

# Structural Characterization of Piscine Globin Superfamily Proteins

Yoshihiro Ochiai

**Abstract**—Globin superfamily proteins including myoglobin and hemoglobin, have welcome new members recently, namely, cytoglobin, neuroglobin and globin X, though their physiological functions are still to be addressed. Fish are the excellent models for the study of these globins, but their characteristics have not yet been discussed to date. In the present study, attempts have been made to characterize their structural uniqueness by making use of proteomics approach. This is the first comparative study on the characterization of globin superfamily proteins from fish.

**Keywords**—Globin, Superfamily, Protein, Fish, Structure

## I. INTRODUCTION

RECENT advances in genome analyses on a wide range of organisms have succeeded in finding the existence of so many new proteins, for some of which even physiological functions have not been annotated. Among them, new globin members have been found in vertebrates [1].

Together with the conventional members, myoglobin (Mgb), and hemoglobin (Hgb), the superfamily consists of neuroglobin (Neugb) [2], cytoglobin (Cygb) [3], globin X [1], globin E [4], and globin Y [5]. The expression levels of these proteins are quite tissue-specific [1,6,7]. The globin proteins appeared about 4000 million years ago [8]. They can bind various gaseous ligands, namely, O<sub>2</sub>, NO, and CO [9]. The structures of classical globins including Mgb are featured by 3/3  $\alpha$ -helical sandwich fold, whereas smaller globins with the truncated N terminus are characterized by 2/2  $\alpha$ -helical structures [8]. Globins seem to occur in the three kingdoms of life [8].

The physiological significance of globins has been believed for a long time in oxygen transport, short- and long-term oxygen storage [10], partial oxygen buffering [11,12], facilitated oxygen diffusion [13], but some new functions have recently been disclosed, namely, scavenging of free radicals like nitric oxide [14] and protection against reactive oxygen species [8]. The functions of the globins seem to be closely related with their tissue-specific localization. For example, Cygb is present in neurons and fibroblast-like cells, and Neugb is found in the nervous system in contrast to Mgb found mostly in muscles [15]. Mgb protein has also been found not only striated muscle but also in the brain of fish [7].

In the presented study, structural characteristics of zebrafish globin proteins were characterized in detail by taking the advantage of bioinformatic approaches.

## II. METHODS

### A. Sequence alignment and data analysis

The obtained amino acid sequences from the NCBI database were aligned using CLUSTAL W in BioEdit program (v 7.0.9) [16]. The sequence data used in this study are as follows: Cygb1 and 2 from zebrafish *Danio rerio* (BC094999 and BC135090) and human *Homo sapiens* Cygb (AJ315162), Neugbs from zebrafish (AJ315610) and mouse (AJ245945), zebrafish globin X (AJ635194), zebrafish Mgb (AY337025), zebrafish Hgba, and Hgb $\beta$  (BC042321 and NM\_131020, respectively).

### B. Bioinformatic analyses

Phylogenetic tree was constructed based on the amino acid sequences by the neighbor-joining method using the program MEGA 4 [17]. Isoelectric point (pI) and molecular weight were calculated based on the deduced amino acid sequences by a web tool using pK values from EMBOSS (<http://isoelectric.ovh.org/>). Hydropathy profiles of the globins were compared according to the hydrophobicity scales by Kyte and Doolittle [18].

Three dimensional homology models of the globin proteins were constructed based on the amino acid sequences and the X-ray structures on Protein Data Bank (PDB) database utilizing Swiss-model program [19]. The structural data used were human Cygb (PDB ID 1UMO), mouse Neugb (2VRY), blackfin tuna *Thunnus atlanticus* Mb (2NRL), and perch *Perca fluviatilis* Hgb (3BJ2). The structures were visualized with Swiss-PdbViewer (<http://www.expasy.org/spdbv/>) and PyMOL (V.1.1).

## III. RESULTS AND DISCUSSION

Sequence data of globins superfamily proteins are so far available for zebrafish *Danio rerio*, puffer *Tetraodon nigroviridis*, goldfish *Carassius auratus* and medaka *Oryzias latipes* as far as fish are concerned. In this study, characterization of those globins was focused on those of zebrafish, since the analysis of globins from the other species gave basically the same features.

