Available online at www.sciencedirect.com
ScienceDirect
Gene 398 (2007) 69–77
GENE
www.elsevier.com/locate/gene

The truncated hemoglobins in the Antarctic psychrophilic bacterium *Pseudoalteromonas haloplanktis* TAC125

Daniela Giordano [a], Ermenegilda Parrilli [b], Agnes Dettaï [c], Roberta Russo [a], Giulia Barbiero [a], Gennaro Marino [b], Guillaume Lecointre [c], Guido di Prisco [a], Luisa Tutino [b], Cinzia Verde [a,*]

[a] *Institute of Protein Biochemistry, CNR, Via Pietro Castellino 111, I-80131 Naples, Italy*
[b] *Dipartimento di Chimica Organica e Biochimica, Università degli Studi di Napoli "Federico II", Complesso Universitario Monte S. Angelo, Napoli, Italy*
[c] *UMR 7138 CNRS, Département Systématique et Evolution, Muséum National d'Histoire Naturelle, CP26, 57, rue Cuvier, 75005 Paris, France*

Received 9 January 2006; received in revised form 23 January 2007; accepted 26 February 2007
Available online 25 April 2007

Abstract

The genome sequence of the Antarctic Gram-negative marine eubacterium *Pseudoalteromonas haloplanktis* TAC125 is a potential source of useful data on proteins from a cold-adapted microorganism. Identifying the bases of protein adaptation to higher or lower temperatures is important to understand the relationship between structure/function and life history on the Earth. The *P. haloplanktis* TAC125 genome contains three genes in distinct positions on chromosome I, named *PSHAa0030, PSHAa2217* and *PSHAa0458*. These genes encode three truncated hemoglobins. The amino-acid identity between the three hemoglobins is less than 25% suggesting that these proteins may have different function(s) in bacterial cellular metabolism.

The hemoglobin encoded by the *PSHAa0030* gene has been cloned, expressed in *Escherichia coli*, purified and structurally characterised. This truncated hemoglobin is monomeric; circular dichroism shows high temperature resistance. The optical spectra of oxygenated and CO forms are similar to those of other truncated hemoglobins.

Phylogenetic analyses show that two truncated globins encoded by the *PSHAa0030* and *PSHAa2217* genes belong to group II, and the third one encoded by *PSHAa0458* to group I.

Keywords: Antarctica; Bacterium; Genome; Truncated-globin gene; Cold-adaptation; Evolution

1. Introduction

Nine tenths of the Earth's biosphere is cold, and most of it is marine, i.e. most of the ocean waters are permanently at temperatures of 5 °C or slightly above this value. Although some marine bacteria have been studied, we only have limited knowledge about life at low temperatures in sea water.

The main challenges that bacteria have to face in cold marine environments are: (i) exponential decrease of biochemical reaction rates, (ii) increased viscosity of aqueous environments, and (iii) increased gas solubility and radical stability (D'Amico et al., 2006).

The genome of the Antarctic Gram-negative marine eubacterium *Pseudoalteromonas haloplanktis* TAC125 was recently sequenced and annotated (Medigue et al., 2005). The information gathered by this analysis contributed to shed light on several molecular features selectively developed in cold environments.

In *P. haloplanktis* TAC125, as observed in other psychrophilic bacteria, proteins are crucial in cold adaptation, since they are involved in the most important cell functions. In general, at

Abbreviations: Hb, hemoglobin; trHb, truncated hemoglobin; *PSHAa*, *Pseudoalteromonas haloplanktis* chromosome I; *PSHAa0030*, *PSHAa0458*, *PSHAa2217*, gene positions on *Pseudoalteromonas haloplanktis* chromosome I; *Ph*-trHbO, *Pseudoalteromonas haloplanktis* TAC125 truncated hemoglobin (encoded by *PSHAa0030* gene); flavoHb, flavohemoglobin; Mb, myoglobin; SW-Mb, sperm-whale myoglobin; β-ME, β-mercaptoethanol; CD, circular dichroism.

* Corresponding author. Tel./fax: +39 081 6132710.
E-mail address: c.verde@ibp.cnr.it (C. Verde).

0378-1119/$ - see front matter © 2007 Elsevier B.V. All rights reserved.
doi:10.1016/j.gene.2007.02.037

low temperature, best performances are made possible through peculiar features which altogether lead to higher flexibility of key parts of the molecular structure of many proteins (D'Amico et al., 2006).

In the evolutionary history of *P. haloplanktis* TAC125, special features linked to life at low temperature have been selected. One of the most interesting examples is related to increased gas solubility. At low temperatures, the enhanced solubility of O_2 significantly increases the production rate of toxic reactive O_2 species (ROS). *Colwellia psychrerythraea*, an obligately psychrophilic Arctic bacterium (Methe et al., 2005), and *Desulfotalea psychrophila*, a marine sulfate-reducing δ-proteobacterium able to grow at *in situ* temperatures below 0 °C (Rabus et al., 2004), seem to have survived this environmental challenge by developing an enhanced antioxidant capacity owing to the presence of several genes that encode catalases and superoxide dismutases. In contrast, the genome sequence reveals that *P. haloplanktis* TAC125 copes with increased O_2 solubility by enhancing production of O_2-scavenging enzymes and deleting entire metabolic pathways, such as those which generate ROS as side products. The remarkable deletion of the ubiquitous molybdopterin-dependent metabolism in the *P. haloplanktis* TAC125 genome (Medigue et al., 2005) can be seen in this perspective.

Bacteria may express three types of hemoglobins (Hbs), i.e. truncated (trHbs), monomeric (Hbs) and flavohemoglobins (flavoHbs). Similar to other bacteria, the *P. haloplanktis* TAC125 genome contains multiple genes encoding trHbs (annotated as *PSHAa0030, PSHAa0458, PSHAa2217*) and a flavoHb gene (*PSHAa2880*), suggesting that specific and different functions may be associated to these two classes of proteins. Interestingly, the *C. psychrerythraea* genome does not contain genes encoding trHbs.

TrHbs are small O_2-binding hemoproteins, generally shorter than vertebrate Hbs by 20–40 amino-acid residues (Pesce et al., 2000). TrHbs show very low amino-acid-sequence homology to vertebrate and non-vertebrate Hbs with few residues conserved throughout the structure. In trHbs the globin fold is based on a "2 on 2" α-helical sandwich edited from the "3 on 3" typical of vertebrate globins (Pesce et al., 2000). Modifications of the classical "3 on 3" fold occur at helix A (almost entirely deleted in all trHbs), and in the CD-D and EF-F regions. The original phylogenetic analysis of trHbs classified them into three groups, denoted I (trHbN), II (trHbO), and III (trHbP) (Wittenberg et al., 2002).

