSSSG----- 179
D. mawsoni Cygb-1 NAIYEEVGVGSKHSSSSG----- 179
T. nigroviridis Cygb-1 AAVYKEAGWTELSSSVE----- 179
T. nigroviridis Cygb-2 TGAYTDVGVWLQVSSSAV----- 190
O. latipes Cygb-1 KAVYEELGVPHLSNSTS----- 177
O. latipes Cygb-2 TGAYTEVGVWLQVSSSAV----- 194
D. rerio Cygb-2 TGAYTEVGVWKLSSSAV----- 179
D. rerio Cygb-1 NRVAEVLGVWENSKK----- 174
CYGB TAAYKEVGVWQVVPNATTPPATLPSGGP 190
                                     H

```

**Figure 1.** Amino-acid sequences of Antarctic fish, *Tetraodon nigroviridis* [30], *Oryzias latipes* [30], *Danio rerio* Cygbs [30] and CYGB [17]. The Cys residues are indicated in red; identical amino-acid residues are in yellow and amino-acid differences between the two Antarctic fish Cygbs are in green. Positions of helices A-H are underlined according to the secondary structure of CYGB [17]. Adapted from Giordano et al. [37].

Additionally, CYGB has two Cys residues positioned at B2 and E9 (Figure 1). Binding kinetic studies on ferrous CYGB showed that it is possible to form an intramolecular disulfide bridge between these two cysteines [16]. While disulfide-bridge formation was shown to induce a decrease in the O<sub>2</sub> affinity by a factor of 10 in human Ngb (NGB), the effect is much smaller for CYGB (at most a factor of 2) [16]. Electron Paramagnetic Resonance (EPR) studies on ferric wild-type (wt) CYGB confirmed that the formation of intramolecular disulfide bridges does not have a

significant influence on the direct environment of the heme group [25]. Recent photoacoustic and transient absorption data indicate that the disulfide-bond formation mainly influences ligand migration and kinetics [26,27]. Surprisingly, recent kinetics studies indicate that the dissociation rate constant of the E7His increases with a factor of ~1000 when the intramolecular disulfide bridge is formed in monomeric CYGB [28]. Furthermore, monomeric CYGB with the internal disulfide bond intact interacts with lipids to induce changes in the heme environment [29]. This is not the case for the dimeric protein with intermolecular disulfide bonds and monomeric protein without an intramolecular disulfide bridge, clearly indicating a regulatory role for the disulfide bridge formation.

Cygb has been discovered also in teleosts, but unlike mammals, which have only a single gene copy [1], fish possess two paralogous Cygb genes (Cygb-1 and Cygb-2), which duplicated early in the teleost evolution [30]. Teleost Cygb-1 typically has 174 to 179 amino acids, while Cygb-2 has 179 to 196 residues. The sequence identity among fish and mammalian Cygbs shows that Cygb-2 is more closely related to mammalian Cygb than Cygb-1 [30]. Like their mammalian orthologs, both fish Cygbs are expressed in a wide range of tissues. However, Cygb-2 has been detected at highest levels in neuronal tissues, suggesting a sub-functionalization of the two globin paralogs after gene duplication [30].

The 16 members of the Antarctic icefish family Channichthyidae, including *Chaenocephalus aceratus* in this study, are the only vertebrates that do not have Hb in their blood (white-blooded fish) [31-33], and six species also lack Mb in the skeletal muscle as well as the heart muscle [34,35]. They, however, have retained Ngb [36] and Cygb [37]. Therefore, the Antarctic fish may provide a valuable system for understanding the mechanisms involved in the evolutionary development and alternative physiology of losing globin genes.

The current work describes the purification and physico-chemical analysis of the Cygbs-1 of *C. aceratus* (C.aceCygb-1) and of *Dissostichus mawsoni* (D.mawCygb-1), a related red-blooded Antarctic notothenioid [38] that has both Mb and Hb. The Cygb-2 gene has been found in the red-blooded Antarctic fishes *Notothenia coriiceps* ([http://www.ncbi.nlm.nih.gov/nuccore/XM\\_010778246.1](http://www.ncbi.nlm.nih.gov/nuccore/XM_010778246.1)) and *D. mawsoni* (Cheng unpublished). Antarctic fish have also Cygb-2 (Giordano, unpublished).

We here evaluate to what extent the heme-pocket structure and properties are affected by the sequence differences between the two Antarctic fish Cygbs and CYGB. Resonance Raman (RR), and continuous-wave (CW) and pulsed EPR spectroscopy have been used to characterize the heme environment, since they have been earlier shown to reveal valuable information on CYGB [25,39,40]. Furthermore, the multimeric state of the Antarctic fish Cygbs is characterized by analytical gel-filtration experiments, native mass spectrometry (MS) and 3D modeling. Finally, the redox potential of the different proteins under study is determined and compared.

## **2 Materials and Methods**

### **2.1. Cloning and sequencing of Cygb cDNA.**

Cygb cDNAs of the red-blooded notothenioid fish *D. mawsoni* and the icefish *C. aceratus* were cloned from retina and brain tissues, respectively. Total RNA was extracted using Ultraspec RNA isolation reagent (Biotecx, USA), and 4 µg was primed with a lock-dock oligodT30 primer for reverse-transcription with Superscript III (Invitrogen, USA) to produce first strand cDNA. A pair of degenerate primers designed to the conserved regions of Cygb-1 and Cygb-2 from percomorph fishes, perco\_cygf1 (5'-AGGGWGATSATCCAGSACWCRTGGG-3'), and perco\_cygf R2 (5'-GGATCTTAAAGTAYAYGGGKTCCACCTTGTGT-3') was first used to amplify a ~315nt Cygb cDNA fragment from the two notothenioid fishes. The PCR products were sequenced and verified to be Cygb cDNA. The partial cDNA sequences were then used to query an assembled transcriptome database of *D. mawsoni* [41], and identified a partial transcript covering the 5' end of Cygb inclusive of 5'UTR sequence. Using this sequence a notothenioid-specific 5'UTR primer, noto\_cyg5UTR\_F (5'-GAAACAGATTTCCATCCTCTCTCACCAGG-3') was designed and paired with lock-dock oligo dT30 to amplify the full length Cygb cDNA. The PCR products were cloned into pGemTeasy (Promega, USA), sequenced and verified to be Cygb. To confirm the validity of the nucleotide variations observed in some clones, a notothenioid-specific 3'UTR primer noto\_cyg3UTR\_R was designed to a conserved region in the 3'UTR of the Cygb cDNA of the two species, and paired with noto\_cyg5UTR\_F to re-amplify full-length Cygb cDNAs using Phusion® high fidelity DNA polymerase (NEB, USA). The fragment was subsequently cloned in SmaI site of pBSII KS- (Stratagene) and sequenced. Sequencing confirmed that the nucleotide variations were valid. All sequencing reactions were performed using BigDye® Terminator v3.1

chemistry (ABI) and electrophoresed on ABI 3730XL capillary sequencer at the Roy Carver Biotechnology Center at the University of Illinois, Urbana-Champaign.

The subcloning of *Cygb* cDNAs was performed in expression vectors pBADA. A PCR was carried out on the above mentioned plasmids using the 5' primer, containing a *Xho*I restriction site (5'-CCGCTCGAGATGGAGAGGATGCAGGGAGAGG-3'), while the 3' primer contained an *Eco*RI restriction site (5'-CCGGAATTCTCACCCACTTGAGCTTGAG-3'). The amplified product was cleaned and cut with *Xho*I and *Eco*RI at 37°C. The fragment was ligated in the similarly cleaved expression vector pBADA. The resulting plasmids were sequenced to verify that *D. mawCygb-1* and *C. aceCygb-1* were successfully cloned in the pBADA vector with the N-terminal His-tag in frame.

Cys residues were replaced by Ser (C38S/C160S double mutant) by using the QuickChange™ site-directed mutagenesis method (Stratagene). These mutants are abbreviated as *D. mawCygb-1\** and *C. aceCygb-1\**.

## **2.2 Protein Expression, Mutagenesis and Purification.**

Recombinant expression plasmids were transformed in the *Escherichia coli* TOP10 strain (Invitrogen). The cells were grown overnight at 37°C in 6 ml L-broth (10 g/L tryptone, 5 g/L yeast extract and 0.5 g/L NaCl) with 50 mg/L ampicillin. The grown cultures were poured into a flask containing 250 ml TB medium (1.2% bactotryptone, 2.4% yeast extract, 0.4% glycerol, 17 mM  $\text{KH}_2\text{PO}_4$  and 72 mM  $\text{K}_2\text{HPO}_4 \cdot 3\text{H}_2\text{O}$ ) and 50 mg/L ampicillin. The cultures were shaken at 160 rpm at 37°C. The cultures were induced at  $A_{550} = 1.0$  O.D by the addition of 20% L-arabinose to a final concentration of 0.2%. The expressions were continued overnight. The grown cells were harvested (20 min at 4000 rpm) and resuspended in 12 ml lysis buffer (50 mM TrisHCl pH7.5, 300 mM NaCl, 1% lysozyme). The resuspended cells were freeze-thawed three times and sonicated (1 minute 70 Hz with a 3-5 sec pulse). The extract was clarified by low speed centrifugation (10 min at 10000 rpm, 4°C). Imidazole was added to an end concentration of 20 mM and another low speed centrifugation was performed. The supernatant was loaded on a nickel Sepharose column (Clontech). The bound *Cygb* was eluted with elution buffer (250 mM imidazole, 50 mM TrisHCl pH 7.5, 300 mM NaCl). The *Cygb* fractions were dialyzed against 50 mM TrisHCl pH 7.5 and concentrated by using a Stirred Cell (Millipore) under 2 bar air pressure.

The recombinant mutants D.mawCygb-1\* and C.aceCygb-1\* were expressed and purified as described above for the wt proteins. The expression, mutagenesis and purification of CYGB and CYGB\*, used as comparative proteins to the Antarctic fish Cygbs, was identical to what was described previously [25].

The purity of the recombinant expressed proteins was analyzed by 15% SDS-PAGE (see Figure S2), MS (see section 2.4) and UV/Vis absorption at 280 nm and 416 nm. The respective  $R_z$  values for C.aceCygb and D.mawCygb are 1.91 and 1.89.

### **2.3 Analytical gel-filtration experiments.**

To determine the quaternary structure of Antarctic fish Cygbs in solution, analytical gel-filtration experiments were performed. The standards and the Antarctic fish Cygbs were passed through a Superose 12 column (25 mL volume). The column was equilibrated at room temperature with 50 mM TrisHCl pH 8.0, 150 mM NaCl and 0.5 mM EDTA. The flow rate was 0.5 ml/min and protein elution was monitored at wavelengths of 280 nm (probing amino acids with aromatic ring structure) and 412 nm [Soret band typical for heme proteins (globins)]. To evaluate whether protein concentration has an influence on the degree of the quaternary structure, different Antarctic fish Cygb-1 concentrations were analyzed (3  $\mu$ M, 12.5  $\mu$ M, 50  $\mu$ M, 100  $\mu$ M, 125  $\mu$ M, 500  $\mu$ M and 1 mM). Also Antarctic fish Cygb-1\* mutants were taken into account. A re-run of the fractionated multimers, after concentrating, was performed to check the equilibrium between the different forms.

Four standards (horse Mb, human Hb, recombinant CYGB and thyroglobulin) were taken into account to make a standard curve. The logarithm of the molecular weights was given as a function of elution time. The molecular weights of the Antarctic fish Cygb-1 proteins could be estimated using the standard curve.

As part of the sample preparation for native mass spectrometry, gel-filtration experiments were also run on a Superdex 200 10/300 GL column (25 mL volume) with a flow rate of 0.75 mL/min (see section 2.4).

### **2.4 Native Mass Spectrometry (MS).**

Proteins were diluted in 100 mM ammonium acetate to a final concentration of 20  $\mu$ M. The samples were desalted using BioSpin columns (Biorad) and were introduced into the vacuum of

the mass spectrometer using nano-electrospray ionization with in-house prepared gold-coated glass capillaries at a voltage of +1.6 kV. Ion mobility MS experiments were performed on a travelling-wave ion mobility Q-TOF instrument (Synapt G2 HDMS, Waters, Manchester, UK) [42,43] or a high mass modified Q-TOF (Micromass, Manchester, UK)[44] tuned to maintain native protein structure in the gas-phase. Critical voltages for ion mobility experiments on the Synapt type instrument were 25 V for the sampling cone, 4 V trap collision energy, 1.5 V transfer collision energy and 45 V for the trap DC bias. Pressures throughout the instrument were 6 mbar,  $4.5 \cdot 10^{-2}$  mbar, 3.1 mbar and  $4.3 \cdot 10^{-2}$  mbar for the source region, trap cell, ion mobility- and transfer cell, respectively. Ion mobility data were calibrated against proteins of known structure, as described elsewhere [45]. Experiments on the modified Q-TOF instrument were performed using 20  $\mu$ M protein in 100 mM AmAc for native or 10  $\mu$ M protein in 49:50:1 H<sub>2</sub>O:acetonitrile:formic acid for denatured experiments, which were obtained by collecting the monomer fraction of a SEC run on a Superdex 200 10/300 GL column (25 mL volume) with a flow rate of 0.75 mL/min. Critical voltages were 1.8 kV for the capillary, 100 V for the sampling cone, 3 V for the extractor and 25 V for the collision cell, with pressures of 4 mbar in the source region and  $1.5 \cdot 10^{-2}$  mbar in the collision cell.