In Fig. 1 is shown the alignment of the amino acid sequences of zebrafish globins. It is noteworthy that the full lengths of the globins differ, namely, Hb is the smallest and globin X is the largest protein among them. This protein has the extended sequences at both the N and C termini. The number of the residues in common was found to be very few among the proteins. It has been reported that distal and proximal histidine residues as well as the phenylalanine residue are conserved in globinX [1].

Y. Ochiai is with School of Marine Science & Technology, Tokai University, Shizuoka 424-8610, Japan (e-mail: aochiai@tokai-u.jp). The present study was supported in part by a Grant-in-Aid from Japan Society for Promotion of Science (KAKENHI #22380015).

GlobinX	MGCAISGSGLTARAPEIRAGEEETPAGLTANHIRLIKESWRLIQEDIAKVGIIIMFVRLFE
Neugb	-----MEKLSEKDKGLIRDSWESLGKNKVPHGIVLFTLRFEE
Cygb	-----MEGDGGVQLTQSPDSLTEEDVCVIQDTWKPVYAERDNAGVAVLVRFET
Hgb	-----VEWTDARTAILGLWGKLN--IDEIGPQALSRLCI
Mgb	-----MADHDLVLKCGAVEADYAANGGEVLNRLFK
	. : * : * : *
GlobinX	THPECKDVFFLFRDVEDLERLRTSRELRAHGLRVMSFIEKSVARLDQLERLETALALELGK
Neugb	LDPALLTLFSYSTNCGDAPECLSSPEFLEHVTKVMLVIDAAVSHLDDLHTLEDFLNLGR
Cygb	NFPSAKQYFEHFRELQDPAEMQQNAQLKEHGQRVLNLTNLVENLRDADKLNTIFNQMGK
Hgb	VYPWTQRYFATFGNLSSPAAIMGNPKVAAHGRTVMGGLERAIKNMDN---IKNTYAALSV
Mgb	EYPTLTKLFPKFSGIS--QGDLAGSPAVAAHGATVLKKLGELLKAKGD---HAALLKPLAN
	* * . . * *: : : : : .
GlobinX	SHYRYN-APPKYYGYVGAEFICAVRPILKDRWTP-ELEEAWKTLFQYVTSIMREGFLEEE
Neugb	KHQAVG-VNTQSFALVGESLLYMLQSSLGPAYTT-SLRQAWLTMYSIIVSAMTRGWAKNG
Cygb	SHALRHKVDVPVYFKILAGVILEVLVEAFPQCFSPEVQSSWSKLMGILYWMNRVYAEVG
Hgb	MHSEKLHVDPDNFKLLADCITVCAAMKFGQAGFNADIQEAQKFLAVVVSALCRQYH---
Mgb	THANIHKVALNNFRLITEVLVKVMAEKAGLDAAG---QGALRRVMDAVIDIGGYKEIG
	* . : : : . : . : : :
GlobinX	RNKRSTQTSSRERPDKRSTAI
Neugb	EHKSN-----
Cygb	WENSKK-----
Hgb	-----
Mgb	FAG-----

Fig. 1 Alignment of the amino acid sequences of the globin proteins from zebrafish *Danio rerio*. Identical residues, conservative replacements and semi-conservative ones are indicated by asterisks, semicolons, and dots, respectively. Neugb, Cygb, Hgb, and Mgb stand for neuroglobin, cytoglobin, hemoglobin, and myoglobin, respectively.

TABLE I  
AMINO ACID SEQUENCE IDENTITY OF ZEBRAFISH  
GLOBINS

	Cygb1	Cygb2	Neugb	GlobinX	Mgb	Hgba	Hgbβ
Cygb1	—	—	—	—	—	—	—
Cygb2	<b>47.7</b>	—	—	—	—	—	—
Neugb	18.2	20.1	—	—	—	—	—
GlobinX	20.1	22.9	27.7	—	—	—	—
Mgb	18.4	17.0	15.0	14.3	—	—	—
Hgba	23.8	26.6	15.4	14.0	20.3	—	—
Hgbβ	22.4	19.0	14.3	15.6	21.1	<b>32.9</b>	—

Refer to the legend of Fig. 1 for the abbreviations.