In mammals, the proteins devoted to O_2 utilisation are the homologues of human Hb and myoglobin (Mb). Mb binds O_2 for intracellular storage whereas Hb transports O_2 to tissues. Vertebrate Hb is a tetramer of two identical α and two identical β globins. The movements and interactions between the α and β subunits lead to the cooperative behaviour. The amino-acid sequences of α and β globins are approximately 50% similar, suggesting that the two genes share a common ancestor present since 450 million years ago in the ancestral jawed vertebrate (Goodman et al., 1987). In fact, the recent discovery of Hbs in virtually all kingdoms has clearly shown that the ancestral gene encoding Hb is very ancient and that Hb does perform other functions beside O_2 transport (Hardison, 1998). Several species containing trHbs exhibit pathogenic properties, fix nitrogen or display distinctive metabolic properties.

The few trHbs studied in detail seem to support these observations (Milani et al., 2005). In contrast to vertebrate Hbs, in most of which the heme pocket is conserved, trHbs display a remarkable variability, especially in the heme-pocket distal side (Wittenberg et al., 2002). All these trHbs are able to bind diatomic ligands such as O_2, CO, and NO, with different affinities (Milani et al., 2005). The high O_2 affinity displayed by most trHbs makes their role as O_2 transporters very unlikely (Wittenberg et al., 2002; Ouellet et al., 2003). Several other functions have been proposed, e.g. trHbs have been suggested to be involved in the response to oxidative and nitrosative stress (Couture et al., 1999) or to function as O_2 sensors (Wittenberg et al., 2002).

In the present work, three genes encoding trHb and one encoding a flavoHb have been identified in *P. haloplanktis* TAC125. Their amino-acid sequences have been compared with a large variety of Hb sequences to explore their assignment to groups of paralogues. Phylogenetic analyses showed that two truncated globins encoded by the *PSHAa0030 and PSHAa2217* genes belong to group II, and the third one encoded by *PSHAa0458* to group I. The *PSHAa0030* gene encoding a group II trHb (hereafter called *Ph*-trHbO) was cloned and over-expressed in *Escherichia coli*. The recombinant protein was purified and its thermal features were characterised by circular dichroism (CD).

2. Materials and methods

2.1. Computational search and identification of globin sequences

The genome sequence of *P. haloplanktis* TAC125 was accessed through the MaGe annotation platform (http://www.genoscope.cns.fr/agc/mage/wwwpkgdb/Login/log.php?pid=7#ancreLogin) and was searched for the presence of putative bacterial Hbs, trHbs and flavoHbs by using the Explore functionality. The database-search criterion was the presence of a globin-like domain, as described by the InterPro family motif IPR009050 http://www.ebi.ac.uk/interpro/IEntry?ac=IPR009050). Four genes were selected for further analysis, all of them located on the larger chromosome I but in different positions. They are *PSHAa0030*, *PSHAa0458*, *PSHAa2217* and *PSHAa2880* (where "*PSHA*" stands for *P. haloplanktis*, "*a*" denotes the chromosome I and the figure indicates the gene position). The present annotation of the selected genes, as displayed in the MaGe database, indicates that the product *Ph*-trHbO encoded by the *PSHAa0030* gene is a "putative Hb-like O_2-binding protein", that *PSHAa0458* is predicted to encode a "putative protozoan/cyanobacterial globin family protein", and that *PSHAa2217* codes for a "conserved protein of unknown function, putative globin-like family". The protein encoded by *PSHAa2880* was classified as a two-domains flavoHb, since it contains the heme-containing O_2-binding domain in the N-terminal region and a FAD-containing reductase domain in the C-terminal region, which is homologous to the InterPro

IPR001433 oxidoreductase FAD/NAD(P)-binding motif (accessible through http://www.ebi.ac.uk/interpro/IEntry?ac=IPR001433).

2.2. Bacterial strains, plasmid and culture condition

The *E. coli* BL21(DE3) (Novagen) strain was routinely used for cloning and expressing recombinant genes. Cells were grown in Luria–Bertani (LB) medium at 25 °C. When required, kanamycin sulfate (Sigma) was added at 50 μg/ml. Plasmid pET-28a was utilised for cloning and expression. Restriction and modifying enzymes were obtained from Promega. The oligonucleotides were custom synthesised from PRIMM.

2.3. Cloning and expression of the PSHAa0030 gene

The primer pairs for the *PSHAa0030* gene (oligoHNtrg(II)-5′ TAAGCTTCCATGGTTAAACGACTTTTTTC 3′ and oligoBtrg(II) rv-5′ AGGATCCGTTAATGCTGGTTAATCATATGG 3′) were designed on the basis of the *P. haloplanktis* TAC125-genome sequence (Medigue et al., 2005). Sequences corresponding to the *NcoI* site and a *BamHI* site were introduced in the forward and reverse primers, respectively. The amplifications were performed in a mixture containing 80 ng of *P. haloplanktis* TAC125-genomic DNA as template, 50 pmol of each oligonucleotide primer, 1.8 mM $MgCl_2$, 50 mM KCl, 20 mM Tris–HCl pH 8.3, 0.1% gelatine, 200 μM dNTP in a final volume of 50 μl. The mixtures were incubated at 95 °C for 10 min, then 1.25 units of *Taq* DNA polymerase were added. Twenty cycles of amplification (consisting of 1 min at 95 °C, 1.5 min at 60 °C and 1 min plus 5 s/cycle at 72 °C) were carried out and followed by a cycle in which the extension reaction at 72 °C was prolonged for 15 min in order to complete DNA synthesis. The amplified fragment was cloned and its nucleotide sequence checked to rule out the occurrence of any mutation during synthesis.

The *Nco*I-*Bam*HI-digested fragment of the *PSHAa0030* gene was further subcloned into the corresponding sites of the expression vector pET-28a and subsequently transformed into *E. coli* BL21(DE3), obtaining the plasmid pET-trHbO. For the overexpression of cold-adapted trHb in *E. coli*, a single colony carrying the respective plasmid construct was inoculated in LB medium supplemented with kanamycin (50 μg/ml) and allowed to grow at 37 °C until A_{600} reached ~ 0.6 OD. The culture was then induced with 1 mM isopropyl-β-D-thiogalactopyranoside and further incubated for another 12 h at 25 °C in the presence of 0.3 mM of D-aminolevulinic acid. The expression of cloned protein was monitored after running the cell lysate of recombinant strains on 15% SDS-PAGE followed by Coomassie Brilliant Blue staining.

