Cytoglobin: A Novel Globin Type Ubiquitously Expressed in Vertebrate Tissues

Thorsten Burmester, Bettina Ebner,† Bettina Weich,† and Thomas Hankeln†*

*Institute of Zoology; and †Institute of Molecular Genetics, Biosafety Research and Consulting, Johannes Gutenberg University

Vertebrates possess multiple respiratory globins that differ in terms of structure, function, and tissue distribution. Three types of globins have been described so far: hemoglobin facilitates the transport of oxygen in the blood, myoglobin serves oxygen transport and storage in the muscle, and neuroglobin has a yet unidentified function in nerve cells. Here we report the identification of a fourth and novel type of globin in mouse, man, and zebrafish. It is expressed in apparently all types of human tissue and therefore has been called cytoglobin (CYGB). Mouse and human CYGBs comprise 190 amino acids; the zebrafish CYGB, 174 amino acids. The human *CYGB* gene is located on chromosome 17q25. The mammalian genes display a unique exon-intron pattern with an additional exon resulting in a C-terminal extension of the protein, which is absent in the fish CYGB. Phylogenetic analyses suggest that the CYGBs had a common ancestor with vertebrate myoglobins. This indicates that the vertebrate myoglobins are in fact a specialized intracellular globin that evolved in adaptation to the special needs of muscle cells.

Introduction

Globins are respiratory proteins that usually bind an oxygen molecule between the iron ion of the porphyrin ring and a histidine of the polypeptide chain (Dickerson and Geis 1983). Globins have been found in bacteria, plants, fungi, and animals and play an important role in the respiratory system (Hardison 1996). In the jawed vertebrates (Gnathostomata), three types of globins have been described. The heterotetrameric hemoglobin in the red blood cells serves to transport oxygen in the circulatory system (Dickerson and Geis 1983). The monomeric myoglobin in the muscle acts as an oxygen buffer and may facilitate oxygen diffusion (Wittenberg and Wittenberg 1989; Wittenberg 1992), and it was recently shown to be involved in the detoxification of NO (Flögel et al. 2001). Because of their high concentrations in the blood and muscles, respectively, the vertebrate hemoglobins and myoglobins are among the best-studied proteins in terms of structure, function, and evolution (Dickerson and Geis 1983; Hardison 1996, 2001). Neuroglobin, expressed in mouse and human brain, has just recently been added to the vertebrate globin family (Burmester et al. 2000). Neuroglobins are phylogenetically ancient, and it has been suggested that they enhance the oxygen supply of nerve tissues. These data suggested that the oxygen metabolism of vertebrates is more complex than previously thought and that other globins of yet unknown function remained to be discovered. Here we report the discovery and molecular evolutionary analyses of a novel vertebrate globin type, named cytoglobin (CYGB).

Abbreviations: CYGB, cytoglobin; RT-PCR, reverse transcription–polymerase chain reaction; SSC, standard saline citrate.

Key words: globin, myoglobin, cytoglobin, evolution, gene duplication.

Address for correspondence and reprints: Dr. Thomas Hankeln, Johannes Gutenberg University Mainz, Institut für Molekulargenetik, Becherweg 32, D-55099 Mainz, Germany. E-mail: hankeln@molgen.biologie.uni-mainz.de.

Correspondence and requests for material should be addressed to T.B. (burmeste@mail.uni-mainz.de) or T.H. (hankeln@molgen.biologie.uni-mainz.de).