As shown in TABLE I, the sequence identities of these proteins were in the range of 14.0 and 47.7%. The highest value was obtained between those of Cygb isoforms 1 and 2, whereas the lowest value was found between Hgb and globin X. As far as Mgb is concerned, the identity of which to other proteins being between 14.3 and 18.4, is quite remotely related to the other globins.

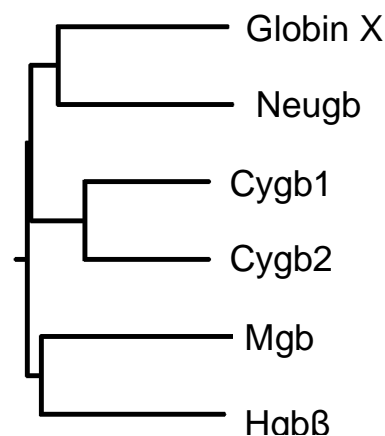


Fig. 2 Phylogenetic tree constructed based on the amino acid sequences of zebrafish *D. rerio* globins. Refer to the legend of Fig. 1 for the abbreviations

Phylogenetic analysis based on the deduced amino acid sequences showed the clades between globin X and Neugb, between Cygb isoforms, and between Mgb and Hgb (Fig. 2). Cygb was diverged from Mgb at the early evolution stage of vertebrates [15].