Collision cross sections (CCS) for comparison of protein structures were derived from the coordinates of the structural model of C.aceCygb-1 that was predicted using the MODELLER software [46] (see further). Oligomeric head-to-tails models were manually docked and subsequently subjected to an energy minimization routine implemented in the Yasara software ([www.yasara.com](http://www.yasara.com)). All CCS from PDB files were calculated with the Mobcal algorithm using the projection approximation (PA) method at 298 K in He and scaled to the experimentally derived factor of  $1.14 \times (MW_{\text{exp}}/MW_{\text{model}})^{2/3} \times \text{CCS}_{\text{PA}}$  [47], to correct for the difference in ion mobility gas (N<sub>2</sub> in IM-MS and He in PA calculations) and the fact that the PA algorithm tends to underestimate cross sections.

## **2.5 Optical Absorption Spectroscopy.**

Optical absorption measurements of ferric (as-expressed), ferrous deoxy, and CO-ligated ferrous Antarctic fish Cygb-1 proteins were performed on a Varian Cary 5E UV-Vis-NIR spectrometer. All spectra were measured in a range from 350 to 700 nm at room temperature with a protein concentration of  $\sim 20 \mu\text{M}$ . The ferric form of the protein was obtained after purification. The

ferrous and CO-ligated ferrous forms were prepared in a sealed cuvette containing 1 mL of buffer solution (50 mM TrisHCl pH 7.5). The sample was then equilibrated with N<sub>2</sub> (ferrous form) or CO (CO-ligated ferrous form). The protein was added to a final concentration of ~20 μM using an airtight syringe and 10 μL of a saturated solution of sodium dithionite was finally added to reduce the protein and remove any residual oxygen.

## **2.6 Resonance Raman (RR) Spectroscopy**

Resonance Raman measurements were performed at room temperature on a Dilor XY-800 Raman spectrometer in low-dispersion mode using a liquid nitrogen-cooled CCD detector. The excitation source is a Kr<sup>+</sup> laser (Spectra-Physics BeamLok 2060) operating at 413.1 nm. The slit width used during the experiments is 200 μm.

During the measurements, the proteins were magnetically stirred at 500 rpm to avoid local heating and photochemical decomposition. Ten spectra were acquired with an integration time of 30-240 s each. In order to remove spikes caused by cosmic rays, we removed the highest and lowest data points for each frequency value, and averaged the remaining values. The power of the Kr<sup>+</sup> laser was 1-50 mW. The ferrous deoxy and CO-ligated forms of the proteins were obtained in the same way as for the optical absorption spectroscopy measurements. The CO-ligated samples were prepared with both <sup>12</sup>CO and <sup>13</sup>CO; the latter was obtained from Cortecnet (99% <sup>13</sup>C). All samples had a final protein concentration of ~50 μM.

## **2.7 Electron Paramagnetic Resonance**

X-band continuous wave (CW) EPR measurements were performed at low temperature (10 K) on a Bruker ESP300E spectrometer with a microwave frequency of ~9.45 GHz, equipped with a gas-flow cryogenic system (Oxford Instruments), allowing for operation from room temperature down to 2.5 K. The magnetic field was measured with a Bruker ER035M NMR Gauss meter. During the experiments, a vacuum pump was attached to the EPR tube in order to remove paramagnetic oxygen from the sample. The spectra of heme proteins are typically measured with a modulation amplitude of 0.5 mT, a modulation frequency of 100 kHz and a microwave power of 100 μW.

The effect of dithiothreitol (DTT) on the heme pocket was determined as follows. DTT was dissolved in the protein buffer to a concentration of 100 mM. This solution was then added to the

protein in the EPR tube to a final concentration of 10 mM DTT, after which the sample was immediately frozen for EPR measurements.

X-band pulsed EPR measurements were performed at low temperature (7K) on a Bruker E580 Elexsys spectrometer with a microwave frequency of 9.74 GHz, equipped with a gas-flow cryogenic system (Oxford Instruments) allowing for operation from room temperature down to 2.5 K. Electron-spin-echo (ESE) spectra [48] were recorded using the  $\pi/2$ - $\tau$ - $\pi$ - $\tau$ -echo sequence with  $t_{\pi/2} = 16$  ns,  $t_{\pi} = 32$  ns and  $\tau$  was varied from 88 ns to 2888 ns in steps of 8 ns. Three-pulse ESEEM spectra [48] were recorded using the  $\pi/2$ - $\tau$ - $\pi/2$ -T- $\pi/2$ - $\tau$ -echo sequence with  $t_{\pi/2} = 16$  ns,  $t_{\pi} = 32$  ns,  $\tau$  was varied from 88 ns to 648 ns in steps of 16 ns and T was varied from 88 ns to 5688 ns in steps of 16 ns. Hyperfine sublevel correlation (HYSCORE) spectra [49] were recorded using the  $\pi/2$ - $\tau$ - $\pi/2$ - $t_1$ - $\pi$ - $t_2$ - $\pi/2$ - $\tau$ -echo sequence with  $t_{\pi/2} = 16$  ns,  $t_{\pi} = 32$  ns and  $t_1$  and  $t_2$  were varied from 88 ns to 5688 ns in steps of 16 ns. HYSCORE measurements were recorded with different  $\tau$ -values and added together as indicated in the figure captions. The HYSCORE spectra are baseline corrected using a third-order polynomial, apodized with a Hamming window and zero-filled. After Fourier transformation, the absolute value spectrum was calculated. The Mims ENDOR (electron nuclear double resonance) spectra [50] were recorded using the  $\pi/2$ - $\tau$ - $\pi/2$ -T- $\pi/2$ - $\tau$ -echo microwave sequence with  $t_{\pi/2} = 16$  ns and  $\tau$  was varied from 88 ns to 488 ns in steps of 16 ns. An rf  $\pi$  pulse between 6.5  $\mu$ s and 11  $\mu$ s was inserted during time T, depending on the magnetic-field position. Spectra taken at the different  $\tau$ -values were added together to remove blind spots.

For all EPR measurements, 20% glycerol was added as a cryoprotectant. All spectra were computer simulated using Easyspin [51], a toolbox for MATLAB (MathWorks, Natick, Mass., USA).

## 2.8 Redox Potential Measurements

All electrochemical measurements were carried out using a  $\mu$ -Autolab III potentiostat (EcoChemie, The Netherlands) controlled by Nova 1.10 software. Gold disk working electrodes (MF-2014 BASi, surface area 2.0 mm<sup>2</sup>) were mechanically polished with abrasive slurries of diamond and alumina (particle sizes of 3, 1, 0.25 and 0.05  $\mu$ m). Residual polishing material was removed from the electrode surface by ultrasonication in ethanol and water for 5 min. The electrochemical pre-treatment was performed by recording 10 successive scans from 0.1 V to 1.5

V versus a saturated calomel electrode (SCE) in 0.5 M H<sub>2</sub>SO<sub>4</sub> at 0.1 Vs<sup>-1</sup>. Finally, the electrodes were incubated overnight in 8 mM 6-mercaptohexanol (Sigma-Aldrich) water solution. The electrodes were washed in copious amount of water before being installed in an electrochemical cell. Electrochemical behavior in the solution was studied using an electrochemical cell designed for measurements in a small sample volume [52]. 50 µl of the protein solutions (20-50 µM) was placed in the cell equipped with a glassy carbon counter and SCE (0.248 V versus SHE at 20 °C, REF401 Radiometer) electrodes. Alternatively a conventional three electrode cell was used, where a thin layer of protein solution was entrapped between the working electrode and a dialysis membrane (MWCO 12 kDa) [53]. The background voltammograms were recorded before introducing a sample of the proteins. All measurements were carried out at (20 ± 2) °C in 0.05 M TrisHCl buffer (pH 8.0) under nitrogen atmosphere.

### **2.9 3D-Modeling of C.aceCygb-1**

The structural model of C.aceCygb-1 was predicted using the MODELLER 9v15 software [46,54] in combination with Chimera [55]. The model was built omitting the N-terminal part of the protein (i.e. assuming C.aceCygb-1(17-179)). Procheck [56] evaluation indicated that the geometry of 100% of the residues is in allowed regions. Verify3D [57,58] showed that 84.66% of the residues have a score higher than 0.2, and ProSA-web [59,60] results confirm this finding through a negative Z-score for every residue, and a calculated Z-score of -7.5 for the global molecule. All these data indicate a reliable model. The predicted model of C.aceCygb-1 is compared to the known high-resolution 3D structure of CYGB (PDB: 2DC3 [A]) (Figure S1, supplementary information).

## **3 Results and discussion**

### **3.1 Amino-acid sequence analysis.**

C.aceCygb-1 and D.mawCygb-1 reported here, are considered as Cygb-1 (accession numbers KR73976 and KR732975, respectively), because their sequence is closely related to the Cygbs-1 of temperate fish with more than 70% of sequence identity, whereas they share from 54 to 59% of sequence identity with fish Cygbs-2 and CYGB (Figure 1 and Supplementary Table S1). The only

exception is for *Danio rerio*, since its Cygb-2 shares 60% of sequence identity with Cygb-1 of Antarctic fish (Supplementary Table S1). Fish Cygb-2 is more closely related to mammalian Cygb than fish Cygb-1 with a sequence identity of 66% [30].

In the liver and brain transcriptome of *C. aceratus*, Shin et al. [61] have identified Cygb-1 mRNA having the same nucleotide sequence found in the muscle of another icefish *Chionodraco hamatus* [62]. The Cygb-1 mRNA has also been identified in the brain of *C. hamatus* and *Trematomus bernacchii* (a red-blooded species of the family *Nototheniidae*) [37]. The latter Antarctic fish Cygb-1 displays very high sequence similarity with *C.aceCygb-1* and *D.mawCygb-1* (unpublished results). As mentioned, the Cygb-2 mRNA has been found in the red-blooded Antarctic fish *N. coriiceps*, *C. hamatus*, *T. bernacchii* and *D. mawsoni* DNA.