Mol. Biol. Evol. 19(4):416–421. 2002

Materials and Methods

Cloning and Sequencing of Mouse, Human, and Zebrafish CYGBs

Human and mouse EST databases were searched using the TBLASTN algorithm (Altschul et al. 1990) with the amino acid sequence of the Drosophila globin (Burmester and Hankeln 1999) as query. Several partial EST matches for Human (accession numbers R87866, AI093531, AL514650, AL516826, AL514972, and BE313504) and mouse (AA469788, AA469798, BE648697, and BF159503) were extracted from the databases and aligned. Specific oligonucleotide primers were designed according to these aligned EST sequences (Human—HsCYGB-N 5′-ATGGAGAAAGTGCCAGG CGAGATG-3′ and HsCYGB-B 5′-TTACGGCCCCGA AGAGGGCAGT-3′; mouse—MmCYGB-N 5′-ATGG AGAAAGTGCCGGGCGACATG-3′ and MmCYGB-B 5′-TTACGGCCCTGAAGAGGGCAGA-3′) and used to amplify *CYGB* cDNAs making up the complete coding regions from mouse and human brain total RNA by reverse transcription–polymerase chain reaction (RT-PCR; using SuperScript™ reverse transcriptase, Life Technologies). The PCR products were cloned into the pGEM-Teasy vector (Promega) and sequenced on both strands using DyeTerminator™ chemistry (Applied Biosystems) on ABI377 sequencers by GENterprise GmbH, Mainz, Germany. Sequences were deposited under the accession numbers AJ315162 (human *CYGB* cDNA) and AJ315163 (mouse CYGB cDNA). A zebrafish (*Danio rerio*) CYGB cDNA clone was identified by EST database searching and obtained from the Resource Center, Berlin, Germany (http://www.rzpd.de), and sequenced by a primer walking strategy (accession number AJ320232).

Expression Analysis of Human *CYGB*

An RNA Master Blot™ (Clontech) containing normalized amounts of polyA$^+$ RNA from 50 human tis-

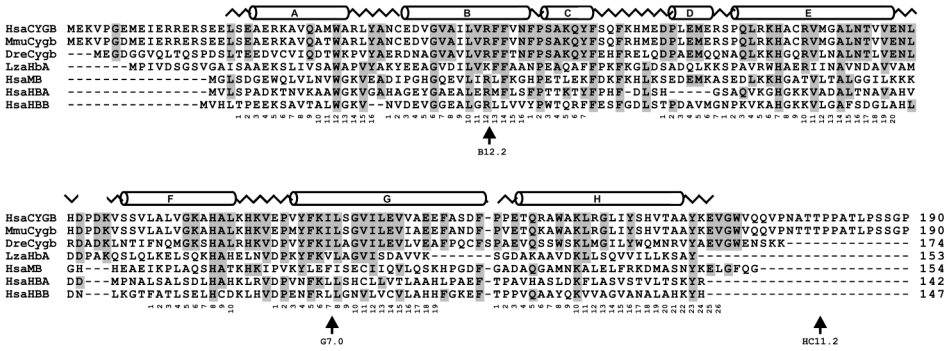

FIG. 1.—Comparison of human, mouse, and zebrafish CYGBs (HsaCYGB, accession number AJ315162; MmuCygb, AJ315163; DreCygb, AJ320232) with human myoglobin (HsaMB, M14603) and hemoglobins α and β (HsaHBA, J00153; HsaHBB, M36640) and a lamprey globin (LzaHbA, Z24746). The globin consensus numbering is given below the sequences; the secondary structure of the sperm whale (*Physeter catodon*) myoglobin is superimposed in the upper row. Alpha-helices are designated A through H, and amino acids strictly conserved between the CYGBs and the myo- or hemoglobins are shaded. The intron positions in the human *CYGB* gene (at B12.2, E11-0, and HC11-2) are indicated by arrows.

sues was hybridized to a ^{32}P-labelled (Feinberg and Vogelstein 1983) subcloned *CYGB* cDNA probe containing the complete coding sequence (accession number AJ315162). Washing was performed at 68°C in 0.1 × standard saline citrate (SSC) solution. Signals were quantified on a Fuji BAS-1800 phosphorimager. No nonspecific binding of the probe was observed to yeast total RNA and tRNA, *E. coli* rRNA, poly rA, and human C_ot 1 repetitive DNA.

Sequence and Phylogenetic Analyses

Sequence analyses were carried out with the programs provided by the Software Package 9.0 from the Genetics Computer Group (GCG), Wisconsin, and the ExPASy web server (http://www.expasy.ch). The genomic organization of the human *CYGB* gene was depicted using the PIPMAKER program (http://nog.cse.psu.edu/pipmaker/; Schwartz et al. 2000). Amino acid sequences of selected vertebrate globins were aligned with ClustalX (Thompson et al. 1997) and corrected using published alignments (Burmester and Hankeln 1999; Burmester et al. 2000) and globin structural data. The software packages PHYLIP 3.6 (Felsenstein 2000) and TREE-PUZZLE 5.0 (Strimmer and von Haeseler 1996) were applied for phylogenetic inference. Gamma-corrected distances were calculated using the PAM250 model with eight rate categories (Dayhoff, Schwartz, and Orcutt 1978). Tree constructions were performed using the neighbor-joining method. The reliability of the trees was tested by bootstrap analysis (Felsenstein 1985) with 100 replications using PUZZLEBOOT (shell script by M. Holder and A. Roger). Synonymous and nonsynonymous nucleotide substitution rates were calculated using the method of Nei and Gojobori (1986).