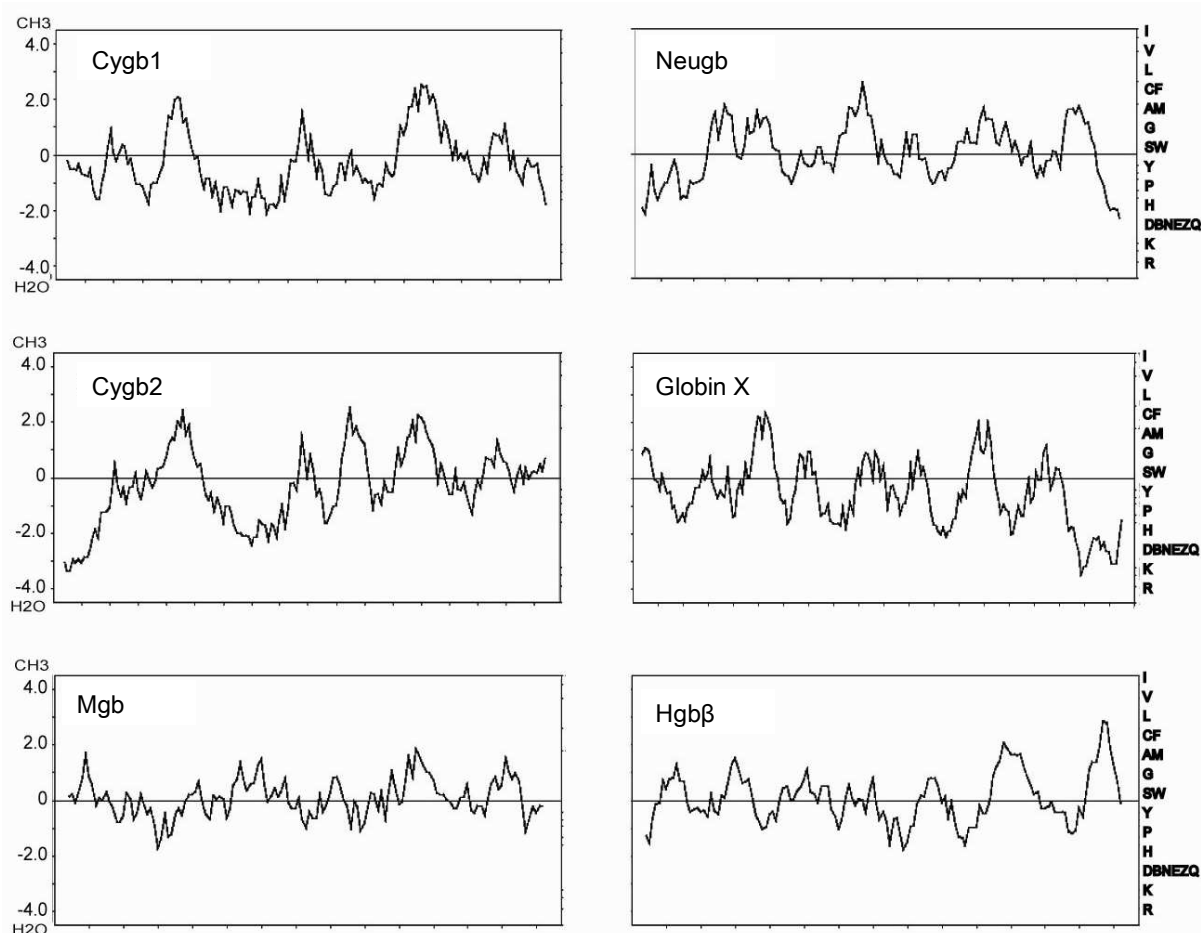


Fig. 3 Hydropathy plot of zebrafish globins. Refer to the legend of Fig. 1 for the abbreviations

TABLE II  
PHYSICOCHEMICAL PARAMETERS CALCULATED FROM THE AMINO ACID SEQUENCES

	Cygb1	Cygb2	Neugb	GlobinX	Mgb	Hgba	Hgbβ
Amino acids	174	179	159	200	147	143	147
Molec. weight	19909	20584	17684	23210	15525	15522	16259
pI	5.22	5.04	5.73	7.73	7.86	7.97	7.84

The fact that Neugb and Cygb are the hexaco-ordinate globins in contrast to Mgb and Hgb belonging to the class of the pentacoordinate ones could be related with the above relations. Globin X and Neugb unexpectedly resembled each other, suggesting that they share the same physiological roles. Such relationship between these two proteins has already been highlighted [1]. It has also been suggested that no out-group is available for globins due to the low sequence conservation (as shown in Fig. 1) and to the generally short sequences [1].

In the next place, hydropathy profiles of the globin proteins were examined (Fig. 3). Cygb isoforms shared similar hydropathy profiles, and were found to contain comparatively high hydrophilic regions unlike the other proteins. However, the N terminus of Cygb2 was more hydrophilic than that of Cygb1.

The physicochemical values were obtained based on the amino acid sequences and are shown in TABLE II. The molecular mass of the globins were as expected from the sequence alignment in Fig. 1. On the other hand, the pI values were found to be as low as around 5 for Cygb isoforms. Neugb was found to be slightly acidic, whereas other globins showed high pI values. As far as pI values are concerned, globin X is quite unlike Neugb, though they were found to be closely related in the phylogenetic analysis (Fig. 2). The pI values of Mgb and Hgb were close to 8.

Comparative homology modeling of the tertiary structures was applied to estimate the structural profiles of zebrafish globins (Fig. 4). The structure of globin X was tentatively modeled based on that of Neugb, because there is no structure available on the database.

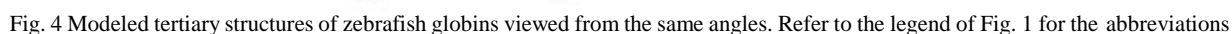


Fig. 5 Alignment of the amino acid sequences of zebrafish and human cytoglobins 1. Refer to the legend of Fig. 1 for the symbols

Fig. 6 Alignment of the amino acid sequences of zebrafish and human neuroglobins. Refer to the legend of Fig. 1 for the symbols

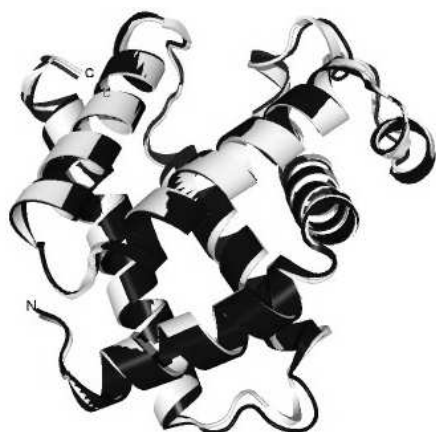


Fig. 7 Superimposed tertiary structures of cytoglobins 1 from zebrafish (grey) and human (black)



Fig. 8 Superimposed structures of neuroglobins from zebrafish (grey) and mouse (dark grey)

The structure of Neugb was used due to the fact that these two proteins were found to be closely related (Fig. 2). Despite the very low sequence identities among the globins (TABLE I) and a large variety of functions, their tertiary structures were found to be very similar. All of them gave typical globin folds consisting of eight helices. Roesner et al. [1] reported that the globin core of globin X is highly conserved based on the analysis on the substitution rates of amino acid residues in fish globins.