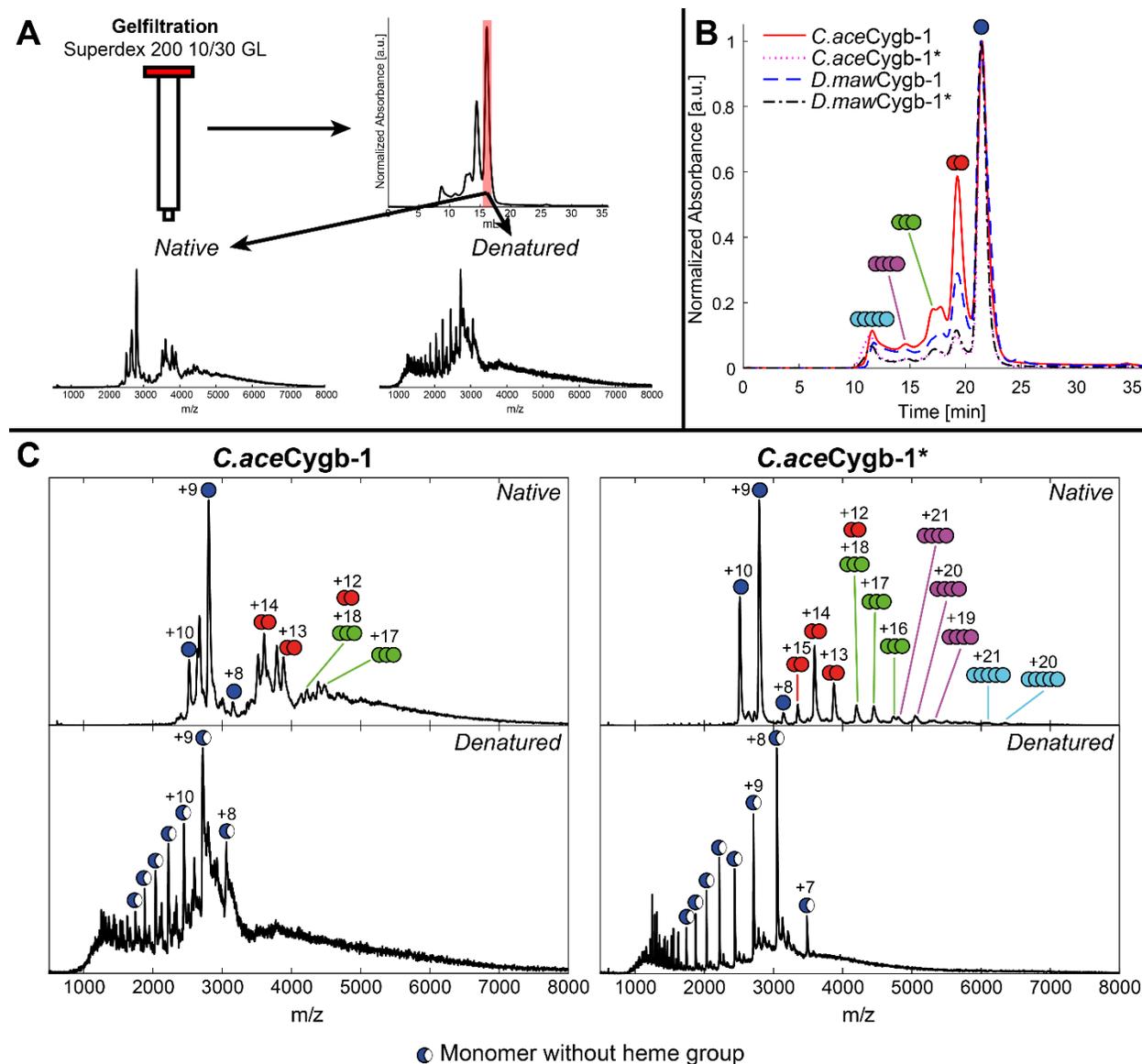
The two Antarctic fish Cygbs are composed of 179 amino-acid residues with a sequence identity of 98%, differing in only four positions: R24K (position A5), N52K (BC loop), I158M (H15 position), V165I (H22 position). The residues suggested to be essential for the function (Leu46 B10, Phe60 CD1, His81 E7, Arg84 E10, Val85 E11 and His113 F8, considered equally important in binding of an exogenous ligand [17,19,63]) are conserved and present in both Antarctic fish Cygbs (Figure 1).

Mammalian Cygbs contain conserved Cys residues, positioned at B2 and E9, which are known to create an intramolecular disulfide bridge in vitro [16] that slightly increases the O<sub>2</sub> affinity of mammalian Cygb under oxidizing conditions. However, these Cys residues are only conserved in *Tetraodon nigroviridis* Cygb-2, whereas in the other teleost Cygbs, as well as in Antarctic fish Cygb-1, the Cys at position E9 is replaced by another amino acid (Figure 1). Moreover, teleost Cygbs have another Cys residue positioned at H17 (Figure 1), too far from the first Cys to form a disulfide bridge [see discussion and modeling (Supplementary Figure S1)]. Therefore, the O<sub>2</sub> affinity of fish Cygb-1 is expected to depend differently on the redox state of the Cys residues than in mammalian Cygbs.

### **3.2 Multimeric state of Antarctic fish Cygbs.**

Analytical gel-filtration experiments for *C.aceCygb-1*\* suggested the presence of monomers and multimers [Figure 2B (low concentrations) and Figure S3 (high concentrations)]. A re-run of the

fractionated peaks (immediately after concentrating the samples) was performed (Supplementary Figure S4). No extra multimers were formed in the fractionated samples. The same results were obtained for *D.mawCygb-1* (data not shown).



**Figure 2.** (A) Cartoon showing measurement approach. (B) Chromatogram of the gel-filtration experiments performed on *C.aceCygb\** and *D.mawCygb\** (Superdex 200 10/300 GL column). (C) Nano-ESI mass spectra of *C.aceCygb-1* (left) and *C.aceCygb-1\** (right) introduced in the spectrometer under non-denaturing conditions (top) and denaturing conditions (bottom). The charges of the formed complexes and signals stemming from charged multimers are indicated.

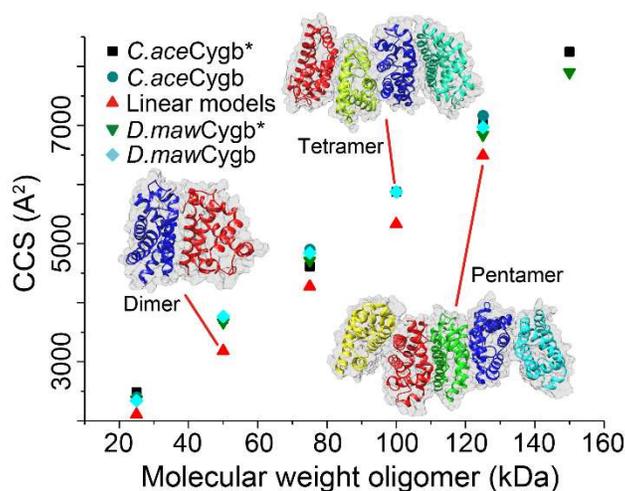
In order to study this in more detail, native MS was performed to assess the degree of heterogeneity present in the Antarctic Cygbs (Figure 2). The experimental approach is outlined in Figure 2A. First, gel-filtration experiments were performed over a Superdex 200 10/300 GL column on diluted samples with a final concentration of 60  $\mu$ M in a 100mM AmAc pH 7.5 buffer (Figure 2B). The volume representing the monomeric species was collected and kept overnight at 4°C at a concentration of 20  $\mu$ M. Nano-electrospray ionization (nESI), which leaves the structure and non-covalent interactions of a protein unaltered upon transfer to the gas phase, was then used to produce the protein ions. MS measurements were performed under both native and denaturing conditions. Figure 2C shows the results for C.aceCyg-1\*, while the corresponding results for D.mawCyg-1\* are shown in Figure S5. nESI generates multiply charged protein ions, with a Gaussian like charge state distribution for each species. For both Cygbs we observe spectra dominated by signals from the monomer (Figure 2 blue circles) and dimer (Figure 2 red circles) species. However, along with monomers and dimers, oligomeric species up to pentamers are observed. When chaotropic solvents (acetonitrile and formic acid), which weaken all non-covalent interactions, were added to the solutions and the solutions were left incubating overnight at 4°C, the signals from oligomeric protein were abolished (Figure 2C and S5). This shows that the driving forces behind oligomerization are non-covalent interactions between the subunits. When the experiments were repeated starting from a sample diluted from a high-concentration aged batch, similar multimers were observed under native conditions (Figure S6). However, under mild denaturing conditions (acetonitrile alone), the signal for oligomeric protein was partially retained (Figure S6C). Degrading in the presence of acetonitrile and formic acid was needed to abolish most of the signals of the multimers, indicating that the non-covalent multimers are quite stable. The initial oligomers formed in a non-covalent way may even be further stabilized by slow disulfide bridge formation. Such disulfide bridge formation may explain the more pronounced broad signals observed at  $m/z$  values larger than 3500 for the Cyg-1 in comparison to Cyg-1\*, both under denaturing conditions (Figures 2C and S5) and can also explain the notable effect of DTT on the MS spectrum of the high-concentration batch of D.mawCyg-1 (Figure S6D). However, these disulfide bridges are clearly not necessary for the oligomerization and are not the key factor in determining their high stability.

Note that in the mass spectrum of C.aceCyg-1 shown in Figure 2C, extra peaks due to a complex 1.1 kDa lower in mass are observed. These peaks are, however, not observed in any of the other

batches (e.g. Figure S6A) and are therefore ascribed to an impurity. Their presence does not influence the oligomerization behavior.

We further assessed the topology of oligomerization with IM-MS (Figure 3, Figure S7). All Antarctic fish Cygbs investigated (both wild-type and mutants) displayed similar CCS for the same oligomeric sizes, suggesting that oligomerization proceeds via a similar pathway in all cases. In addition, native IM-MS revealed that the Antarctic fish Cygb forms oligomers for which the CCS increases linearly with oligomer size, suggesting a chain-like growth of the oligomers. The obtained results are congruent with a head-to-tail oligomerization. In order to confirm the linear oligomerization pathway, we considered models of head-to-tail oligomeric *C.aceCygb-1* (Figure 3). Although the CCSs predicted for these oligomeric models were found to be somewhat lower than the experimental values, such differences are most likely due to the missing N-terminus in the models.

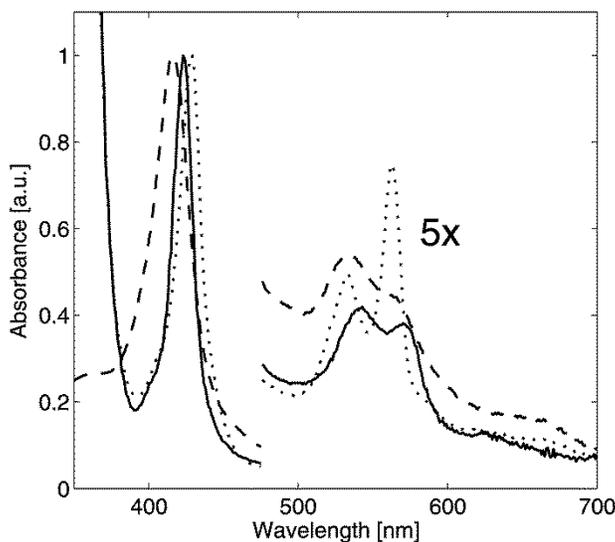
CYGB was also investigated with native MS (Figure S8). Compared to the spectra of Antarctic fish Cygbs-1, no extensive oligomeric CYGB was observed. The spectrum is dominated by monomeric CYGB, with only low intense signals for dimer CYGB.



**Figure 3.** Ion-mobility MS shows that Antarctic fish Cygbs form oligomers that grow linearly in size. Head-to-tail model of oligomeric *C.aceCygb-1* confirms a linear oligomerization pathway.

### 3.3 Optical absorption spectra.

Figure 4 shows the optical absorption spectra of the ferric, the ferrous deoxy and the CO-ligated ferrous forms of *C.aceCygb-1*. The corresponding spectra of *D.mawCygb-1* are shown in the supplementary information (Supplementary Figure S9) and are similar to those of *C.aceCygb-1*. Ferric *C.aceCygb-1* shows the Soret band at 415 nm and the  $\beta$  and  $\alpha$  Q-bands at 531 and 562 nm, respectively. This is typical for hexacoordination of the heme iron (low-spin (LS) state  $S=1/2$ ) and indicates bis-histidine ligation to the iron as was earlier observed for ferric CYGB [39]. Similarly, the Soret (425 nm) and Q-bands (531/559 nm) of the ferrous deoxy Antarctic fish Cygbs are typical for a bis-histidine coordination of the ferrous iron [39]. The Soret (423 nm) and Q-bands (543/471 nm) of the CO-ligated ferrous forms of the Antarctic fish Cygbs are also identical to those of the corresponding form of CYGB [39].