Results

Identification, Cloning, and Expression Analysis of CYGB

A systematic search of the expressed sequence tag (EST; Boguski, Lowe, and Tolstoshev 1993) databases using the Drosophila globin (Burmester and Hankeln 1999) as query resulted in several partial globin-like cDNA sequences from Human, mouse, and zebrafish that did not correspond to the vertebrate hemoglobins, myoglobins or neuroglobins. Specific primers (see *Materials and Methods*) were used to amplify the mouse and human globin cDNAs via RT-PCR from brain mRNA. They each code for a novel globin of 190 amino acids (20.9 kDa; fig. 1). A homologous cDNA clone from the zebrafish *D. rerio* was also obtained by EST database searching, and was sequenced completely. It covers 5,188 bp with an open reading frame of 522 bp, giving rise to a 174–amino acid protein (fig. 1). The cDNAs from all three vertebrates encode proteins that fit well into a globin alignment based on the conserved alpha-helices A to H of the globin fold. The protein sequences of the novel globin type are clearly longer than those of the typical vertebrate myo- and hemoglobin chains, which comprise about 150 amino acids. The difference is because of unusually long N- and C-termini, whereas no sequence insertions interrupt the globin fold (fig. 1). Computer predictions using the PSORT II program (Nakai and Horton 1999) indicate that the novel globins do not contain any signal peptide and are most likely localized in the cytoplasm.

The expression of the newly identified globin gene was analyzed by Northern hybridization to a filter containing mRNA from different human tissues and developmental stages (fig. 2). We observed a ubiquitous expression of the globin mRNA, with the strongest signals seen in heart, stomach, bladder, and small intestine. Because of its apparently widespread expression, the novel globin was designated as *cytoglobin.*

The Human and Mouse *CYGB* Genes

The genomic sequence of the human *CYGB* gene was derived from contigs of the human genome project (accession number AC015802; International Human Genome Sequencing Consortium 2001). By using our cDNA sequence information, it was possible to reveal the complete gene structure which comprises four exons (fig. 3), whereas in the gene prediction by the EN-

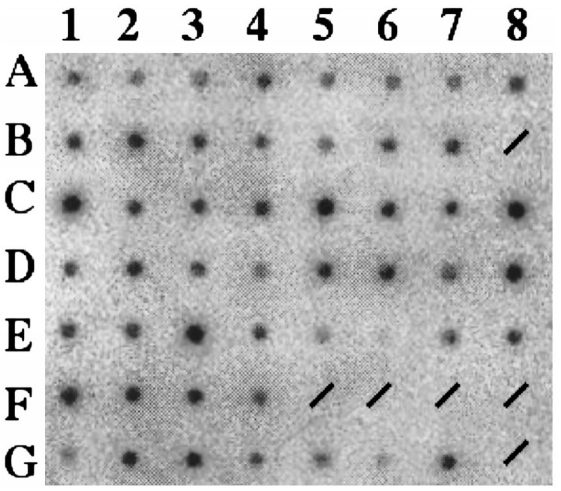

FIG. 2.—Northern dot blot analysis of human mRNA with *CYGB* cDNA as hybridization probe. All human tissues and developmental stages show the presence of *CYGB* mRNA. A1, whole brain; A2, amygdala; A3, caudate nucleus; A4, cerebellum; A5, cerebral cortex; A6, frontal lobe; A7, hippocampus; A8, medulla oblongata; B1, occipital pole; B2, putamen; B3, substantia nigra; B4, temporal lobe; B5, thalamus; B6, subthalamic nucleus; B7, spinal cord; C1, heart; C2, aorta; C3, skeletal muscle; C4, colon; C5, bladder; C6, uterus; C7, prostate; C8, stomach; D1, testes; D2, ovary; D3, pancreas; D4, pituitary gland; D5, adrenal gland; D6, thyroid gland; D7, salivary gland; D8, mammary gland; E1, kidney; E2, liver; E3, small intestine; E4, spleen; E5, thymus; E6, peripheral leukocyte; E7, lymph node; E8, bone marrow; F1, appendix; F2, lung; F3, trachea; F4, placenta; G1, fetal brain; G2, fetal heart; G3, fetal kidney; G4, fetal liver; G5, fetal spleen; G6, fetal thymus; G7, fetal lung. Blot positions indicated by dashes contain no RNA sample.