2.4. Purification of Ph-trHbO

All buffers were in MilliQ water. The cells were harvested by centrifugation at 10,000 rpm for 15 min at 4 °C, resuspended in a minimal volume of lysis buffer [50 mM Tris–HCl pH 8.0, 1.0 mM EDTA, 1.0 mM β-mercaptoethanol (β-ME), 1.0 mM phenylmethylsulfonylfluoride and protease-inhibitor cocktail (SIGMA P8465)] and sonically disrupted until the supernatant was reddish and clear (Giangiacomo et al., 2005). The cell debris were removed by centrifugation at 14,000 rpm for 1 h at 4 °C. The supernatant was loaded on a DEAE-cellulose column (Whatman International Ltd., Maidstone, UK), equilibrated with 20 mM Tris–HCl pH 7.6, 1.0 mM β-ME and 1.0 mM EDTA [Akta Explorer system (GE Healthcare Biosciences, Amersham Biosciences Ltd, U.K.)].

Ph-trHbO was eluted with a NaCl gradient (from 0 to 0.25 M) in 20 mM Tris–HCl pH 7.6, 1.0 mM EDTA and 1.0 mM β-ME (Giangiacomo et al., 2005). The fraction was chosen on the basis of heme absorbance at 418 nm and 546 nm and protein absorbance at 280 nm. The eluate was concentrated using Ultracel YM-3 Centricon Millipore with a molecular mass cut-off of 3 kDa, dialysed against 20 mM Tris–HCl pH 7.6, 1.0 mM EDTA and 1.0 mM β-ME, and further purified by gel filtration. A Superdex-200 column (1.6 × 60 cm, GE Healthcare Biosciences, Amersham Biosciences Ltd, U.K.) was equilibrated with 20 mM Tris–HCl pH 7.6 containing 0.2 M NaCl, 1.0 mM EDTA and 1.0 mM β-ME (Wainwright et al., 2006). The protein obtained was > 98% pure on SDS-PAGE; it was stored at 4 °C. The N-terminal sequence was elucidated by automatic sequencing performed with an Applied Biosystems Procise 494 automatic sequencer, equipped with on-line detection of phenylthiohydantoin amino acids.

2.5. Spectrophotometry

Visible absorption spectra were acquired on a Cary 300 UV–visible spectrophotometer (Varian) at room temperature. In order to obtain fully oxygenated spectra, *Ph*-trHbO was reduced with a few crystals of Na dithionite for 15 min. CO binding was achieved by gently bubbling the gas into reduced samples for 2 min. The side-reaction products and excess reagent were removed on a Sephadex G-25 column (Pharmacia) equilibrated with 20 mM Tris–HCl pH 7.6, 1.0 mM EDTA and 1.0 mM β-ME. To obtain deoxygenated spectra, excess solid dithionite was added to the oxygenated protein.

2.6. Thermostability

CD spectra were obtained on a Jasco J-715 spectropolarimeter with 350 μl of 0.1 μg/μl protein in 5 mM Tris–HCl pH 7.6. Hellma quartz cells of 0.1-cm-path length were used in the far UV (190–250 nm). The temperature of the sample cell was regulated by a PTC-348 WI thermostat, under constant nitrogen flow.

Thermal denaturation experiments were performed by raising the cell temperature from 2 °C to 99 °C at 0.5 °C min^{-1} and monitoring the CD signal at 222 nm. Spectra were monitored in the far UV during heating from 2 °C to 99 °C. They were signal-averaged by adding three scans and baseline corrected by subtracting a buffer spectrum. The sample was then cooled back to 2 °C in order to monitor the final folding after denaturation. Melting points were identified by taking the maximum of the first-derivative plot and the zero of the second-derivative plot.

2.7. Sequence alignment and phylogenetic analyses

Some trHb sequences were recovered through keyword searches and blast searches in GenBank (Altschul et al., 1990), but most sequences and all the outgroup sequences were extracted from either the sequences of Vuletich and Lecomte (2006) or the additional data for Fig. 3 of Vinogradov et al. (2006); most accession numbers and sequences were re-checked with the deposited originals. Duplicates present in several of the datasets were eliminated. All accession numbers of sequences are in Table 1 (Supplementary Data). A high number (130) of outgroup sequences were included to represent all groups of globins. The purpose was to ensure that the placement of the three sequences of *P. haloplanktis* TAC125 did