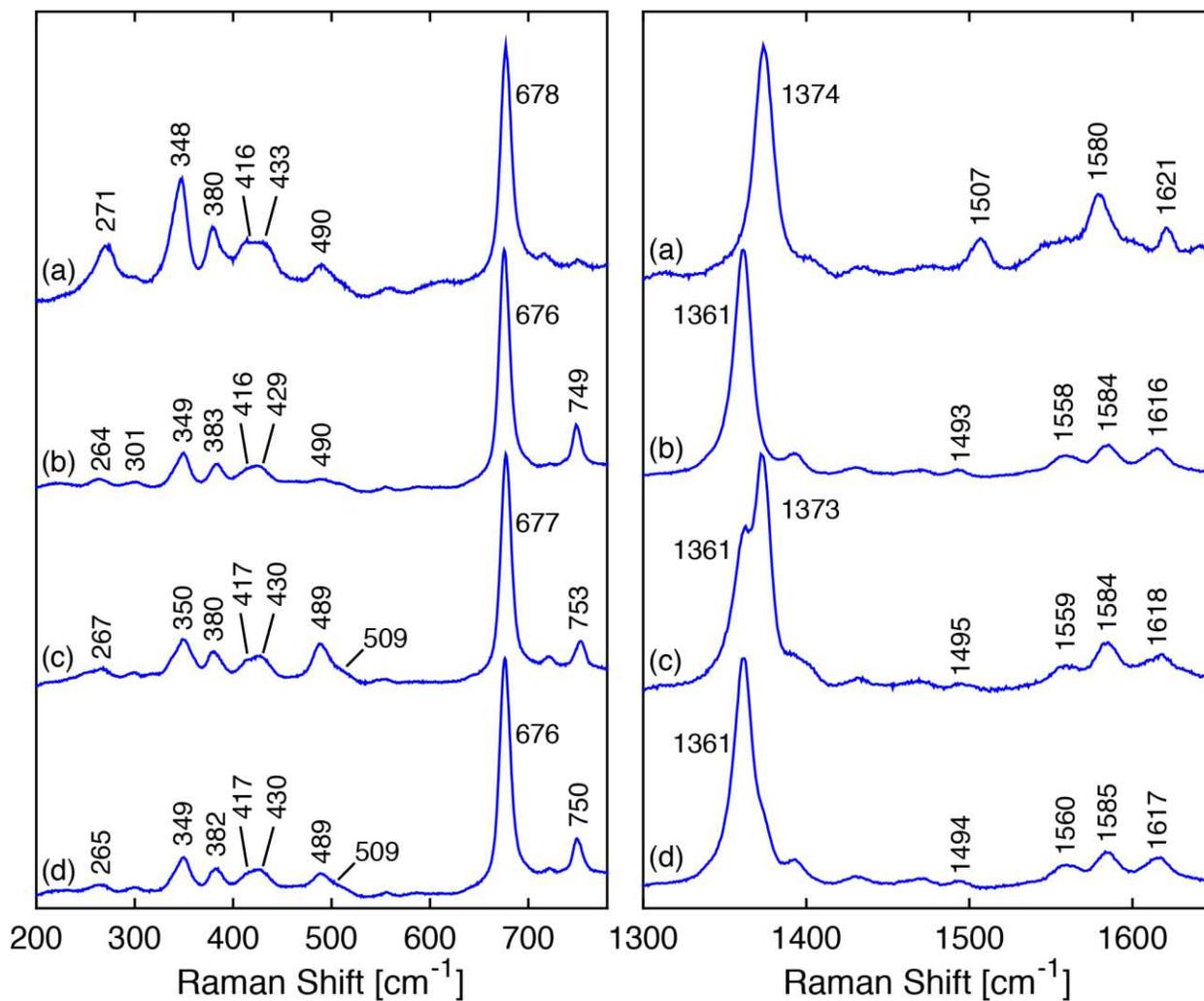


**Figure 4.** Optical absorption spectra of the ferric (dashed line), ferrous deoxy (dotted line) and CO-ligated ferrous (solid line) forms of *C.aceCygb-1* at pH 7.5. The spectra were normalized to the Soret band in order to allow better visual comparison of the shifts of the spectral lines.

### 3.4 Resonance Raman spectroscopy.

The porphyrin in-plane vibrational modes in the high-frequency region ( $1000-1700\text{ cm}^{-1}$ ) of the RR spectra of heme proteins present marker bands for the oxidation, coordination and spin state of the heme iron. In the low-frequency region ( $250-900\text{ cm}^{-1}$ ) different bending modes of the propionate, pyrrole and vinyl groups can be found [64]. Figure 5 shows the RR spectra of ferric

(a), ferrous deoxy (b), and CO-ligated ferrous (c,d) *C.aceCygb-1*. The corresponding spectra for *D.mawCygb-1* are shown in the supplementary information (Supplementary Figure S10). The spectra are very similar.



**Figure 5.** RR spectra in the low- and high-frequency range of *C.aceCygb-1* in the ferric (a), ferrous deoxy (b) and CO-ligated ferrous (c,d) form. The spectra were measured with a laser power of 12 mW (a,b), 1 mW (c) and 50 mW (d).

The main vibration modes of the ferric and ferrous ligated forms of the Antarctic fish *Cygb-1* are given in Supplementary Table S2 and compared to those of *CYGB* and other globins exhibiting heme bis-histidine coordination. The characteristic marker bands  $\nu_4$ ,  $\nu_3$  and  $\nu_2$  clearly confirm this coordination state for both the ferrous deoxy and the ferric state of the two proteins.

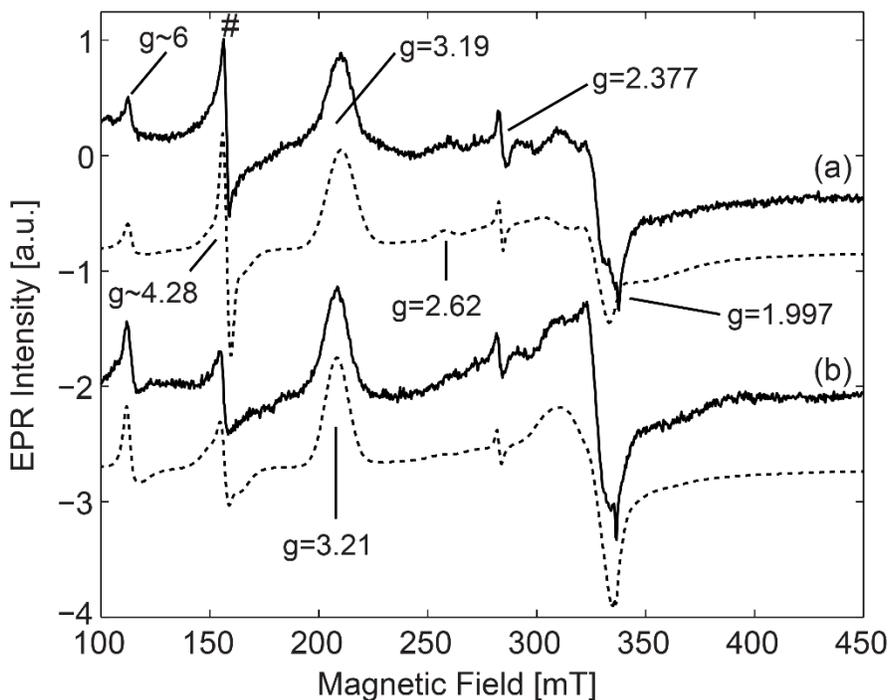
The size of the porphyrin ring can be calculated using the following formula  $v=K(A-d)$  where  $v$  represents the peaks correlated with  $\nu_{10}$ ,  $\nu_2$ ,  $\nu_{38}$  and  $\nu_3$  [65]. The distance between the porphyrin center and the pyrrole nitrogen is given by  $d$  (Å) and the values of  $K$  ( $\text{cm}^{-1}$ ) and  $A$  (Å) were determined by Choi et al. [65].  $\nu_{38}$  was not taken into account for the calculation of  $d$  for C.aceCygb-1 and D.mawCygb-1 because of the uncertainty on the peak position determination (Supplementary Table S2). For C.aceCygb-1 and D.mawCygb-1 the predicted distance is 0.1987(1) nm. This is comparable to the heme core dimension in CYGB, but somewhat lower than the one found for the tomato globin SOLly GLB1 [66]. The out-of-plane modes  $\gamma_6$ ,  $\gamma_7$ ,  $\gamma_{12}$  and  $\gamma_{21}$  are not visible in the ferric and ferrous heme forms indicating that the central iron is located in the heme plane due to the bis-histidine coordination.

The propionate bending mode  $\delta(\text{C}_\beta\text{C}_\alpha\text{C}_d)$  is found at 380-383  $\text{cm}^{-1}$  for all forms of the Antarctic fish Cygbs under study, in agreement with a strong hydrogen bond between the heme propionate and the surrounding amino acids [67]. A similar value was found for CYGB [39]. The vinyl bending modes are observed at 416 and 433  $\text{cm}^{-1}$  (ferrous deoxy form), 416 and 429  $\text{cm}^{-1}$  (ferric form) and 417 and 430  $\text{cm}^{-1}$  (CO-ligated form). Similar vinyl bending modes are observed for CO-ligated ferrous CYGB [39]. The two vinyl bending modes of the Cygbs are less separated than in the Mb case (408 and 439  $\text{cm}^{-1}$ ) [64].

Figure 5 reports the RR spectra of CO-ligated ferrous C.aceCygb-1 recorded with a laser power of 1 mW (Figure 5c) and 50 mW (Figure 5d). Upon increase of the laser power, the  $\nu_4$  component at 1361  $\text{cm}^{-1}$  (ferrous form) increases with respect to the  $\nu_4$  component at 1373  $\text{cm}^{-1}$  (CO-ligated ferrous form) in line with the photolysis of the heme-bound CO. At the same time, a marked change occurs in the 470-530  $\text{cm}^{-1}$  area. This allows an identification of the Fe-CO stretching modes ( $\nu_{\text{Fe-CO}}$ ). Two bands can be discerned: one strong band at 489  $\text{cm}^{-1}$  (77%) and one lower intense peak at 509  $\text{cm}^{-1}$  (23%) (Supplementary Figure S11 for fit). Analogous observations were done for the CO-ligated form of D.mawCygb-1 (Supplementary Figure S10). The identification of these bands is further corroborated by comparison of the RR spectra of the  $^{12}\text{CO}$ - and  $^{13}\text{CO}$ -ligated ferrous forms of the Antarctic fish Cygbs (Supplementary Figure S12), revealing a down-frequency shift of  $\sim 3$   $\text{cm}^{-1}$  for both bands in the  $^{13}\text{CO}$ -ligated Cygb case.

The RR spectra of CO-ligated ferrous C.aceCygb-1\* and D.mawCygb-1\* were found to be identical to those of the wild-type proteins (Figure S13). This shows that the mutations and putative formation of intermolecular disulfide bonds do not influence the stabilization of the CO ligand.

### 3.5 Electron Paramagnetic Resonance.



**Figure 6.** CW-EPR spectra of a frozen solution of ferric C.aceCygb-1 (a) and C.aceCygb-1\* (b) at pH 7.5. Solid line is the experiment, dashed line is the simulation. # indicates the contribution of a non-heme iron, the feature around  $g \sim 6$  is due to a small amount of HS ferric heme; the spectra are rescaled to the same microwave frequency and normalized to equal intensity for comparison.

Figure 6A shows the CW-EPR spectra of ferric C.aceCygb-1, and C.aceCygb-1\* at 10 K with their corresponding simulations. The corresponding spectra for D.mawCygb-1 and D.mawCygb-1\* are given in the supplementary information (Supplementary Figure S14) as well as the ESE-detected EPR spectra (Supplementary Figures S15-S16). Similar to earlier observations for ferric CYGB [25], all EPR spectra are dominated by a low-spin (LS,  $S=1/2$ ) ferric heme component that agrees with the bis-histidine coordinated ferric form (HisF8–Fe(III)–HisE7) of the protein (Table 1). The EPR spectrum also shows minor contributions of a non-heme iron (indicated with an asterisk,  $g \approx 4.28$ ), a batch-dependent high-spin (HS) ferric heme component ( $g_x = 5.96 \pm 0.01$ ,  $g_y$

=  $5.85 \pm 0.01$ ,  $g_z = 1.997 \pm 0.01$ ) and a minor LS ferric form ( $g_x = 1.85 \pm 0.01$ ,  $g_y = 2.37 \pm 0.01$ ,  $g_z = 2.62 \pm 0.01$ ) due to ligation of buffer molecules to the heme iron [25]. The batch-dependent HS form is related to protein denaturation and/or the breaking of the Fe-E7His bond. It stems from less than 1% of the globin proteins and is of little biological relevance. Furthermore, all spectra show a Cu(II) cavity background signal ( $g_z = 2.28 \pm 0.01$ ,  $g_y = g_x = 2.06 \pm 0.01$ ;  $|A_x|=|A_y|= 45 \pm 5$  MHz,  $|A_z|= 500 \pm 5$  MHz).

**Table 1.** Principal g values of the dominant LS heme form of the ferric Antarctic fish Cygbs-1 under study in comparison with other globins.