The amino acid sequences of Mbs are known to be closely related to their stability such as in autoxidation [20], heme retention [21], structural stability [22,23], thermostability [24], oxygen affinity [11], and so on. Piscine Mbs are generally labile compared with the mammalian counterparts [25]. It is thus expected the structural stabilities of the other fish globins are lower than those of mammalian ones. Since the amount of these proteins present in the respective tissues are very low, recombinant proteins are essential in order to reveal the instability of fish proteins.

Finally, detailed comparison was attempted for the tertiary structure differences of Cygb and Neugb between zebrafish and mammals (human and mice). Because the available structural

data are limited, the structures from the different sources were used in the study. The sequence alignments are shown in Fig. 5 for Cygb and in Fig. 6 for Neugb. Zebrafish Cygb was slightly shorter than the human counterpart, whereas zebrafish Neugb was longer than that of mouse. The modeled structures of Cygb and Neugb are given in Fig. 7 and Fig. 8, respectively. In Fig. 7, the structures of zebrafish and human Cygb are superimposed. Though slight differences were recognized between them, they seem to have very similar tertiary structures. On the other hand, the structures of zebrafish and mouse Neugbs showed essentially the same tertiary structures (Fig. 8).

Neugb content is intrinsically high in fish. The fact might be related with hypoxia tolerance of the nervous system by virtue of immediate availability of O<sub>2</sub> upon the onset of hypoxia. Hypoxia-induced strong enhancement of Mgb protein level has also been reported in several fish species [6,7,26,27]. The results so far obtained suggest that different roles of globins in these fish species having different strategies against hypoxia and other acute abnormal conditions.

All the findings in the present study demonstrated that fish globin proteins are structured very similarly, and that fish globins are hardly distinguishable from the mammalian counterparts as far as the tertiary structures are concerned.

#### REFERENCES

- [1] A. Roesner, "A globin gene of ancient evolutionary origin in lower vertebrates: Evidence for two distinct globin families in animals," *Mol. Biol. Evol.*, vol. 22, 12-20, 2005
- [2] T. Burmester, B. Weich, S. Reinhardt, and T. Hankeln, "A vertebrate globin expressed in the brain," *Nature*, vol. 407, 520-523, 2000
- [3] T. Burmester, B. Ebner, B. Weich, and T. Hankeln, "Cytoglobin: a novel globin type ubiquitously expressed in vertebrate tissues," *Mol. Biol. Evol.*, vol. 19, 416-421, 2002.
- [4] D. Kugelstadt, M. Haberkamp, T. Hankeln, and T. Burmester, "Neuroglobin, cytoglobin, and a novel, eye-specific globin from chicken," *Biochem. Biophys. Res. Commun.*, vol. 325, 719-725, 2004.
- [5] C. Fuchs, T. Burmester, and T. Hankeln, "The amphibian globin gene repertoire as revealed by the *Xenopus* genome," *Cytogenet. Genome Res.*, 112, 296-306, 2006
- [6] A. Roesner, C. Fuchs, T. Hankeln, and T. Burmester, "A globin gene of ancient evolutionary origin in lower vertebrates: evidence for two distinct globin families in animals," *Mol. Biol. Evol.*, vol. 22, 12-20, 2005.
- [7] J. Fraser, L.V. de Mello, D. Ward, H.H. Rees, D.R. Williams, Y. Fang, A. Brass, A.Y. Gracey, and A.R. Cossins, "Hypoxia-inducible myoglobin expression in nonmuscle tissues," *Proc. Natl. Acad. Sci. USA*, vol. 103, 2977-2981, 2006.
- [8] H. Wajcman, L. Kiger, and M. C. Marden, "Structure and function evolution in the superfamily of globins," *Com. Rend. Biol.*, vol. 332, pp.273-282, 2009
- [9] S. Dewilde, L. Kiger, T. Burmester, T. Hankeln, V. Baudin-Creuzat, T. Aerts, M.C. Marden, R. Caubergs, and L. Moens, "Biochemical characterization and ligand binding properties of neuroglobin, a novel member of the globin family," *J. Biol. Chem.*, vol. 276, 2001.
- [10] Y. Enoki, K. Matsumura, Y. Ohga, H. Kohzuki, and M. Hattori, "Oxygen affinities (*p*50) of myoglobins from four vertebrate species (*Canis familiaris*, *Rattus norvegicus*, *Mus musculus* and *Gallus domesticus*) as determined by a kinetic and an equilibrium method," *Comp. Biochem. Physiol. B*, vol. 110, 193-199, 1995.
- [11] R.A. Meyer, H.L. Sweeney, and M.J. Kushmerick, "A simple analysis of the phosphocreatine shuttle," *Am. J. Physiol.*, vol. 246, C365-C377, 1984.
- [12] P.W. Hochachka, "The metabolic implications of intracellular circulation," *Proc. Natl. Acad. Sci. USA*, vol. 96, 12233-12239, 1999.