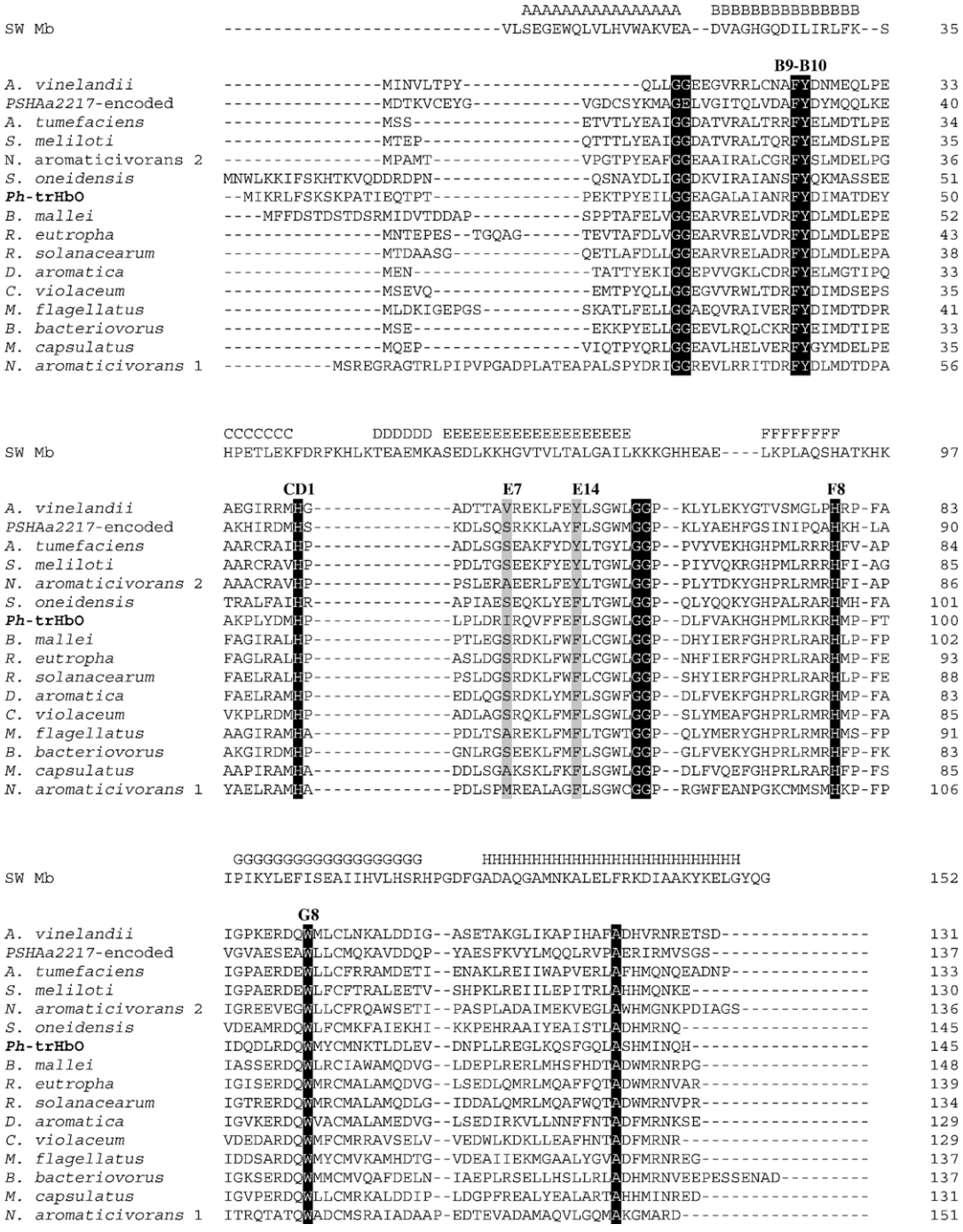

Fig. 1. Sequence alignment of representative trHbs (group II). The sequence comparison was carried out by Clustal W (Thompson et al., 1997) and manual adjustments are based on known crystal structures. Amino-acid residues have been written using one-letter codes. The helical positions of some residues, designated on the globin-fold topological positions of sperm-whale Mb (SW-Mb) are denoted on the alignment top. Identical functionally important residues in group II trHbs are in black and similar functionally important residues are in grey. The sequences from other species than *P. haloplanktis* TAC125 are from Vuletich and Lecomte (2006). This is not the same alignment analysed in the phylogenetic study; the latter is available upon request. For species full names, see Table 1 (Supplementary Data).

not simply result from the failure to include their sister group. Alignment was made using that of Vinogradov et al. (2006) as a basis, and then adjusted by hand. The beginning of regions A and H of the globins and all sequences of *Chaenorhabditis elegans* from Vinogradov et al. (2006) were not included in the analysis because of alignment problems. The final alignment is available upon request. The sequences were analysed using PAUP* (Swofford, 2001). Maximum-parsimony analysis was performed through 1000 random addition sequence replicates with TBR search, and a second search with an equal number of replicates to check that the most parsimonious tree had been found. One thousand bootstrap replicates were also computed. Bootstrap values over 50 are indicated on the strict consensus of the most parsimonious trees.

3. Results and discussion

3.1. Primary structure

Fig. 1 shows the alignment of *Ph*-trHbO with other representative bacterial globins of group II. *Ph*-trHbO is closely related to the other trHbs. The sequence identity between *Ph*-trHbO and the protein encoded by the *PSHAa2217* gene (belonging to the same group) is 24%. The identity between trHbs from different groups is generally low, but may be higher within a given group. All trHbs, without exception, retain His F8 as the proximal ligand to the heme. Distal Tyr B10 participates in heme stabilisation (Wittenberg et al., 2002) and is almost invariant in all trHbs (Milani et al., 2005). Position CD1 hosts Phe in group I and group III trHbs, whereas in group II trHbs there may be Phe, Tyr or His. Phe is often present at E14 fulfilling a role comparable to that of Phe CD1 in other Hbs (Wittenberg et al., 2002). Similar to all trHbs known to date, *Ph*-trHbO has His F8. Position E7, more variable in group II, is occupied by Ile.

At the B9-B10 positions there is the strongly conserved Phe-Tyr motif. His in CD1 is involved in hydrogen bonding. In other trHbs belonging to the same group, CD1 may be occupied by Tyr or Phe. According to replacement studies in *Mycobacterium tuberculosis* (*Mt*-trHbO), when wild-type Tyr is replaced by Phe in CD1, Tyr B10 becomes a H-bond donor (Ouellet et al., 2003). This happens in wild-type *Bacillus subtilis* trHbII (*Bs*-trHbO), also displaying a Leu (arrow) Gln replacement at E11 (Giangiacomo et al., 2005). At E11, *Ph*-trHbO has hydrophobic Phe.

The crystal structures of the two cyanomet trHbO solved to date (Milani et al., 2003; Giangiacomo et al., 2005) have revealed a link between CD1 and E11. All trHbO having Tyr (and not His) at CD1 have a non-polar residue at E11. Tyr B10 and Trp G8 are conserved in all trHbO investigated. The arginyl residues at F5 and F7 are conserved in *Ph*-trHbO, but are replaced by Pro and Ala, respectively in the other trHb of *P. haloplanktis* belonging to the same group (encoded by *PSHAa2217*).

In *Ph*-trHbO, there are the two conserved Gly-Gly motifs located in the AB and EF hinges, the latter at the end of the pre-F loop, involved in stabilising the globin fold (Pesce et al., 2000). These structural glycyl residues are well conserved in groups I and II, but they are absent in group III (Nardini et al., 2006).

3.2. Expression of the PSHAa0030 gene

PSHAa0030 from *P. haloplanktis* TAC125 was expressed in *E. coli* under T7 promoter. This procedure resulted in the accumulation of heme protein inside the cell yielding a reddish-brown colour to the recombinant *E. coli* cells. SDS-PAGE confirmed the presence of a ~ 17 kDa protein corresponding to the expected size of *Ph-trHbO* (data not shown).

3.3. Purification of Ph-trHbO and optical properties

Ph-trHbO was cloned and over-expressed in *E. coli*, then purified by ion-exchange and gel-filtration chromatographies (not shown). The latter step did not provide evidence for the presence of a dimeric species, and the molecular mass was ~ 17 kDa in agreement with the primary structure.

Partially purified *Ph*-trHbO by DEAE-cellulose chromatography underwent unexpected proteolytic cleavage, probably occurring in solution. SDS/PAGE revealed two bands, one of which displaying a lower molecular mass. Mass spectrometry (MALDI/MS and electrospray mass spectrometric measurements; not shown) confirmed the presence of two components of molecular masses 16.0 kDa, and 16.79 kDa in accordance with the sequence. The presence of the former additional band was attributed to proteolysis of the native protein by trace amounts of proteases co-purified with *Ph*-trHbO. Direct sequencing showed two N-terminal sequences, namely that of the full-length protein in higher amount, and that of the cleaved form in which the first six residues had been removed. Addition of protease-inhibitor cocktails in the lysis buffer was not sufficient to fully inhibit cleavage.