	$g_z$ ( $\pm 0.005$ )	$g_y$ ( $\pm 0.02$ )	$g_x$ ‡ ( $\pm 0.05$ )
D.mawCyg-1 (44mg/mL)	3.250	2.08	$\leq 1.05$
D.mawCyg-1 (5.5mg/mL)	3.230	2.08	$\leq 1.11$
D.mawCyg-1 (44mg/mL) + DTT	3.230	2.08	$\leq 1.11$
D.mawCyg-1*	3.210	2.08	$\leq 1.17$
C.aceCyg-1 (27.95 mg/mL)	3.200	2.08	$\leq 1.22$
C.aceCyg-1 + DTT	3.190	2.08	$\leq 1.25$
C.aceCyg-1*	3.210	2.08	$\leq 1.17$
CYGB [25]	3.200	2.08	1.20
CYGB* [25]	3.200	2.08	1.20
D.mawNgb [76]	3.125	2.15	1.27
C.aceNgb [76]	3.125	2.15	1.27
NGB [25]	3.10	2.17	1.30
	3.26	2.06	1.05
NGB* [25]	3.10	2.16	1.30
T. bernacchii Hb [77]	3.15	2.26	0.98
T. newnesi Hb [77]	3.13	2.24	1.08

‡ calculated assuming  $g_x^2 + g_y^2 + g_z^2 = 16$ , ESE-detected EPR indicates that this is an upper limit (see supplementary information, Figures S15-S16)

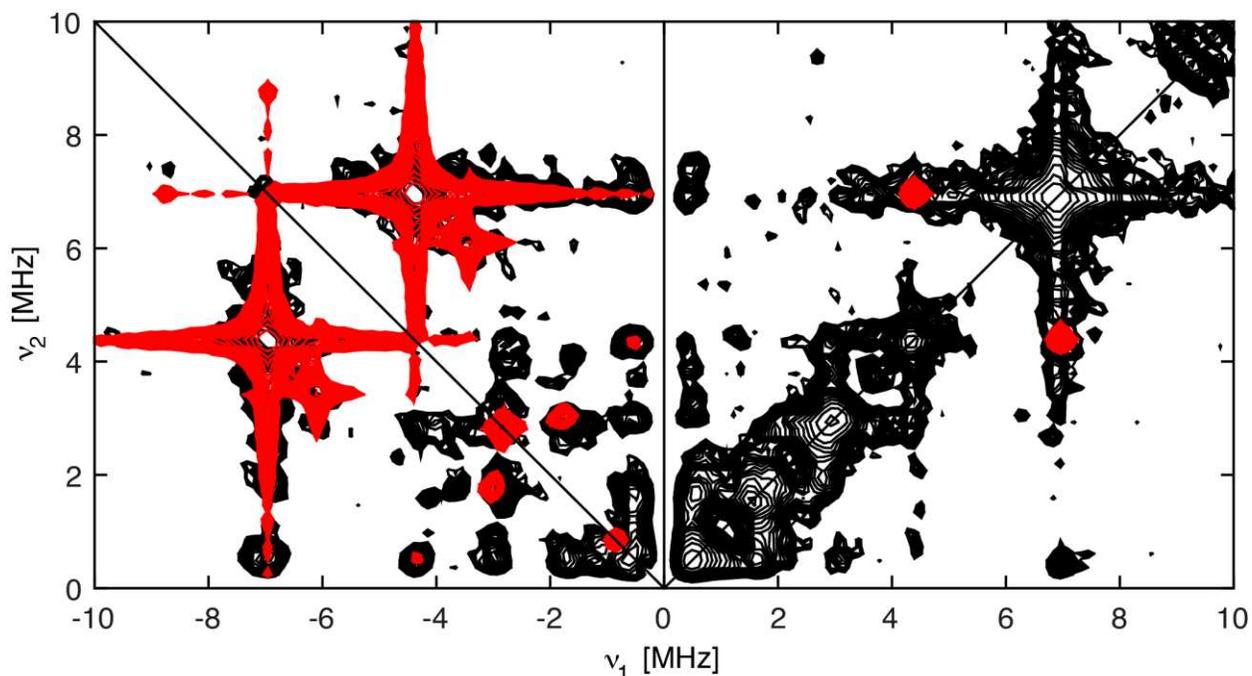
The principal  $g$  values of the dominant LS component of ferric C.aceCygb-1 and D.mawCygb-1 (Table 1) are in line with those found for ferric - CYGB [25]. They are typical for heme proteins with bis-histidine ligation of the ferric heme and a large dihedral angle between the imidazole planes. The  $g_z$  value increases with increasing dihedral angle, suggesting that D.mawCygb-1 has a slightly higher dihedral angle between the two heme-binding His imidazoles.

Interestingly, the principal  $g_z$  value of the C38S/C160S mutant D.mawCygb-1\* shows small but significant differences with those of the corresponding wt variant for the concentrated protein solutions (44mg/mL) (Table 1 and supplementary information, Figure S17). Furthermore, the EPR parameters are also influenced by addition of DTT to this high concentration sample. A similar, but more pronounced effect was observed earlier for NGB [25]. Indeed, the EPR parameters of NGB with or without an internal disulfide bridge differ strongly (Table 1), thus indicating a significant change in the heme environment upon disulfide-bridge formation. Although intramolecular disulfide-bridge formation was demonstrated to slightly influence the dioxygen affinity of CYGB, no change in the EPR parameters was observed when comparing the spectra of ferric wt CYGB and CYGB\* or when DTT was added to ferric wt CYGB [25]. For the Antarctic fish Cygbs intramolecular disulfide bridge formation can be excluded based on the 3D modeling (Supplementary Figure S1). However, MS experiments on samples diluted from the D.mawCygb-1 batch used for EPR with and without DTT showed a change in the multimer composition (Figure S6B,D). The observed change of the EPR spectrum after addition of DTT may thus agree with the break-up of the disulfide bridges. This reveals that, although disulfide bridges are not necessary for the oligomerization, their presence influences the heme-pocket region. More specifically, the dihedral angle between the two imidazole planes of E7His and F8His is slightly changed. Since the amino-acid residue at position B10 is located in the heme pocket of globins, formation of a disulfide bridge between B9Cys and H17Cys may induce a movement of the B-helix (and thus the position of B10), which can in turn influence the orientation of the E7His residue. Furthermore, the low-field EPR feature of the ferric wt forms shifts in solutions with lower proteins concentration (Table 1, Figure S17d), which may point to the loss of non-covalent multimers and/or reflects the concentration dependence of the disulfide-bridge formation.

The changes are more subtle and within the experimental error for C.aceCygb-1 (Table 1). This agrees with the fact that the MS data did not change notably for this sample upon addition of DTT.

It highlights further that disulfide bridges are not the key factors determining oligomerization for these proteins.

Earlier work has shown that a combination of CW-EPR, HYSCORE and ENDOR experiments can be used to deduce information on the relative orientation of the imidazole planes of the heme-ligating histidine residues in ferric globins [40,66]. There exist two conflicting X-ray structures of CYGB. One shows staggering imidazole planes of E7His and F8His with a dihedral angle between the two planes of 60-80° and different Fe-N<sub>ε</sub> distances for the two His [17]. The other shows a bis-histidine coordinated heme with only negligible differences between the Fe-N<sub>ε</sub> distances and a dihedral angle of 60° between the two imidazole planes [19]. The HYSCORE analysis of CYGB allowed to establish that the second structure was the one found in solution [40]. Different pulse EPR experiments were therefore set up for both ferric *C.aceCygb-1* and *D.mawCygb-1* and compared to our results obtained earlier for ferric wt CYGB [40]. Figure 7 shows the experimental <sup>14</sup>N HYSCORE spectrum of a frozen solution of *C.aceCygb-1* taken at an observer magnetic field corresponding to  $g = g_z$ . In the supplementary information, the HYSCORE spectra taken at other magnetic field positions and the corresponding spectra of ferric *D.mawCygb-1* are shown (Supplementary Figures S19-S23). Since the electron spin echo is very small at high field, three-pulse ESEEM experiments were also recorded to supplement the HYSCORE data (Supplementary Figures S24-S25). The <sup>14</sup>N HYSCORE spectra reflect the interactions of the unpaired electron on the iron and the nearby heme and His nitrogens (for more details and peak assignment, see [40,68] and supplementary information (Figure S18)). From the simulations of the HYSCORE and three-pulse ESEEM spectra (Figure 7 and Supplementary Figures S18-S23), the hyperfine and nuclear quadrupole tensors of these <sup>14</sup>N nuclei can be determined (Table 2). For each of the Antarctic fish Cygbs only one set of hyperfine/quadrupole parameters is found for the Fe-binding <sup>14</sup>N nuclei of the two His residues. This indicates that in each protein the Fe-N<sub>ε</sub> distance is the same for the two heme-iron His ligands. The hyperfine values of these nuclei are slightly lower (in absolute value) for Antarctic fish Cygbs than those found earlier for CYGB. This could indicate a slightly longer Fe-N<sub>ε</sub> distance in the former case.

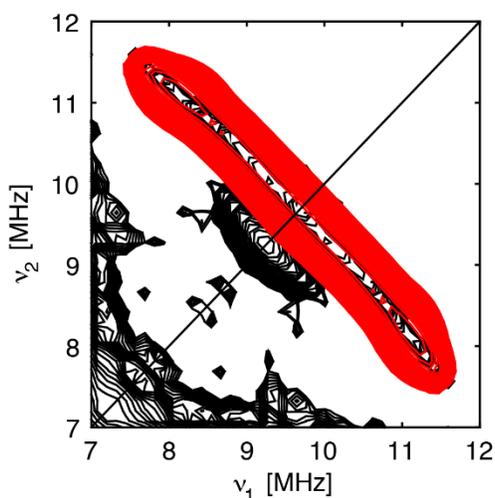


**Figure 7.** Experimental (black) and simulated (red)  $^{14}\text{N}$  HYSCORE spectrum of a frozen solution of ferric C.aceCygb-1 (pH 7.5) recorded at a magnetic field position corresponding to  $g = g_z$  ( $B_0 = 218$  mT). The experimental HYSCORE spectrum is recorded at 7.5 K and is the sum of those recorded for different  $\tau$ -values (88, 184 and 264 ns). The same procedure was applied for the simulated spectrum.

As shown earlier,  $^1\text{H}$  HYSCORE [40,68] and  $^1\text{H}$  ENDOR [69] reveal the hyperfine interaction with the nearby His protons that reflect the distance and rotation of the His imidazole planes versus the  $\mathbf{g}$  tensor frame. The  $^1\text{H}$  ENDOR spectra of ferric C.aceCygb-1 and D.mawCygb-1 were found to be the same (Supplementary Figure S26). Figure 8 shows the  $^1\text{H}$  HYSCORE and the simulation of the contribution of the His protons closest to the heme iron for the case of the ferric C.aceCygb-1 at one field setting (Supplementary Figure S27). The hyperfine tensor used for the simulation is  $[-3.7, -4.8, 6.8]$  MHz for the closest protons with Euler angles  $[-45, 36, \pm 30^\circ]$ . These values are very similar to those found for ferric CYGB [40], indicating a similar arrangement of the two imidazole planes of the HisE7 and HisF8 in the Antarctic fish Cygbs with maximum dihedral angle between the planes being  $60^\circ$  and equal Fe- $\text{N}_\epsilon$  distances (Supplementary Figure S28). This also indicates that the increase in dihedral angle going from CYGB to D.mawCygb-1 as follows from the  $\mathbf{g}$ -tensor analysis is only a few degrees.

**Table 2.** Hyperfine (**A**) and nuclear quadrupole (**Q**) tensors of the heme and histidine nitrogens of ferric D.mawCygb-1 and C.aceCygb-1 used for the spectral simulations of Figure 7 and Supplementary Figures S19-S25. An Euler angle  $\alpha = (22 \pm 10)^\circ$  was used for the tensors of the histidine nuclei in the **g**-tensor frame, for all other tensors, all Euler angles were  $(0 \pm 10)^\circ$ .