- [13] S.E. Flonta, S. Arena, A. Pisacane, P. Michieli, and A. Bardelli, "Expression and functional regulation of myoglobin in epithelial cancers," *Am. J. Pathol.*, vol. 175:201-206, 2009.
- [14] M. Brunori, "Myoglobin strikes back," *Protein Sci.*, vol. 19, 195-201, 2010.
- [15] T. Burmester and T. Hankeln, "What is the function of neuroglobin?," *J. Exp. Biol.*, vol. 212, 1423-1428, 2009.
- [16] T.A. Hall, "BioEdit: a user-friendly biological sequence alignment editor and analysis program for Windows 95/98/NT," *Nucleic Acids Symp. Ser.* vol. 41, 95-98, 1999.
- [17] K. Tamura, J. Dudley, M. Nei, and S. Kumar, "MEGA4: Molecular Evolutionary Genetics Analysis (MEGA) software version 4.0," *Mol. Biol. Evol.*, vol. 24, 1596-1599, 2007.
- [18] J. Kyte and R.F. Doolittle, "A simple method for displaying the hydrophobic character of a protein," *J. Mol. Biol.*, vol. 157, 105-132, 1982.
- [19] T. Schwede, J. Kopp, N. Guex, and M.C. Peitsch, "SWISS-MODEL: an automated protein homology-modeling server," *Nucleic Acids Res.*, vol. 31, 3381-3385, 2003.
- [20] J.M. Stewart, J.A. Blakely, P.A. Karpowicz, E. Kalanxhi, B.J. Thatcher, and B.M. Martin, "Unusually weak oxygen binding, physical properties, partial sequence, autoxidation rate and a potential phosphorylation site of beluga whale (*Delphinapterus leucas*) myoglobin," *Comp. Biochem. Physiol. B*, vol. 137, 401-412, 2004.
- [21] E.W. Grunwald and M.P. Richards, "Studies with myoglobin variants indicate that released heme is the primary promoter of lipid oxidation in washed fish muscle," *J. Agric. Food Chem.*, vol. 54, 4452-4460, 2006.
- [22] W.C.B. Regis, J. Fattori, M.M. Santoro, M. Jamin, and C.H.I. Ramos, "On the difference in stability between horse and sperm whale myoglobins," *Arch. Biochem. Biophys.*, 436, 168-177, 2005.
- [23] Y. Ochiai, N. Ueki, and S. Watabe, "Effects of point mutations on the structural stability of tuna myoglobins," *Comp. Biochem. Physiol. B*, vol. 153, 223-228, 2009.
- [24] N. Ueki and Y. Ochiai, "Primary structure and thermostability of bigeye tuna myoglobin in relation to those from other scombridae fish," *Fish. Sci.*, vol. 70, 875-884, 2004.
- [25] H.S. Ajoula, M.T. Wilson, and I.E.G. Morrison, "Functional consequences of haem orientational disorder in sperm-whale and yellow-fin-tuna myoglobins," *Biochem. J.*, vol. 243, 205-210, 1987.
- [26] A. Wawrowski, F. Gerlach, T. Hankein, and T. Burmester, "Changes of globin expression in the Japanese medaka (*Oryzias latipes*) in response to acute and chronic hypoxia," *J. Comp. Physiol. B*, vol. 188, 199-208, 2011.
- [27] A.Roesner, S.A. Mitz, T. Hankeln, and T. Burmester, "Globins and hypoxia adaptation in the goldfish, *Carassius auratus*," *FEBS J.*, vol. 275, 3633-3643, 2008.