The *Ph*-trHbO fraction from gel-filtration chromatography (native form) was a mixture of the ferric and ferrous forms (data not shown). The ferric form was slowly reduced by Na dithionite over ~ 15 min. The addition of carbon monoxide to the

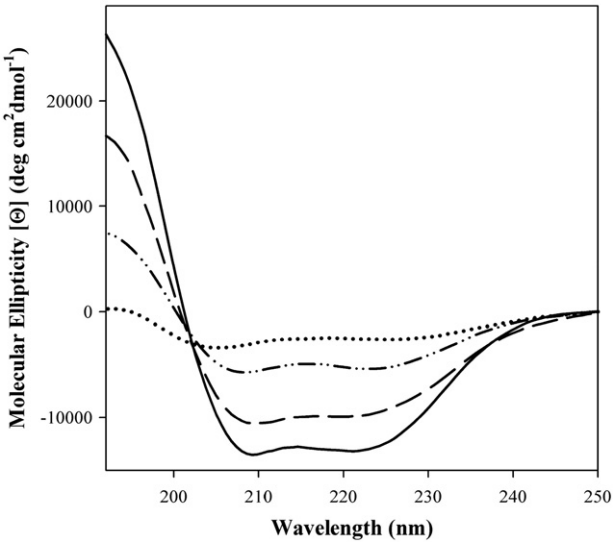

Fig. 2. Temperature-dependent CD measurements of *Ph*-trHbO. Far-UV CD spectra at 2 °C (solid line), 70 °C (long-dash line), heated to 99 °C (dotted line), cooled back to 2 °C (dash–dot–dot line). The amount of CD is expressed as molecular ellipticity per mean residue [Θ].

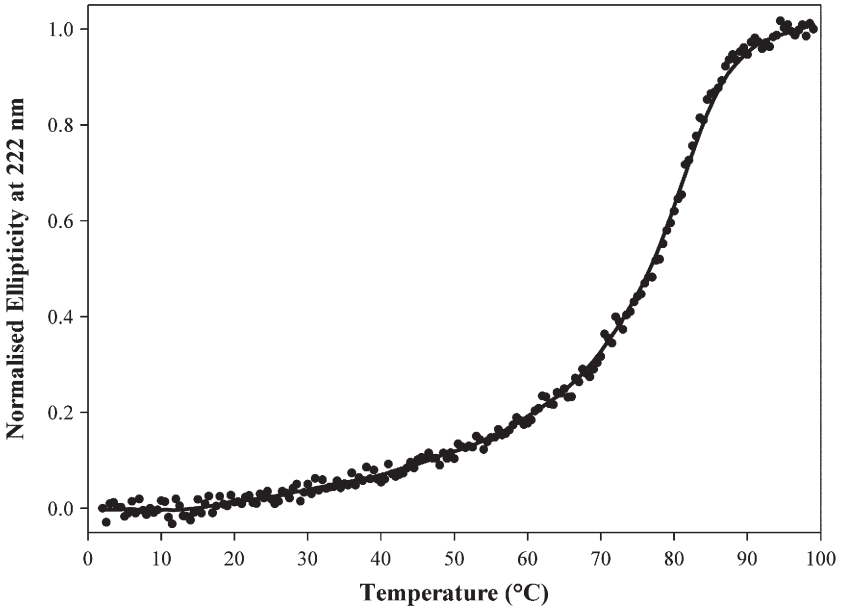

Fig. 3. Thermal denaturation of *Ph*-trHbO. Far-UV CD melting profiles recorded at 222 nm. Normalised ellipticity corresponds to $\Delta\Theta_{max}$. FS is the difference between molecular ellipticity at a given temperature and molecular ellipticity measured at the minimal temperature. $\Delta\Theta_{max}$ is the difference between molecular ellipticities at 99 °C and 2 °C.

reduced protein yielded the carbomonoxy ferrous-heme form (CO-*Ph*-trHbO). It was characterised by a Soret peak at 422 nm and two peaks at 538 and 569 nm, corresponding to the low-spin configuration bands.

3.4. Thermostability and resistance to thermal denaturation

The thermal stability of *Ph*-trHbO was monitored by CD in the temperature range 2 °C–99 °C. The α-helical signal gradually decreased as the protein was slowly heated from 2 °C to 99 °C (0.5 °C min^{-1}), even though *Ph*-trHbO still contained high α-helical content at 70 °C, where 75% of the signal at 222 nm is still present (Fig. 2). After cooling back from 99 °C to 2 °C, the spectrum was not completely superimposable to that obtained from the initial acquisition at 2 °C, demonstrating that unfolding is largely irreversible.

To further characterise the thermal stability of *Ph*-trHbO, the CD signal converted to normalised molecular ellipticity at

```
                 ....|....| ....|....| ....|....| ....|....| ....|....| ....|....| ....|....| ....|....|
                     5          15         25         35         45         55         65         75
Microbdeg1       MFSAIRTVSI GMIATLLLLG CSSLQGQNQT LYKQLGG[-] QDGVSALTRQ LLINMAADN- RIAPR---[-] FRGV[-]--
PSHAa0458        ----MRVFIL FLLATSVLAG CSSKP--KQS LYQQING[-] NAGIEKLVDS FIYQIGNDR- QVFHY---[-] FEHS[-]--
Methylccap       ---------- ----MQEPV- -------IQT PYQRLGG[-] EAVLHELVER FYGYMDELPE AAPIRA--[-] MHAD[-]--
PSHAa2217        ---------- ----MDTKVC ---EYGVGDC SYKMAGE[-] LVGITQLVDA FYDYMQQLKE AKHIRDMH[-] S---[-]--
Shewon           ----MNWLKK IF--SKHTKV QDDRDPNQSN AYDLIGG[-] DKVIRAIANS FYQKMASSEE TRALFA--[-] IHRA[-]--
Ph-trHbO         ------MIKR LFSKSKPATI -EQTPTPEKT PYEILGG[-] EAGALAIANR FYDIMATDEY AKPLYD--[-] MHPL[-]--
Mesorhiz3        ---------- ----MSKQSA GRTTIIVDGV PLPEILD[-] ERMIHDVVHG FYEEIRNDD- LLGPI---[-] FN--[-]GA

                 ....|....| ....|....| ....|....| ....|....| ....|....| ....|....| ....|....| ....|....|
                     85         95        105        115        125        135        145        155
Microbdeg1       ----NIGKFK TGLDTYLCSI TDGG[-]--- CVYGG-DSIK TIHSGYN-[-] Y[-]TATEF NALVENLMQA METLQIP[-]
PSHAa0458        ----NITHFR QGFISHLCAL TNGP[-]--- CEYKG-DSMV AIHTGMN-[-] I[-]NEKDF NHVVDLLINA MDEQSIP[-]
Methylccap       ----DLSGAK SKLFKFLSGW LGGP[-]DLF VQEFGHPRLR ARHFPFS-[-] I[-]GVPER DQWLLCMRKA LDDIPL-[-]
PSHAa2217        ---KDLSQSR KKLAYFLSGW MGGP[-]KLY AEHFGSINIP QAHKHLA-[-] V[-]GVAES EAWLLCMQKA VDDQPYA[-]
Shewon           ----PIAESE QKLYEFLTGW LGGP[-]QLY QQKYGHPALR ARHMHFA-[-] V[-]DEAMR DQWLFCMKFA IEKHIK-[-]
Ph-trHbO         ----PLDRIR QVFFEFLSGW LGGP[-]DLF VAKHGHPMLR KRHMPFT-[-] I[-]DQDLR DQWMYCMNKT LDLEVD-[-]
Mesorhiz3        INSEAWPHHL AKMCDFWSAT LLRT[-]--- KRYEG--RPL PPHLAIPG[-] L[-]GEVHF RRWLMLFRAT VDRLCPA[-]

                 ....|....| ....|....| ....|....
                     165        175        185
Microbdeg1       --TATQNKLL AKLAPSYQ-D VVYQ-----
PSHAa0458        --HPVQNKII SKMAPLRS-E IIKI-----
Methylccap       -DGPFREALY EALARTAH-H MINRED---
PSHAa2217        --ESFKVYLM QQLRVPAE-R IRMVSGS--
Shewon           -KPEHRAAIY EAISTLAD-H MRNQ-----
Ph-trHbO            -NPLLREGLK QSFGQLAS-H MINQH----
Mesorhiz3        DVAVLFMDRA LRIAHSFR-L AVAFSRGDD
```