	$A_1$ [MHz]	$A_2$ [MHz]	$A_3$ [MHz]	$Q_1$ [MHz]	$Q_2$ [MHz]	$Q_3$ [MHz]
	( $\pm 0.2$ )	( $\pm 0.2$ )	( $\pm 0.05$ )	( $\pm 0.10$ )	( $\pm 0.10$ )	( $\pm 0.10$ )
<b>Histidine</b>						
D.mawCygb-1	-4.7	-5.55	-4.80	0.76	0.12	-0.88
C.aceCygb-1	-4.7	-5.55	-4.75	0.76	0.12	-0.88
CYGB [40]	-4.7	-5.90	-5.00	0.56	0.34	-0.90
<b>Heme</b>						
D.mawCygb-1	-4.0	-4.05	-5.53	0.82	-0.40	-0.42
C.aceCygb-1	-4.0	-4.05	-5.55	0.82	-0.40	-0.42
CYGB [40]	-4.0	-4.10	-5.45	0.92	-0.50	-0.42



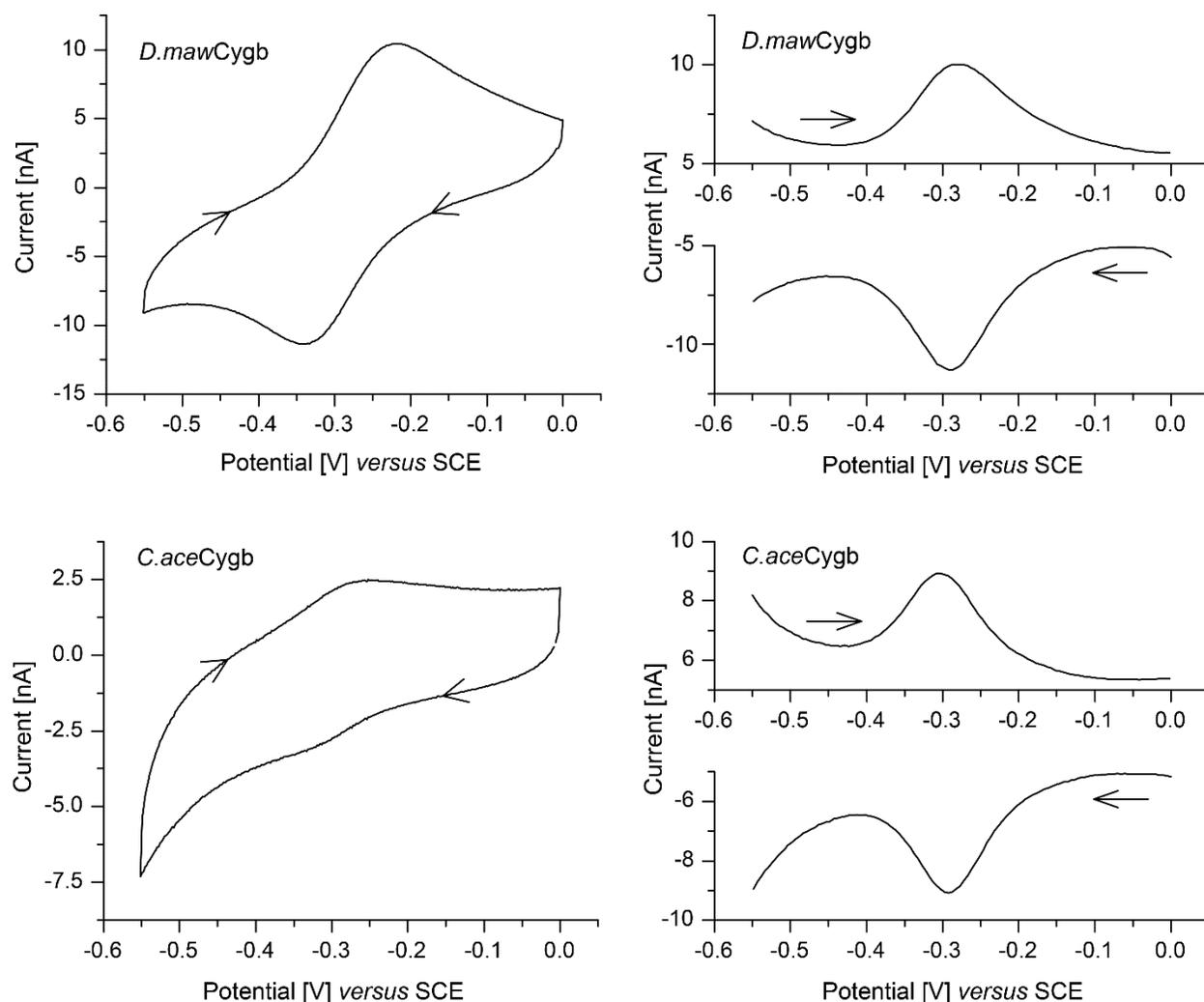
**Figure 8.** Experimental (black) and simulated (red)  $^1\text{H}$  HYSCORE spectrum of a frozen solution of ferric C.aceCygb-1 (pH 7.5) recorded at a magnetic field position corresponding to  $g=g_z$  (215 mT). The experimental HYSCORE spectrum is recorded at 7.5 K and is the sum of those recorded for different  $\tau$ -values (88, 184 and 264 ns). The same procedure was applied for the simulated

spectrum. Only the contributions of the HisE7 and HisF8 protons, nearest to the iron, were assumed.

**Redox potential of Antarctic fish Cygbs.** The electrochemical behavior of the Antarctic fish Cygbs-1 was studied by cyclic voltammetry (CV) and differential pulse voltammetry (DPV) in a 50 mM TrisHCl buffer (pH 8.0). The voltammograms (Figure 9, Supplementary Figures S29-S30) reveal one redox process for each protein at approximately -0.04 V (versus SHE) related to reduction/oxidation of the heme active site. The entrapped proteins gave clear peaks in both CV and DPV measurements. Table 3 lists values of the redox potential including CYGB and CYGB\* for comparison. The value of the wt proteins equals that of the mutant forms, whereas a small difference in the redox potential (10-15 mV) between D.mawCygbs-1 and C.aceCygbs-1 is noticed. The redox potential of CYGB lies closer to the one of D.mawCygbs-1 than of C.aceCygbs-1. The redox potentials found for dissolved and entrapped protein by DPV are identical (Table S2), however, CV measurements for dissolved D.mawCygbs-1 and C.aceCygbs-1 showed no clear peaks (Supplementary Figure S29), probably because of aggregation that slows down the protein diffusion and hinders the redox active site. It is the only difference found between wt and mutant Cygbs.

**Table 3.** Redox potential ( $E^\circ$ ) of the studied Cygbs obtained by differential pulse (DPV) and cyclic voltammetry (CV).

	$E^\circ$ [V] versus SHE $\pm$ 0.005V	
	DPV	CV
D.mawCygbs-1	-0.037	-0.031
D.mawCygbs-1*	-0.037	-0.035
C.aceCygbs-1	-0.048	-0.046
C.aceCygbs-1*	-0.043	-0.040
CYGB	-0.033	-0.026
CYGB*	-0.027	-0.025



**Figure 9.** Cyclic and differential pulse voltammograms for D.mawCygb-1 and C.aceCygb-1 entrapped in a thin layer between the electrode surface and the dialysis membrane (MWCO 12 kDa). Scan rate, 5 mVs<sup>-1</sup>; step potential, 5 mV; DPV modulation amplitude, 20 mV; 50 mM TrisHCl buffer (pH 8.0).

**Interpretation of obtained results.** The optical absorption spectra, RR spectra and CW-EPR spectra (Figures 4-6) confirm that C.aceCygb-1 and D.mawCygb-1 are ‘hexacoordinated’ in both the ferric and ferrous deoxy form, i.e. exhibiting a HisF8-Fe<sup>II/III</sup>-HisE7 ligation in agreement with CYGB [17,19,25,39]. Interestingly, a recent study revealed that zebrafish Cygb-1 is mainly pentacoordinated in its ferrous deoxy form [70], in contrast to what is observed here for the Antarctic fish Cygbs-1, whereas zebrafish Cygb-2 displays spectral features consistent with a

hexacoordinated heme [70]. Nothing is at present known about the spectroscopic characteristics of Antarctic fish Cygb-2, except that they share about 50% and 70% of sequence identity with *D. rerio* Cygb-1 and Cygb-2, respectively (Giordano, unpublished).

The RR spectra of the Antarctic fish Cygbs show similar propionate bending modes as for the CYGB case [39], indicative of a strong hydrogen bond between the propionate heme groups and the surrounding amino acids (Figure 4). In CYGB, the A propionate group interacts with the Lys116 and His117 of the FG loop, while the D propionate interacts with Gln77 at position E3 and two water molecules [17,19]. Furthermore, the nearby guanidinium group of Arg84 (E10) may be involved in electrostatic stabilization through the same propionate group [17]. Gln77 (E3), Arg84 (E10) and His117 are conserved in *C.ace*Cygb-1 and *D.maw*Cygb-1 (Figure 1), while the Lys116 is replaced by an Arg. This suggests a common heme-stabilization mechanism in the different Cygbs. Small differences in the EPR data for the Antarctic fish Cygbs-1 and CYGB indicate slight differences in the dihedral angle between the imidazole planes of the two His residues ligating to the heme iron (Table 1). Overall, the heme environment of the bis-histidine coordinated ferric form of the three proteins is very similar.

This changes when an external ligand, like CO, is bound. Two Fe-CO stretching modes are observed in the CO-ligated ferrous form of the Antarctic fish Cygbs. The dominant form (77%) has  $\nu_{\text{Fe-CO}}$  at  $489\text{ cm}^{-1}$  and agrees with an open structure, where no group is sufficiently close to the CO to affect the electronic structure of the Fe-C-O unit. This has been observed for Mb at low pH [71], for *C.ace*Ngb\* and *D.maw*Ngb\* [36] and for CYGB [39], although in the latter case this form constituted less than 40% of the Fe-CO forms. The second stretching mode ( $\nu_{\text{Fe-CO}} = 509\text{ cm}^{-1}$ ) agrees with a closed conformation where a positively charged group near the iron-bound CO stabilizes an  $\text{Fe}=\text{C}=\text{O}\cdots\text{X}^+$  structure [72]. This mode is very similar to the one observed in CO-ligated Mb at neutral pH and the  $\text{X}^+$  residue is the His at position E7. A small contribution of this mode was visible in CYGB (10%). Surprisingly, the CO-ligated form of *C.ace*Cygb-1 and *D.maw*Cygb-1 does not show an Fe-CO stretching mode at  $518\text{-}523\text{ cm}^{-1}$ , as was the case for CYGB (~50%) [39,73]. This mode was associated with a second closed  $\text{Fe}=\text{C}=\text{O}\cdots\text{X}^+$  structure, whereby stronger hydrogen bonding and electrostatic interactions occur than in the closed form characterized by  $\nu_{\text{Fe-CO}}$  at  $\sim 509\text{ cm}^{-1}$ . The X-ray structure of carbonmonoxy CYGB did indeed show a shorter distance between the HisE7 and the CO ligand than in the carbonmonoxy Mb case [24]. If the mode at  $520\text{ cm}^{-1}$  is present, an Fe-CO bending mode is visible at  $584\text{ cm}^{-1}$  [73]. In the

RR spectra of CO-ligated ferrous *C.aceCygb-1* and *D.mawCygb-1* this mode is also not present, thus confirming the absence of the second closed Fe-CO structure. This indicates different arrangements of the residues in the heme pocket. Note that CYGB has a Cys residue at position E9 (Cys83), while this is Asn in the two Antarctic fish Cygbs under study (Figure 1). CYGB has a second Cys at position B2 (Cys38) that is spatially close to Cys83 (E9), thus allowing formation of a disulfide bridge [16,18]. Although EPR revealed that the formation of the disulfide bond in CYGB does not induce a marked change in the structure of the heme pocket [25], the ligand binding affinity changes [16] and a switch in the ligand migration pathway is induced [26]. The lack of CysE9 (and hence the lack of similar disulfide bridges as in CYGB) may account for the different stabilization of the CO ligand in *C.aceCygb-1* when compared to CYGB. Indeed, a combined RR and FT-IR study of the ferrous-CO state of CYGB and mutants revealed that point mutation of Arg84 at position E10 to Ala induced a change in the relative contributions of the three Fe-CO stretching modes in favor of the open Fe-CO mode ( $492\text{ cm}^{-1}$ ) [73]. This proves that a point mutation in the E-helix may indeed affect the relative contributions of the different Fe-CO forms. Interestingly, the cold-adapted *C.aceNgb\** and *D.mawNgb\** have only two Fe-CO-modes ( $489$  and  $522\text{ cm}^{-1}$ ), while there are three CO-binding motives in NGB ( $494/505/521\text{ cm}^{-1}$ ) [36].

Another striking difference between the Antarctic fish Cygb-1 proteins and CYGB lies in their oligomerization behavior. For CYGB, only monomers and dimers have been reported. From gel-filtration experiments it was assumed that CYGB forms a dimer in solution [16,17]. However, a more detailed study revealed that CYGB is monomeric in a diluted solution and that the protein has a loose structure with high flexibility at the N- and C-termini instead of a more compact globular structure [18]. Native MS studies showed that low fractions of dimers in CYGB could form intermolecular disulfide bridges in vitro [18], as is confirmed in this work (Supplementary Figure S8). Furthermore, the formation of intermolecular disulfide bridges was shown to have an effect on the heme-pocket region, since clear structural differences were found between the crystal structures of wt CYGB (dimer with an intermolecular disulfide bridge between Cys38 and Cys83) [19] and CYGB\* (where the Cys are mutated to Ser) [17]. For Antarctic fish Cygbs-1, we here report the observation of oligomers (up to pentamers). These oligomers are found to be stable also at low concentrations and driven by non-covalent interactions. Although intermolecular disulfide bridges may further stabilize the oligomers, they are not essential for the oligomerization. Moreover, combined EPR and MS data of the protein with and without DTT indicates that the

disulfide bridges, if formed, induce small changes in the heme-pocket region. This may not be so surprising, giving the fact that one of the Cys is located on the B helix. Interestingly, the predicted 3D structure of *C.aceCygb-1* (Supplementary Figure S1), based on the known XRD structure of CYGB (PDB:1URV [17]), shows that the two Cys ligands are located at the surface of the protein pointing outwards at opposite sides of the protein. Whether the oligomerization of the Antarctic fish Cygbs occurs in vivo is rather difficult to ascertain and deserves further investigation, but it is worth noting that occurrence of the polymerization phenomenon displayed by other cold-adapted (hemo)globins may be a response to stressful environmental conditions [74].