SEMBL annotation project the small, last coding exon is missing (http://www.ensembl.de). The human *CYGB* gene is located on chromosome 17q25. It reveals the presence of three introns (figs. 1 and 3), at position B12-2 (i.e., between codon positions 2 and 3 of the 12th amino acid of globin helix B), at position G7-0, and at position HC11-2 downstream of the H helix, close to the C-terminus of the protein sequence. This genomic organization is confirmed by the comparative sequencing of the murine *Cygb* gene (to be published elsewhere).

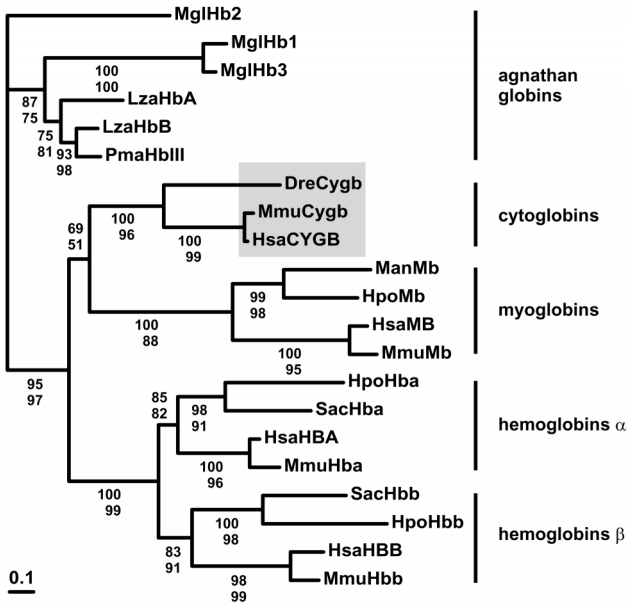

FIG. 4.—Phylogenetic analysis of the vertebrate globins using amino acid sequence data. Bootstrap (upper number) and quartet support (lower number) values are given at the branches; the bar equals 0.1 PAM distance. The CYGBs are shaded. In addition to the proteins used in figure 1, the tree includes *Myxine glutinosa* hemoglobins 1 and 3 (MglHb1, AF156936; MglHb2, AF157494; MglHb3, AF184239), *Lampetra zanandreai* hemoglobin B (LzaHbB, Z24748), *Petromyzon marinus* hemoglobin III (PmaHbIII, P09968), *Heterodontus portusjacksoni* myoglobin (HpoMb, P02206) and hemoglobins α and β (HpoHba, P02021; HpoHbb, P02143), *Mustelus antarcticus* myoglobin (ManMb, P14399), *Squalus acanthias* hemoglobins α and β (SacHba, A24653; SacHbb, B24653), and *Mus musculus* myoglobin (MmuMb, P04247) and hemoglobins α and β (MmuHba, A45964; MmuHbb, P02088).

Phylogenetic Analyses

The cDNA-derived amino acid sequences of the CYGBs were aligned with selected vertebrate hemoglobins and myoglobins. Molecular phylogenetic analyses using the neighbor-joining, maximum parsimony, or maximum likelihood method result in four well-supported clades, representing α- and β-hemoglobins, myoglobins, agnathan globins, and the CYGBs, respectively (fig. 4). Because of the high degree of divergence in the

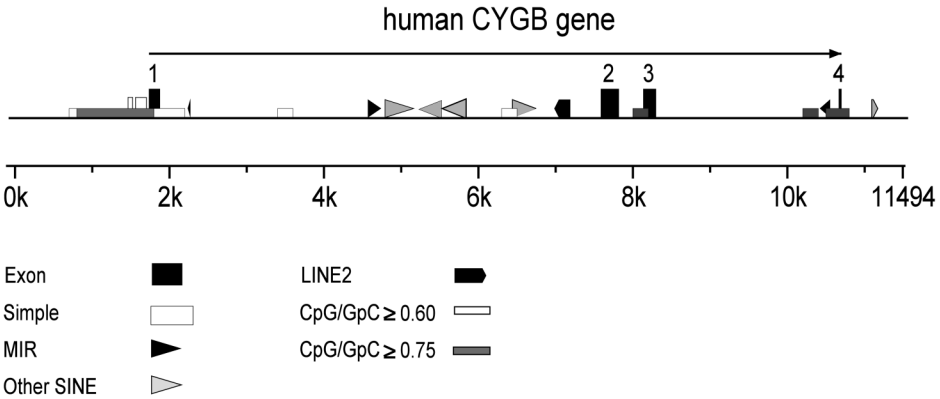

FIG. 3.—Genomic organization of the human *CYGB* gene. Exons 1 to 4 are boxed. Various types of repetitive sequences present and GC-rich regions are also indicated.