Fig. 4. Alignment of trHbs encoded by *PSHAa0030*, *PSHAa0458* and *PSHAa2217* with representative trHbs from the three groups. For group II and I, the putative sister groups of the trHbs sequences of *P. haloplanktis* TAC125 have been selected. The randomly selected representative of group III is *Mesorhizobium* sp. BNC1 (see also Fig. 5). [-] indicate a larger gap due to sequences not shown here. The part of the sequences not included in the analyses has been highlighted in grey. For species full names, see Table 1 (Supplementary Data).

222 nm was recorded as a function of temperature in 5 mM Tris–HCl pH 7.6, at a heating rate of 0.5 °C min^{-1}. The rise of the molecular ellipticity at 222 nm indicated loss of secondary structure. The melting profile of *Ph*-trHbO (Fig. 3) showed a sigmoidal transition at 81 °C, the value of the temperature-induced unfolding of the native structure.

Although the exceptionally high thermostability of a cold-adapted protein was unexpected, other similar cases have been reported. One example is psychrophilic L-glutamate dehydrogenase from the Antarctic bacterium *Psychrobacter* sp. TAD1, which becomes irreversibly denatured only at temperatures higher than 55 °C (Di Fraia et al., 2000), namely the same typical of the bovine enzyme.

The stability of mesophilic globins in vertebrates and invertebrates has been investigated in the past (Wittenberg et al., 2002; Hughson and Baldwin, 1989; Hughson et al., 1991). Mbs are very stable to thermal denaturation when they are in the holo-structure. In the absence of heme, some intra-protein