The sequences of *C.aceCygb-1* and *D.mawCygb-1* differ only in 4 amino-acids (indicated in green in Figure 1). Not surprisingly, the spectroscopic parameters of both Antarctic fish Cygbs are very similar (Tables 1- 2 and Supplementary Table S2). The EPR parameters of ferric *C.aceCygb-1*\* and *D.mawCygb-1*\* are identical. It is unclear whether the small differences in the  $g_z$  values of ferric *C.aceCygb-1* and *D.mawCygb-1* (Table 1) are due to differences in the heme pocket of the *Cygb-1* monomer or reminiscent of different relative amounts of the formed multimers. The hyperfine tensor and nuclear quadrupole tensor obtained for the heme-ligating histidine nitrogens in ferric CYGB are slightly different from those obtained for the ferric Antarctic fish Cygbs-1 (Table 2). This may indicate local differences in the heme-pocket structure, possibly the difference in the Fe-N<sub>ε</sub> distance and slight differences in the dihedral angle between the imidazole planes of the Fe-bound His ligands. The nuclear quadrupole tensors are influenced by the full electronic structure, and not only by the localization of the unpaired electron. Comparison of the sequences of CYGB and the Antarctic fish Cygbs-1 (Figure 1) indicate small differences in the B, E and F helices surrounding the heme. The differences at position E9 (CYGB: Cys, Antarctic fish Cygbs: Asn) and B12 (CYGB: Phe, Antarctic fish Cygbs: Leu) are the most striking, which may lead to small local variations in the heme-pocket environment. CYGB is longer at the C-terminal and Lechauve et al. [18] showed that the C- and N-terminal ends are responsible for the loose structure of the protein, which may have again an effect on the heme-pocket structure. Next to crystal structures in which the C- and N-terminal segments were not resolved due to disorder (PDB:1URV and 1V5H), a crystal structure of wt CYGB could be obtained that showed an additional helix in the N-terminal region prior to the A helix and an ordered loop in the C-terminal region (PDB:2DC3) [20]. The two crystal forms showed substantial differences in the conformation of

the residues in the heme environment, thus suggesting that the terminal region is structurally important.

The redox potentials of C.aceCygb-1 and D.mawCygb-1 are in the range of -0.03 to -0.05 V (versus SHE). Mutation of the Cys to Ser did not alter the redox potential, but a small difference (10-15 mV) between C.aceCygb-1 and D.mawCygb-1 was observed. The difference is close to the expected experimental error and is negligible for mutant forms. Notably, the largest difference is observed between the redox potential of C.aceCygb-1 and CYGB (15-20 mV). The redox potential of CYGB ( $-0.033 \pm 0.005$  V versus SHE) found in the present work by the direct electrochemical measurements equals to the previously reported value obtained by redox titration ( $-0.028 \pm 0.005$  V) [75]. To conclude, the direct electrochemical measurements show no significant variation in the redox potential between mutant and wild type forms of the Antarctic fish Cygbs, but minor differences among Cygbs of different origin are observed. This may be related to the earlier mentioned small differences in the heme environment.

The impact of the primary structure on the spectral behavior of Antarctic fish Cygbs-1 is clearly small since their sequence identity is very high, although C. aceratus and D. mawsoni differ hugely in terms of globin composition in their genome. Other mechanisms, such as differences in gene regulation and/or protein expression, may account for Cygbs function and regulation under specific physiological requirements in Antarctic fish.

The difference between CYGB and Antarctic fish Cygbs-1 are clearer. They are most pronounced in the multimerization behavior (see MS data) and the binding of exogenous ligands (see RR of CO-ligated forms), and to a lesser extent in their redox properties. The ability of Antarctic fish Cygbs-1 to form stable, non-covalently linked multimers up to pentamers even at low concentrations is intriguing and may mark a common feature with other cold-adapted (hemo)globins [74].

#### ACKNOWLEDGEMENT

The authors acknowledge the support of the University of Antwerp GOA-BOF funding (28312), FWO funding (G.0687.13) and the Hercules foundation for funding of the Synapt G2 instrument. This study was carried out in the framework of the SCAR program “Antarctic Thresholds – Ecosystem Resilience and Adaptation” (AnT-ERA). It was financially supported by the Italian

National Program for Antarctic Research (PNRA). Research of A. De Schutter is funded by a PhD grant of the Agency for Innovation by Science and Technology (IWT, Belgium). C-H C. Cheng acknowledges funding support from US National Science Foundation Polar Programs (ANT-1142158).

**Supporting Information Available.** Amino-acid sequence identity analysis, 3D modeling of C.aceCygb-1 structure, extra gel-filtration and MS data on multimeric state of Antarctic fish Cygbs, extra optical absorption and RR spectra, full EPR and hyperfine spectroscopy analysis and detailed electrochemical experiments. This material is available free of charge via the Internet.

## REFERENCES

1. 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
2. E. Geuens, I. Brouns, D. Flamez, S. Dewilde, J. P. Timmermans, L. Moens, A globin in the nucleus!, *J. Biol. Chem.* 278 (2003) 30417-30420.
3. N. Kawada, D. B. Kristensen, K. Asahina, K., Nakatani, Y. Minamiyama, S. Seki, K. Yoshizato, Characterization of a stellate cell activation-associated protein (STAP) with peroxidase activity found in rat hepatic stellate cells, *J. Biol. Chem.* 276 (2001) 25318-25323.
4. U. Oleksiewicz, T. Liloglou, J. K. Field, G. Xinarianos, Cytoglobin: biochemical, functional and clinical perspective of the newest member of the globin family, *Cell. Mol. Life Sci.* 68 (2011) 3869-3883.
5. H. Li, C. Hemann, T. M. Abdelghany, M. A. El-Mahdy, J. L. Zweier, Characterization of the Mechanism and Magnitude of Cytoglobin-mediated Nitrite Reduction and Nitric Oxide Generation under Anaerobic Conditions, *J. Biol. Chem.* 287 (2012) 36623-36633.
6. N. Shivapurkar, V. Stastny, N. Okumura, L. Girard, Y. Xie, C. Prinsen, F. B. Thunnissen, I.I. Wistuba, B. Czerniak, E. Frenkel, J.A. Roth, T. Liloglou, G. Xinarianos, J. K. Field, J. D., Minna, A. F. Gazdar, Cytoglobin, the newest member of the globin family, functions as a tumor suppressor gene, *Cancer Res.* 68 (2008) 7448-7456.
7. A. Avivi, F., Gerlach, A., Joel, S., Reuss, T., Burmester, E., Nevo, T. Hankeln, Neuroglobin, cytoglobin, and myoglobin contribute to hypoxia adaptation of the subterranean mole rat *Spalax*, *Proc. Natl. Acad. Sci. U. S. A.* 107 (2010) 21570-21575.
8. D. Li, X. Q. Chen, W.-J. Li, Y.-H., Yang, J. Z. Wang, A.C.H. Yu., Cytoglobin up-regulated by hydrogen peroxide plays a protective role in oxidative stress, *Neurochem. Res.* 32 (2007) 1375-1380.
9. K. Nakatani, H. Okuyama, Y. Shimahara, S. Saeki, D. H. Kim, Y. Nakajima, S. Seki, N. Kawada, K. Yoshizato, Cytoglobin/STAP, its unique localization in splanchnic fibroblast-like cells and function in organ fibrogenesis, *Lab. Invest.* 84 (2004) 91-101.
10. M. Schmidt, F. Gerlach, A. Avivi, T. Laufs, S. Wystub, J. C. Simpson, E. Nevo, S. Saaler-Reinhardt, S. Reuss, T. Hankeln, T. Burmester, Cytoglobin is a respiratory protein

- in connective tissue and neurons, which is up-regulated by hypoxia, *J. Biol. Chem.* 279 (2004) 8063-8069.
11. J. I. Stagner, R. S. Seelan, R. N. Parthasarathy, K. White, Reduction of ischemic cell death in cultured islets of Langerhans by the induction of cytoglobin, *Islets* 1 (2009) 50-54.
  12. H. Nishi, R. Inagi, N. Kawada, K. Yoshizato, I. Mimura, T. Fujita, M. Nangaku, Cytoglobin, a Novel Member of the Globin Family, Protects Kidney Fibroblasts against Oxidative Stress under Ischemic Conditions, *Am. J. Pathol.* 178 (2011) 128-139.
  13. I. Okayasu, T. Mikami, T. Yoshida, K. Hana, M. Yokozawa, M. Sada, M. Fujiwara, N. Kawada, Cytoglobin expression of rectal subepithelial myofibroblasts: Significant alterations of cytoglobin(+) stromal cells in long-standing ulcerative colitis, *Histol. Histopathol.* 26 (2011) 679-688.
  14. W. Cui, M. Wang, H. Maegawa, Y. Teranishi, N. Kawada, Inhibition of the activation of hepatic stellate cells by arundic acid via the induction of cytoglobin, *Biochem. Biophys. Res. Commun.* 425 (2012) 642-648.
  15. T. C. Bhola, V. S. Neergheen-Bhujun, N. J. Hodges, S. D. Dyal, T. Bahorun, Cytoglobin as a biomarker in cancer: potential perspective for diagnosis and management, *Biomed. Res. Int.*, (2015) Article ID 824514.
  16. 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.
  17. D. de Sanctis, S. Dewilde, A. Pesce, L. Moens, P. Ascenzi, T. Hankeln, T. Burmester, M. Bolognesi, Crystal structure of cytoglobin: The fourth globin type discovered in man displays heme hexa-coordination, *J. Mol. Biol.* 336 (2004) 917-927.
  18. C. Lechauve, C. Chauvierre, S. Dewilde, L. Moens, B.N. Green, M. C. Marden, C. Celier, L. Kiger, Cytoglobin conformations and disulfide bond formation, *FEBS J.* 277 (2010) 2696-2704.
  19. H. Sugimoto, M. Makino, H. Sawai, N. Kawada, K. Yoshizato, Y. Shiro, Structural basis of human cytoglobin for ligand binding, *J. Mol. Biol.* 339 (2004) 873-885.

20. M. Makino, H. Sugimoto, H. Sawai, N. Kawada, K. Yoshizato, Y. Shiro, High-resolution structure of human cytoglobin: identification of extra N- and C-termini and a new dimerization mode, *Acta Crystallogr. Sect. D-Biol. Crystallogr.* 62 (2006) 671-677.
21. B.J. Reeder, D. A. Svistunenko, M. T. Wilson, Lipid binding to cytoglobin leads to a change in haem co-ordination: a role for cytoglobin in lipid signalling of oxidative stress, *Biochem. J.* 434 (2011) 483-492.
22. A. Pesce, S. Dewilde, M. Nardini, L. Moens, P. Ascenzi, T. Hankeln, T. Burmester, M. Bolognesi, Human brain neuroglobin structure reveals a distinct mode of controlling oxygen affinity, *Structure*, 11 (2003) 1087-1095.
23. S. Kakar, F. G. Hoffman, J. F. Storz, M. Fabian, M. S. Hargrove, Structure and reactivity of hexacoordinate hemoglobins, *Biophys. Chem.* 152 (2010) 1-14.
24. M. Makino, H. Sawai, Y. Shiro, H. Sugimoto, Crystal structure of the carbon monoxide complex of human cytoglobin, *Proteins: Struct. Funct. Bioinform.* 79 (2011) 1143-1153.
25. E. Vinck, S. Van Doorslaer, S. Dewilde, L. Moens, Structural change of the heme pocket due to disulfide bridge formation is significantly larger for neuroglobin than for cytoglobin, *J. Am. Chem. Soc.* 126 (2004) 4516-4517.
26. L. Astudillo, S. Bernad, V. Derrien, S. Pierre, J. Miksovska, Reduction of the internal disulfide bond between Cys 38 and 83 switches the ligand migration pathway in cytoglobin, *J. Inorg. Biochem.* 129 (2013) 23-29.
27. H. Tsujino, T. Yamashita, A. Nose, K. Kukino, H. Sawai, Y. Shiro, T. Uno, Disulfide bonds regulate binding of exogenous ligand to human cytoglobin, *J. Inorg. Biochem.* 135 (2014) 20-27.
28. P. Beckerson, B. J. Reeder, M. T. Wilson, *FEBS Lett.* 589 (2015) 507-512.
29. P. Beckerson, M. T. Wilson, D. A. Svistunenko, B. J. Reeder Cytoglobin ligand binding regulated by changing haem-co-ordination in response to intramolecular disulfide bond formation and lipid interaction, *Biochem. J.* 465 (2015) 127-137.
30. C. Fuchs, A. Luckhardt, F. Gerlach, F. Burmester, T. Hankeln, Duplicated cytoglobin genes in teleost fishes, *Biochem. Biophys. Res. Commun.* 337 (2005) 216-223.