globin superfamily, the interrelationship between these clades cannot be resolved with the help of either invertebrate globins or neuroglobins as an outgroup. However, because it is known that distinct myoglobins and hemoglobins most likely differentiated after the split of the Agnatha and the Gnathostomata (jawless and jawed vertebrates; Goodman 1981), the hemoglobins of lamprey and hagfish, agnathan species, may be used as the outgroup (fig. 4). When placing the root at the agnathan globins, there is consistent support for a close phylogenetical relationship of the CYGBs with the vertebrate myoglobins, although the bootstrap support values are only moderate (50% to 70%).

Discussion

We report the identification of a fourth and novel type of globin present in the vertebrates. This molecule, dubbed cytoglobin, is ubiquitously expressed in human tissue and fulfils a yet undefined role. Taking into account the recent findings of a specialized globin in nerve cells (Burmester et al. 2000) and a novel orphaned β-hemoglobin (ω-globin) from the Marsupialia (Wheeler et al. 2001), this in fact indicates that the evolution of globins in higher organisms may be more complex than hitherto thought.

Protein Structure and Genomic Organization of CYGBs

Within the conserved globin fold, which covers the standard alpha helices A through H, the key residues important for the function of CYGB as a typical oxygen-binding protein are strictly conserved. The proximal and distal histidines in the positions E7 and F8 as well as the phenylalanine at the CD1 corner are present in the CYGBs (fig. 1). The lengths of the mammalian (190 amino acids) and fish (174 amino acids) CYGBs exceed those of vertebrate myo- and hemoglobins. The length differences are exclusively because of N- and C-terminal protein extensions, which occasionally have been observed previously in invertebrate globins (e.g., in *Caenorhabditis elegans*; Neuwald et al. 1997) but whose functional relevance is unclear. Whereas the C-terminal extension of CYGBs may be caused by the recruitment of an additional exon (see subsequent discussion), the N-terminal extension of murine and human CYGBs seems to have partly resulted from a direct duplication of 21 nucleotides (with 7/21 mismatches) at the 5′ end of the coding region (data not shown).

The antiquity of introns within globin genes and their positional stability during evolution have been a matter of intense debate (e.g., Hankeln et al. 1997; Logsdon, Stoltzfus, and Doolittle 1998). The human *CYGB* gene displays the B12-2 and G7-0 introns which are typically found in many globins, including the vertebrate hemo-, myo- and neuroglobins, and which must therefore be considered phylogenetically ancient (Dixon and Pohajdak 1992; Burmester et al. 2000). However, the additional intron in the 3′-most region of the murine and human *CYGB* coding sequences (corresponding to the C-terminal position HC11-2) is unprecedented. The origin of the small exon 4 sequence, which only encodes the 10 most–C-terminal amino acid residues, is unclear. The C-terminus of zebrafish CYGB is shorter and lacks the additional exon 4 sequence (fig. 1). In the murine and human *CYGB*s, the HC11-2 intron occurs just downstream of the C-terminal end of other globin sequences, and we therefore consider that exon 4 might have been acquired only during the evolution of the tetrapod CYGBs.

Molecular Evolution of CYGBs

Mouse and human CYGB share 92.8% of the nucleotides and 95.3% of the amino acids in the coding region. The zebrafish CYGB shows 49% amino acid identity to the mammalian proteins. CYGBs display the highest degree of amino acid sequence similarity to the hemoglobins of the Agnatha (26% to 33% identity). Somewhat lower scores were observed when the CYGBs were compared with the myoglobins and hemoglobins (~30% identity). Assuming that mice and humans diverged about 80 MYA (Kimura 1987), an amino acid substitution rate of about 0.3×10^{-9} replacements per site per year was inferred for the mammalian CYGBs. This is much lower than calculated for the orthologous mammalian hemoglobins (0.9×10^{-9} to 1.2×10^{-9}) and myoglobins (0.8×10^{-9} to 1.2×10^{-9}) but lies in the range of the neuroglobins (0.4×10^{-9}). These values are in agreement with our calculations of very low nonsynonymous nucleotide substitution rates in human and mouse cyto- and neuroglobin (dn = 0.02 and 0.03 nonsynonymous substitutions per site, respectively), compared with human and mouse hemoglobins (dn = 0.09 for α globin, 0.21 for β globin) and myoglobin (dn = 0.09). In the case of *CYGB*, the low nonsynonymous substitution rate is correlated to an unusually low substitution rate at synonymous codon positions (ds = 0.28 synonymous substitutions per site), the reason for which is unclear. The ds:dn ratio (Nei and Gojobori 1986) of $\gg 1$, however, clearly demonstrates that mammalian CYGBs evolve under strong purifying selection.