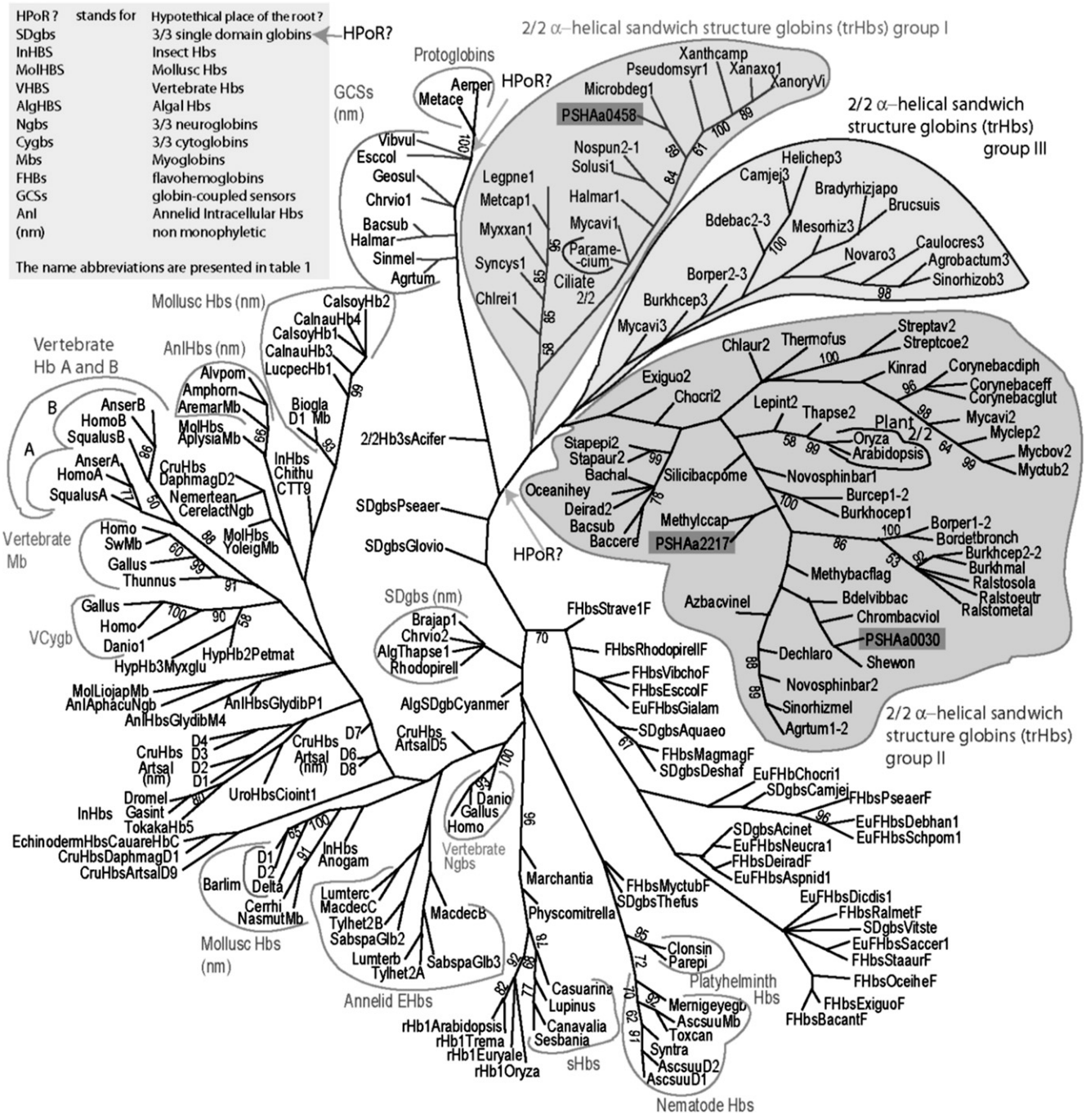

Fig. 5. Strict consensus of the 296 equi-parsimonious trees obtained with a dataset of 206 globin sequences (maximum-parsimony analysis and boostrap with PAUP* CI=0.208; RI=0.478). The branch lengths are not representative of the rate of evolution, and have been modified by hand to improve the readability of the tree. The numbers associated to the branches are bootstrap values; values under 50 are not indicated. The arrow indicates one possibility for the hypothetical outgroup. The three sequences of *P. haloplanktis* TAC125 are on darker grey background.

interactions may not be strong enough to preserve the stability of the classical eight-helix globin fold (Harrison and Blout, 1965). Apo-globins generally display lower resistance to thermal or chemical denaturation.

Hb recently identified in *Aquifex aeolicus* is called thermoglobin because it is resistant to thermal denaturation, still containing ~ 75% of its physiological α-helical content at 90 °C (Miranda et al., 2005). Similar to thermoglobin, trHb from the thermophilic actinobacterium *Thermobifida fusca* is significantly more thermostable (as expected) than its mesophilic homologue (Bonamore et al., 2005). However, it currently seems virtually impossible to ascribe thermostability of these two trHbs to substitution of some residues, or to invoke a single explanation to distinguish between mesophile and psychrophile proteins. A commonly accepted view for protein cold adaptation is the activity–stability–flexibility relationships. Active-site sequences are generally highly conserved among homologous proteins due to structural constraints in ligand/substrate binding. In some proteins, e.g. enzymes, adaptive changes in the structure may occur at subunit interfaces, distant from the active site (Johns and Somero, 2004). These alterations in the strength of subunit interactions may affect thermal stability and energy changes associated with conformational transitions due to ligand binding (Somero, 1995; D'Amico et al., 2006).

3.5. Phylogenetic analyses

Fig. 4 shows alignment of the three sequences of interest (*PSHAa0030*, *PSHAa0458*, *PSHAa 2217*) with representatives from the three groups of trHbs. Fig. 5 shows the strict consensus of the 296 equi-parsimonious trees obtained with a dataset of 206 globin sequences. The parsimony analysis yielded 296 equi-parsimonious trees of 9823 steps. The strict consensus of these trees is relatively well resolved, but few bootstrap values are above 50. The tree has not been initially rooted because the position of the root in the globin tree is still controversial. Two of the hypotheses (Freitas et al., 2005; Vinogradov et al., 2005) for the position of the root have been represented by arrows in Fig. 5. As we are merely interested in the position of the three sequences of *P. haloplanktis* TAC125, this root is of little consequence in assigning *PSHAa0030* and *PSHAa2217* to the globin group II of 2/2 α-helical sandwich structure and *PSHAa0458* to group I. The positions of these three sequences in the tree are consistent with what is known about the tertiary structure and interactions adopted by amino-acid sequences of these groups.

4. Concluding remarks

The presence of trHbs in cold-adapted *P. haloplanktis* TAC125 stimulates questions on which structural and functional features are evolved by proteins (in comparison with mesophilic and thermophilic counterparts) in response to the challenges of low temperature and high O_2 (and other ligands) concentration in extreme environments. Features of protein structure and function are progressively fitting the puzzle of cold adaptation, helping to improve our knowledge. However, although a coherent picture of the molecular changes involved in evolutionary adaptation to temperature in some proteins is emerging, much remains unknown as yet. Relatively few cold-adapted enzymes have been examined, and many important cellular processes have hardly been studied at all, although progress is likely to occur as scientists exploit new molecular approaches, and become increasingly concerned about environmental changes in polar regions (di Prisco and Verde, 2006).

Understanding the mechanisms of phenotypic response to cold exposure and the genetic regulatory mechanisms that enable organisms to live in polar habitats will offer fundamental insights into guiding principles for environmental adaptations.

Most of the studies of protein thermostability take advantage of one of two approaches. The structural/mutational approach produces a detailed portrait (although often controversial) of the interactions stabilising the high- and low-temperature-adapted proteins. However this approach, due to much labour and costs, restricts analyses to a limited number of proteins, leading to a potentially biased view of thermal stabilising mechanisms. The second approach, less expensive yet more comprehensive, uses sequence comparisons of families of homologous high- and low-temperature-adapted proteins. In the past, the latter approach has been hampered by limitations, due for instance to paucity of genomic sequences from extremophile organisms. Nowadays, the possibility to sequence whole genomes may provide the necessary amount of data, allowing to reject or accept some of the classical hypotheses currently invoked to explain protein thermal adaptation.

The first whole-genome sequences, that of Antarctic marine *P. haloplanktis* TAC125 and of obligately psychrophilic Arctic *C. psychrerythraea*, are now available. Only the former is endowed with three distinct trHbs. These molecules are thus bound to fulfill an important physiological role, which may be related to the peculiar features of the Antarctic habitat.

Preliminary comparative genome analyses of *C. psychrerythraea* (Methe et al., 2005) with related mesophilic proteobacteria genomes have allowed to pinpoint abundance of specific amino-acid residues for cold activity. Such studies are likely to lead to unequivocal conclusions more easily than similar analyses on vertebrates, due to the much higher complexity of physiology and metabolism of the latter. Consequently, temperature-sensitive proteins, such as bacterial Hbs (in part mirroring thermal adaptations encountered by species during their evolutionary histories) are targets for future research. Studies on structure/function and evolution of *Ph*-trHbO will yield additional tools for the elucidation of the general rules to environmental adaptation.

Separation of the cleaved component will be attempted in order to establish whether it is endowed with different physico-chemical stability.

Acknowledgements

We thank Angela Amoresano and Giovanni Chiappetta for the molecular mass measurements by mass spectrometry, and Vito Carratore for the skillful technical assistance in amino-acid sequencing. This study is financially supported by the Italian

National Programme for Antarctic Research (PNRA). It is in the framework of the SCAR programme Evolution and Biodiversity in the Antarctic (EBA).