31. B. D. Sidell, K. M. O'Brien, When bad things happen to good fish: the loss of hemoglobin and myoglobin expression in Antarctic icefishes, *J. Exp. Biol.* 209 (2006) 1791-1802.

32. T. J. Near, S. K. Parker, H. W., III Detrich, A genomic fossil reveals key steps in hemoglobin loss by the Antarctic icefishes, *Mol. Biol. Evol.* 23 (2006) 2008-2016.
33. D. J. Small, T. Moylan, M. E. Vayda, B. D. Sidell, The myoglobin gene of the Antarctic icefish, *Chaenocephalus aceratus*, contains a duplicated TATAAAA sequence that interferes with transcription, *J. Exp. Biol.* 206 (2003) 131-139.
34. B. D. Sidell, M. E. Vayda, D. J. Small, T. J. Moylan, R. L. Londraville, M. L. Yuan, K. J. Rodnick, Z. A. Eppley, L. Costello, Variable expression of myoglobin among the hemoglobinless Antarctic icefishes, *Proc. Natl. Acad. Sci. U. S. A.* 94 (1997) 3420-3424.
35. M. E. Vayda., D. J. Small, M. L. Yuan, L. Costello, B. D. Sidell, Conservation of the myoglobin gene among Antarctic notothenioid fishes, *Mol. Mar. Biol. Biotechnol.* 6 (1997) 207-216.
36. D. Giordano, I. Boron, S. Abbruzzetti, W. Van Leuven, F. P. Nicoletti, F. Forti, S. Bruno, C. H. C. Cheng, L. Moens, G. di Prisco, A. D. Nadra, D. Estrin, G. Smulevich, S. Dewilde, C. Viappiani, C. Verde, Biophysical Characterisation of Neuroglobin of the Icefish, a Natural Knockout for Hemoglobin and Myoglobin. Comparison with Human Neuroglobin, *PLoS One* 7 (2012) e44508.
37. D. Giordano, R. Russo, D. Coppola, G. Altomonte, G. Di Prisco, S. Bruno, C. Verde, Cool adaptations to cold environments: globins in Notothenioidei, *Hydrobiologia*, 761 (2015) 293-312.
38. A. J. Kiss, C. H. C. Cheng, Molecular diversity and genomic organisation of the alpha, beta and gamma eye lens crystallins from the Antarctic toothfish *Dissostichus mawsoni*, *Comp. Biochem. Phys. D* 3 (2008) 155-171.
39. H. Sawai, N. Kawada, K. Yoshizato, H. Nakajima, S. Aono, Y. Shiro, Characterization of the heme environmental structure of cytoglobin, a fourth globin in humans, *Biochemistry* 42 (2003) 5133-5142.
40. A. I. Ioanimescu, S. Van Doorslaer, S. Dewilde, B. Endeward, L. Moens, Probing the heme-pocket structure of the paramagnetic forms of cytoglobin and a distal histidine mutant using electron paramagnetic resonance, *Mol. Phys.* 105 (2007) 2073-2086.
41. Z. Chen, C. H. C. Cheng, J. Zhang, L. Cao, L. Chen, L. Zhou, Y. Jin, H. Ye, C. Deng, Z. Dai, Q. Xu, P. Hu, S. Sun, Y. Shen, L. Chen, Transcriptomic and genomic evolution

- under constant cold in Antarctic notothenioid fish, *Proc. Natl. Acad. Sci. U. S. A.* 105 (2008) 12944-12949.
42. F. Sobott, M. G. McCammon, H. Hernandez, C. V. Robinson, The flight of macromolecular complexes in a mass spectrometer, *Philosophical Transactions of the Royal Society of London Series a-Mathematical Physical and Engineering Sciences* 363 (2005) 379-389.
43. A. Konijnenberg, A. Butterer, F. Sobott, Native ion mobility-mass spectrometry and related methods in structural biology, *Biochim. Biophys. Acta-Proteins and Proteomics* 1834 (2013) 1239-1256.
44. R. H. van den Heuvel, H. H Robert, E. van Duijn, H. Mazon, S. A. Synowsky, K. Lorenzen, C. Versluis, S. J. J. Brouns, D. Langridge, J. van der Oost, J. Hoyes, A. J. R. Heck, Improving the performance of a quadrupole time-of-flight instrument for macromolecular mass spectrometry, *Anal. Chem.* 78 (2006) 7473-7483.
45. M. F. Bush, Z. Hall, K. Giles, J. Hoyes, C.V. Robinson, B. T. Ruotolo, Collision Cross Sections of Proteins and Their Complexes: A Calibration Framework and Database for Gas-Phase Structural Biology, *Anal. Chem.* 82 (2010) 9557-9565.
46. N. Eswar, B. Webb, M. A. Marti-Renom, M. S. Madhusudhan, D. Eramian, M.-Y. Shen, U. Pieper, A. Sali, Comparative Protein Structure Modeling Using Modeller, In *Current Protocols in Bioinformatics*, John Wiley & Sons, Inc , 2006, Unit 5.6.
47. Z. Hall, A. Politis, M. F. Bush, L. J. Smith, C. V. Robinson, Charge-State Dependent Compaction and Dissociation of Protein Complexes: Insights from Ion Mobility and Molecular Dynamics, *J. Am. Chem. Soc.* 134 (2012) 3429-3438.
48. A. Schweiger, G. and Jeschke, Principles of pulse electron paramagnetic resonance, Oxford University Press, Oxford, 2001.
49. P. Höfer, A. Grupp, H. Nebenführ. M. Mehring, Hyperfine Sublevel Correlation (HYSCORE) Spectroscopy – A 2D Electron-Spin-Resonance Investigation of the Squaric Acid Radical. *Chem. Phys. Lett.*, 132 (1986) 279-282.
50. W. B. Mims, Pulsed ENDOR experiments, *Proc. Roy. Soc. London Ser. A* 283 (1965) 452-457.
51. S. Stoll and A. Schweiger, EasySpin, a Comprehensive Software Package for Spectral Simulation and Analysis in EPR. *J. Magn. Reson.* 178 (2006) 42-55.

52. W. R. Hagen, Direct Electron-Transfer of redox proteins at the bare glassy-carbon electrode, *Eur. J. Biochem.* 182 (1989) 523-530.
53. J. Haladjian, P. Bianco, F. Nunzi, M. Bruschi, A permselective-membrane electrode for the electrochemical study of redox proteins - Application to Cytochrome c (552) from *Thiobacillus-ferrooxidans*, *Anal. Chim. Acta* 289 (1994) 15-20.
54. A. Sali, Comparative Protein Modeling By Satisfaction Of Spatial Restraints, *Mol. Med. Today* 1 (1995) 270-277.
55. E. F. Pettersen, T. D. Goddard, C. C. Huang, G. S. Couch, D. M. Greenblatt, E. C. Meng, T. E. Ferrin, UCSF chimera - A visualization system for exploratory research and analysis, *J. Comput. Chem.* 25 (2004) 1605-1612.
56. R. A. Laskowski, M. W. Macarthur, D. S. Moss, J. M. Thornton, PROCHECK - A Program to check the stereochemical quality of protein structures, *J. Appl. Crystallogr.* 26 (1993) 283-291.
57. J. U. Bowie, R. Luthy, D. Eisenberg, A method to identify protein sequences that fold into a known three-dimensional structure, *Science* 253 (1991) 164-170.
58. D. Eisenberg, R. Luthy, J. U. Bowie, VERIFY3D: Assessment of protein models with three-dimensional profiles, *Methods Enzymol.* 277 (1997) 396-404.
59. M. J. Sippl, Recognition of errors in three-dimensional structures of proteins, *Proteins: Struct. Funct. Bioinform.* 17 (1993) 355-362.
60. M. Wiederstein, M.J. Sippl, ProSA-web: interactive web service for the recognition of errors in three-dimensional structures of proteins, *Nucleic Acids Res.* 35 (2007) W407-W410.
61. S. C. Shin, S. J. Kim, J. K. Lee, D. H. Ahn, M. G. Kim, H. Lee, J. Lee, B.-K. Kim, H. Park, Transcriptomics and Comparative Analysis of Three Antarctic Notothenioid Fishes, *PLoS One* 7 (2012) e43762.
62. A. Coppe, C. Agostini, I. A. M. Marino, L. Zane, L. Bargelloni, S. Bortoluzzi, T. Patarnello, Genome Evolution in the Cold: Antarctic Icefish Muscle Transcriptome Reveals Selective Duplications Increasing Mitochondrial Function, *Genome Biol. Evol.* 5 (2013) 45-60.

63. A. Roesner, C. Fuchs, T. Hankeln, T. Burmester, A globin gene of ancient evolutionary origin in lower vertebrates: Evidence for two distinct globin families in animals, *Mol. Biol. Evol.* 22 (2005) 12-20.
64. S. Z. Hu, K. M. Smith, T. G. Spiro, Assignment of protoheme Resonance Raman spectrum by heme labeling in myoglobin, *J. Am. Chem. Soc.* 118 (1996) 12638-12646.
65. S. Choi, T. G. Spiro, K. C. Langry, K. M. Smith, D. L. Budd, G. N. Lamar, Structural correlations and vinyl influences in Resonance Raman Spectra of protoheme complexes and proteins, *J. Am. Chem. Soc.* 104 (1982) 4345-4351.
66. A. I. Ioanitescu, S. Dewilde, L. Kiger, M. C. Marden, L. Moens, S. Van Doorslaer, Characterization of nonsymbiotic tomato hemoglobin, *Biophys. J.* 89 (2005) 2628-2639.
67. T. Uchida, E. Sato, A. Sato, I. Sagami, T. Shimizu, T. Kitagawa, CO-dependent activity-controlling mechanism of heme-containing CO-sensor protein, neuronal PAS domain protein, *J. Biol. Chem.* 280 (2005) 21358-21368.
68. E. Vinck, E., S. Van Doorslaer, Analysing low-spin ferric complexes using pulse EPR techniques: a structure determination of bis (4-methylimidazole) (tetraphenylporphyrinato)iron(III), *Phys. Chem. Chem. Phys.* 6 (2004) 5324-5330.
69. F. Desmet, L. Thijs, H. El Mkami, S. Dewilde, L. Moens, G. Smith, S. Van Doorslaer, The heme pocket of the globin domain of the globin-coupled sensor of *Geobacter sulfurreducens* - An EPR study, *J. Inorg. Biochem.* 104 (2010) 1022-1028.
70. P. Corti, M. Ieraci, J. Tejero, Characterization of zebrafish neuroglobin and cytoglobins 1 and 2: Zebrafish cytoglobins provide insights into the transition from six-coordinate to five-coordinate globins, *Nitric Oxide* 53 (2016) 22-34.
71. S. Han, D. L. Rousseau, G. Giacometti, M. Brunori, M. Metastable intermediates in myoglobin at low pH, *Proc. Natl. Acad. Sci. U. S. A.* 87 (1990) 205-209.
72. N. T. Yu, Resonance Raman studies of ligand-binding, *Methods Enzymol.* 130 (1986) 350-409.
73. H. Sawai, M. Makino, Y. Mizutani, T. Ohta, H. Sugimoto, T. Uno, N. Kawada, K. Yoshizato, T. Kitagawa, Y. Shiro, Structural characterization of the proximal and distal histidine environment of cytoglobin and neuroglobin, *Biochemistry* 44 (2005) 13257-13265.

74. A. Riccio, G. Mangiapia, D. Giordano, A. Flagiello, R. Tedesco, S. Bruno, A. Vergara, L. Mazzearella, G. di Prisco, P. Pucci, L. Paduano, C. Verde Polymerization of hemoglobins in Arctic fis: *Lycodes reticulatus* and *Gadus morhua*. *IUBMB Life* 63 (2011) 346-354.
75. P. Halder, J. T., III Trent, M. S. Hargrove, Influence of the protein matrix on intramolecular histidine ligation in ferric and ferrous hexacoordinate hemoglobins, *Proteins: Struct. Funct. Bioinform.* 66 (2007) 172-182.
76. W. Van Leuven, B. Cuypers, F. Desmet, D. Giordano, C. Verde, L. Moens, S. Van Doorslaer, S. Dewilde, Is the heme pocket region modulated by disulfide-bridge formation in fish and amphibian neuroglobins as in humans?, *Biochim. Biophys. Acta-Proteins and Proteomics* 1834 (2013) 1757-1763.
77. A. Vergara, M. Franzese, A. Merlino, L. Vitagliano, C. Verde, G. di Prisco, H. C. Lee, J. Peisach, L. Mazzearella, Structural characterization of ferric hemoglobins from three Antarctic fish species of the suborder notothenioidei, *Biophys. J.* 93 (2007) 2822-2829.