A Model of Globin Evolution in Vertebrates

Phylogenetic analyses suggest that the CYGBs share a common clade with the vertebrate myoglobins (fig. 4). An independent confirmation of the common ancestry of cyto- and myoglobins may come from data suggesting that the chromosomal regions encompassing *CYGB* (17q25) and myoglobin (22q12) represent long, paralogous stretches of genomic DNA, which are thought to have originated by an ancient duplication event (A. McLysaght, K. Hokamp, and K. H. Wolfe, personal communication).

Taking into account the antiquity of the neuroglobins, the last common ancestor of all vertebrates most likely possessed two different types of globins (fig. 5). Neuroglobin maintained its function in the nervous system, which it had acquired early in the evolution of the Bilateria (Burmester et al. 2000). The other globin likely differentiated into a cellular globin, which later gave rise to the myoglobins and CYGBs on the one hand and to

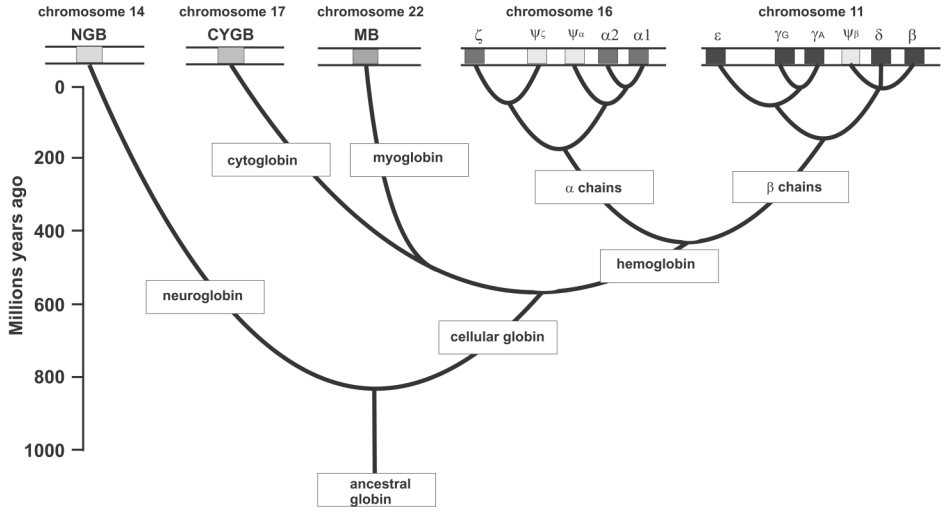

FIG. 5.—Model of vertebrate globin evolution (see *Discussion* for detailed explanation).

hemoglobin on the other. The hemoglobins obtained their function in the circulatory system of the gnathostomian vertebrates after their divergence from the lineage leading to the myoglobins and CYGBs, probably as early as 500 to 600 MYA (Goodman et al. 1987). This event was probably correlated with an increase in body size and the evolution of an efficient circulatory system. Myoglobin and CYGB separated later, but before the divergence of the Chondrichthyes and the other gnathostomes more than 450 MYA (Benton 1990, p. 44). It is conceivable that myoglobin, which is present in high concentrations in skeletal and smooth muscle (Wittenberg and Wittenberg 1989; Qiu, Sutton, and Riggs 1998), and which supplies the cells with high amounts of oxygen, is in fact an offspring of a more general tissue-globin of similar or other function.