Appendix A. Supplementary data

Supplementary data associated with this article can be found, in the online version, at doi:10.1016/j.gene.2007.02.037.

References

Altschul, S.F., Gish, W., Miller, W., Myers, E.W., Lipman, D.J., 1990. Basic local alignment search tool. J. Mol. Biol. 215, 403–410.

Bonamore, A., Ilari, A., Giangiacomo, L., Bellelli, A., Morea, V., Boffi, A., 2005. A novel thermostable hemoglobin from the actinobacterium *Thermobifida fusca*. FEBS J. 272, 4189–4201.

Couture, M., et al., 1999. A cooperative O2-binding haemoglobin from *Mycobacterium tuberculosis*. Proc. Natl. Acad. Sci. U. S. A. 96, 11223–11228.

D'Amico, S., Collins, T., Marx, J.C., Feller, G., Gerday, C., 2006. Psychrophilic microorganisms: challenges for life. EMBO Rep. 7, 385–389.

Di Fraia, R., et al., 2000. $NADP^+$-dependent glutamate dehydrogenase in the Antarctic psychrotolerant bacterium *Psychrobacter* sp. TAD1. Eur. J. Biochem. 267, 121–131.

di Prisco, G., Verde, C., 2006. Predicting the impacts of climate change on the evolutionary adaptations of polar fish. Rev. Environ. Sci. Biotechnol. 5, 309–321.

Freitas, T.A., Saito, J.A., Hou, S., Alam, M., 2005. Globin coupled sensors, protoglobins, and the last universal common ancestor. J. Inorg. Biochem. 99, 23–33.

Giangiacomo, L., Ilari, A., Boffi, A., Morea, V., Chiancone, E., 2005. The truncated oxygen-avid hemoglobin from *Bacillus subtilis*. J. Biol. Chem. 280, 9192–9202.

Goodman, M., Czelusniak, J., Koop, B., Tagle, D., Slightom, J., 1987. Globins: a case study in molecular phylogeny. Cold Spring Harbor Symp. Quant. Biol. 52, 875–890.

Hardison, R., 1998. Hemoglobins from bacteria to man: evolution of different patterns of gene expression. J. Exp. Biol. 201, 1099–1117.

Harrison, S.C., Blout, E.R., 1965. Reversible conformational changes of myoglobin and apomyoglobin. J. Biol. Chem. 240, 299–303.

Hughson, F.M., Baldwin, R.L., 1989. Use of site-directed mutagenesis to destabilize native apomyoglobin relative to folding intermediates. Biochemistry 28, 4415–4422.

Hughson, F.M., Barrick, D., Baldwin, R.L., 1991. Probing the stability of a partly folded apomyoglobin intermediate by site-directed mutagenesis. Biochemistry 30, 4113–4118.

Johns, G.C., Somero, G.N., 2004. Evolutionary convergence in adaptation of proteins to temperature: A4-lactate dehydrogenases of pacific damselfishes (*Chromis* spp.). Mol. Biol. Evol. 21, 314–320.

Medigue, C., et al., 2005. Coping with cold: the genome of the versatile marine Antarctica bacterium *Pseudoalteromonas haloplanktis* TAC125. Genome Res. 15, 1325–1335.

Methe, B.A., et al., 2005. The psychrophilic lifestyle as revealed by the genome sequence of *Colwellia psychrerythraea* 34H through genomic and proteomic analyses. Proc. Natl. Acad. Sci. U. S. A. 102, 10913–10918.

Milani, M., Savard, P.Y., Ouellet, H., Ascenzi, P., Guertin, M., Bolognesi, M., 2003. A TyrCD1/TrpG8 hydrogen bond network and a TyrB10/Tyr CD1 covalent link shape the heme distal site of *Mycobacterium tubercolosis* hemoglobin O. Proc. Natl. Acad. Sci. U. S. A. 100, 5766–5771.

Milani, M., et al., 2005. Structural bases for heme binding and diatomic ligand recognition in truncated hemoglobins. J. Inorg. Biochem. 99, 97–109.

Miranda, J.J.L., Maillett, D.H., Soman, J., Olson, J.S., 2005. Thermoglobin, oxygen-avid hemoglobin in a bacterial hyperthermophile. J. Biol. Chem. 280, 36754–36761.

Nardini, M., et al., 2006. Structural determinants in the group III truncated hemoglobin from *Campylobacter jejuni*. J. Biol. Chem. 281, 37803–37812.

Ouellet, H., et al., 2003. Reactions of *Mycobacterium* tubercolosis truncated hemoglobin O with ligands reveal a novel ligand-inclusive hydrogen bond network. Biochemistry 42, 5764–5774.

Pesce, A., et al., 2000. A novel two-over-two alpha-helical sandwich fold is characteristic of the truncated hemoglobin family. EMBO J. 19, 2424–2434.

Rabus, R., et al., 2004. The genome of *Desulfotalea psychrophila*, a sulfate-reducing bacterium from permanently cold Arctic sediments. Environ. Microbiol. 6, 887–902.

Somero, G.N., 1995. Proteins and temperature. Annu. Rev. Physiol. 57, 43–68.

Swofford, D.L., 2001. PAUP*. Phylogenetic Analysis Using Parsimony (*and Other Methods). Version 4. Sinauer Associates, Sunderland, Massachusetts, U.S.A.

Thompson, J.D., Gibson, T.J., Plewniak, F., Jeanmougin, F., Higgins, D.G., 1997. The CLUSTAL X windows interface: flexible strategies for multiple sequence alignment aided by quality analysis tools. Nucleic Acids Res. 25, 4876–4882.

Vinogradov, S.N., et al., 2005. Three globin lineages belonging to two structural classes in genomes from the three kingdoms of life. Proc. Natl. Acad. Sci. U. S. A. 102, 11385–11389.

Vinogradov, S.N., et al., 2006. A phylogenomic profile of globins. BMC Evol. Biol. 6, 31–67.

Vuletich, D.A., Lecomte, J.T.J., 2006. A phylogenetic and structural analysis of truncated hemoglobins. J. Mol. Evol. 62, 196–210.

Wainwright, L.M., Wang, Y., Park, S.F., Yeh, S., Poole, R.K., 2006. Purification and spectroscopic characterization of Ctb, a group III truncated haemoglobin implicated in oxygen metabolism in the food-borne pathogen *Campylobacter jejuni*. Biochemistry 45, 6003–6011.

Wittenberg, J.B., Bolognesi, M., Wittenberg, B.A., Guertin, M., 2002. Truncated hemoglobins: a new family of hemoglobins widely distributed in bacteria, unicellular eukaryotes and plants. J. Biol. Chem. 277, 871–874.