Functional Implications of CYGB

The physiological function of CYGBs still has to be investigated. During the preparation of this manuscript, Kawada et al. (2001) reported the finding of a protein that we believe represents the rat ortholog of CYGB. The protein was identified in a proteomics approach by virtue of its heavily upregulated expression in stellate cells of rat liver (and it was therefore dubbed stellate cell activation–associated protein, STAP). The authors could demonstrate that STAP possesses peroxidase activity, and they speculate upon its role as a scavenger of peroxides in fibrotic liver. In fact, it has been reported before that the dehaloperoxidase enzyme of the marine worm Amphritite is phylogenetically related to globins (Lebioda et al. 1999; LaCount et al. 2000) and has retained its ability to bind oxygen (Roach et al. 1997). However, without additional data the possible spectrum of physiological role(s) of CYGB within the broad variety of vertebrate tissues can only be hypothesized. CYGB may be involved in intracellular oxygen storage or transfer, have an enzymatic (peroxidase) function, play a role in O_2 sensing (Goldberg, Dunning, and Bunn 1988), or may bind and detoxify NO (Flögel et al. 2001). Finally, a combination of several of these functions also seems conceivable.

Acknowledgments

We wish to thank E. R. Schmidt, J. Markl, and H. Decker for excellent working facilities, continuous support, and valuable suggestions. This work is supported by the Deutsche Forschungsgemeinschaft (DFG Ha 2103-3 and Bu 956-3) and the Naturwissenschaftlich-Medizinisches Forschungszentrum (NMFZ) Mainz.

LITERATURE CITED

ALTSCHUL, S. F., W. GISH, W. MILLER, E. W. MYERS, and D. J. LIPMAN. 1990. Basic local alignment search tool. J. Mol. Biol. **215**:403–410.

BENTON, M. J. 1990. Vertebrate paleontology. Unwin Hyman Ltd., London.

BOGUSKI, M. S., T. M. LOWE, and C. M. TOLSTOSHEV. 1993. dbEST—database for ''expressed sequence tags.'' Nat. Genet. **4**:332–333.

BURMESTER, T., and T. HANKELN. 1999. A globin gene of *Drosophila melanogaster*. Mol. Biol. Evol. **16**:1809–1811.

BURMESTER, T., B. WEICH, S. REINHARDT, and T. HANKELN. 2000. A vertebrate globin expressed in the brain. Nature **407**:520–523.

DAYHOFF, M. O., R. M. SCHWARTZ, and B. C. ORCUTT. 1978. A model of evolutionary change in proteins. Pp. 345–352 *in* M. O. DAYHOFF, ed. Atlas of protein sequence structure, Vol. 5, Suppl. 3. National Biomedical Research Foundation, Washington, DC.

DICKERSON, R. E., and I. GEIS. 1983. Hemoglobin: structure, function, evolution, and pathology. Benjamin/Cummings Publ. Co., Menlo Park, Calif.

DIXON, B., and B. POHAJDAK. 1992. Did the ancestral globin gene of plants and animals contain only two introns? TIBS **17**:486–488.

FEINBERG, A., and B. VOGELSTEIN. 1983. A technique for radiolabeling DNA restriction endonuclease fragments to high specific activity. Anal. Biochem. **132**:6–13.

FELSENSTEIN, J. 1985. Confidence limits on phylogenies: an approach using the bootstrap. Evolution **39**:783–791.

———. 2000. PHYLIP (phylogeny inference package). Version 3.6. Distributed by the author, Department of Genetics, University of Washington, Seattle.

FLÖGEL, U., M. W. MERX, A. GOEDECKE, U. DECKING, and J. SCHRADER. 2001. Myoglobin: a scavenger of bioactive NO. Proc. Natl. Acad. Sci. USA **98**:735–740.

GOLDBERG, M. A., S. P. DUNNING, and H. F. BUNN. 1988. Regulation of the erythropoietin gene: evidence that the oxygen sensor is a heme protein. Science **242**:1412–1415.

GOODMAN, M. 1981. Globin evolution was apparently very rapid in early vertebrates: a reasonable case against the rate-constancy hypothesis. J. Mol. Evol. **17**:114–120.

GOODMAN, M., J. CZELUSNIAK, B. F. KOOP, D. A. TAGLE, and J. L. SLIGHTOM. 1987. Globins: a case study in molecular phylogeny. Cold Spring Harbor Symp. Quant. Biol. **52**:875–890.

HANKELN, T., H. FRIEDL, I. EBERSBERGER, J. MARTIN, and E. R. SCHMIDT. 1997. A variable intron distribution in globin genes of *Chironomus*: evidence for recent intron gain. Gene **205**:151–160.

HARDISON, R. C. 1996. A brief history of hemoglobins: plant, animal, protist, and bacteria. Proc. Natl. Acad. Sci. USA **93**:5675–5679.

———. 2001. New views of evolution and regulation of vertebrate beta-like globin gene clusters from an orphaned gene in marsupials. Proc. Natl. Acad. Sci. USA **98**:1327–1329.

KAWADA, N., D. B. KRISTENSEN, K. ASAHINA, K. NAKATANI, Y. MINAMIYAMA, S. SEKI, and K. YOSHIZATO. 2001. Characterization of a stellate cell activation-associated protein (STAP) with peroxidase activity found in rat hepatic stellate cells. J. Biol. Chem. **276**:25318–25323.

KIMURA, M. 1987. Molecular evolutionary clock and the neutral theory. J. Mol. Evol. **26**:24–33.

LACOUNT, M. W., E. ZHANG, Y. P. CHEN, K. HAN, M. W. WHITTON, D. E. LINCOLN, S. A. WOODIN, and L. LEBIODA. 2000. The crystal structure and amino acid sequence of dehaloperoxidase from *Amphritite ornata* indicate common ancestry with globins. J. Biol. Chem. **275**:18712–18716.

LEBIODA, L., M. W. LACOUNT, E. ZHANG, Y. P. CHEN, K. HAN, M. W. WHITTON, D. E. LINCOLN, and S. A. WOODIN. 1999. An enzymatic globin from a marine worm. Nature **401**:445.

LOGSDON, J. M., A. STOLTZFUS, and W. F. DOOLITTLE. 1998. Molecular evolution: recent cases of spliceosomal intron gain? Curr. Biol. **8**:R560–R563.

NAKAI, K., and P. HORTON. 1999. PSORT: a program for detecting sorting signals in proteins and predicting their subcellular localization. Trends Biochem. Sci. **24**:34–36.

NEI, M., and T. GOJOBORI. 1986. Simple methods for estimating the number of synonymous and nonsynonymous nucleotide substitutions. Mol. Biol. Evol. **3**:418–426.

NEUWALD, A. F., J. S. LIU, D. J. LIPMAN, and C. E. LAWRENCE. 1997. Extracting protein alignment models from the sequence database. Nucleic Acids Res. **25**:1665–1677.

QIU, Y., L. SUTTON, and A. F. RIGGS. 1998. Identification of myoglobin in human smooth muscle. J. Biol. Chem. **273**:23426–23432.

ROACH, M. P., Y. P. CHEN, S. A. WOODIN, D. E. LINCOLN, C. R. LOVELL, and J. H. DAWSON. 1997. *Notomastus lobatus* chloroperoxidase and *Amphitrite ornata* dehaloperoxidase both contain histidine as their proximal heme iron ligand. Biochemistry **36**:2197–2202.

SCHWARTZ, S., Z. ZHANG, K. A. FRASER, A. SMIT, C. RIEMER, J. BOUCK, R. A. GIBBS, R. HARDISON, and W. MILLER. 2000. PIPMAKER—a web server for aligning two genomic DNAs. Genome Res. **10**:577–586.

STRIMMER, K., and A. VON HAESELER. 1996. Quartet puzzling: a quartet maximum likelihood method for reconstructing tree topologies. Mol. Biol. Evol. **13**:964–969.

THOMPSON, J. D., T. J. GIBSON, F. PLEWNIAK, F. JEANMOUGIN, and D. G. HIGGINS. 1997. The ClustalX Windows interface: flexible strategies for multiple sequence alignment aided by quality analysis tools. Nucleic Acids Res. **25**:4876–4882.

WHEELER, D., R. HOPE, S. J. B. COOPER, G. DOLMAN, G. C. WEBB, C. D. K. BOTTEMA, A. A. GOOLEY, M. GOODMAN, and R. A. B. HOLLAND. 2001. An orphaned mammalian globin gene of ancient evolutionary origin. Proc. Natl. Acad. Sci. USA **98**:1101–1106.

WITTENBERG, J. B. 1992. Functions of cytoplasmatic hemoglobins and myohemerythrin. Adv. Comp. Environ. Physiol. **13**:60–85.

WITTENBERG, J. B., and B. A. WITTENBERG. 1989. Mechanisms of cytoplasmic hemoglobin and myoglobin function. Annu. Rev. Biophys. Biophys. Chem. **19**:217–241.

KENNETH WOLFE, reviewing editor

Accepted October 12, 2001