Available online at www.sciencedirect.com
ScienceDirect
Gene 398 (2007) 29–34
GENE
www.elsevier.com/locate/gene

Crystal structures of two hemoglobin components from the midge larva *Propsilocerus akamusi* (Orthocladiinae, Diptera)

Takao Kuwada [a,b], Tomokazu Hasegawa [c], Shu Sato [d], Isamu Sato [a,b], Koichi Ishikawa [e], Takashi Takagi [d], Fumio Shishikura [b,f,*]

[a] *Advanced Research Institute for the Sciences and Humanities, Nihon University, Chiyoda-ku, Tokyo 102-8251, Japan*
[b] *Laboratory for Electron Beam Research and Application, Institute of Quantum Science, Nihon University, Funabashi 274-8501, Japan*
[c] *X-ray Research Laboratory, Rigaku Corporation, Akishima 196-8666, Japan*
[d] *Graduate School of Life Sciences, Tohoku University, Sendai 980-8578, Japan*
[e] *Department of Pharmacology, Nihon University School of Medicine, Itabashi-ku, Tokyo 173-8610, Japan*
[f] *Department of Chemistry, Nihon University School of Medicine, Itabashi-ku, Tokyo 173-8610, Japan*

Received 24 October 2006; received in revised form 8 February 2007; accepted 13 February 2007
Available online 10 May 2007

Abstract

The polymorphic components of hemoglobin (Hb) of the midge larva *Propsilocerus akamusi* were classified into two distinct types dependent on their spectroscopic properties, normal absorption (N) and low absorption (L). Analyses of the amino acid sequences of component VII (N-type Hb) and component V (L-type Hb) from *P. akamusi* indicated that one remarkable difference is the replacement of the distal histidine (His) with isoleucine (Ile) in component V. To clarify the structural differences between the two Hb components, we determined the crystal structures of components V and VII at resolutions of 1.64 Å and 1.50 Å, respectively. These crystal structures indicated a short additional helix comprising three amino acid residues at the C-terminal region in component V, and a typical globin fold including eight helices in component VII. Comparison of the heme regions of the Hb components suggests that the structural changes of the heme region in component V on ligation differ from that of usual Hb.

Keywords: X-ray crystallography; Insect hemoglobin; Additional helix; Distal histidine; Amino acid replacement

1. Introduction

Both the biochemical nature and protein structure of invertebrate hemoglobins (Hbs) are unique because they are present in the coelomic fluid or hemolymph as well as in the cellular components (Vinogradov, 1985; Riggs, 1998). Crystallographic analyses of invertebrate Hbs have revealed the characteristics of their tertiary and quaternary structures and structural changes dependent on ligand binding (Bolognesi et al., 1997). The crystal structures of Hb from three species of Diptera, *Chironomus thummi thummi* (Chironomidae), *Gasterophilus intestinalis* (Oestridae), and *Drosophila melanogaster* (Drosophilidae), were determined and the structural features of each Hb component were characterized (Steingemann and Weber, 1979; Pesce et al., 2005; de Sanctis et al., 2005, 2006). However, within one insect species, the structural differences between polymorphic Hb components have not yet been analyzed by X-ray crystallography techniques.

Propsilocerus akamusi (old species name: *Tokunagayusurika akamusi*) is a common species of midge found in eutrophic lakes in Japan that belongs to a different subfamily Orthocladiinae from *C. thummi thummi* in Chironominae. The coelomic fluid of the 4th-instar midge larva contains polymorphic hemoglobin with as many as eleven separable hemoglobin components, as discriminated on DEAE-cellulose column chromatography (Fukuda et al., 1993). All the components are of low molecular mass,

Abbreviations: Hb, hemoglobin; L, low absorption; N, normal absorption; PDB, Protein Data Bank; SAD, single-wavelength anomalous diffraction.

* Corresponding author. Department of Chemistry, Nihon University School of Medicine, 30-1 Oyaguchi-kami-machi, Itabashi-ku, Tokyo 173-8610, Japan. Tel.: +81 3 3972 8111; fax: +81 3 3972 0027.
E-mail address: fshishi@med.nihon-u.ac.jp (F. Shishikura).

0378-1119/$ - see front matter © 2007 Elsevier B.V. All rights reserved.
doi:10.1016/j.gene.2007.02.049

17 kDa or less, and classified into two distinct types on the basis of their spectroscopic properties: low absorbance (L) and normal (N) absorbance. Component V is representative of the six members of the L type and component VII of the N type, they comprise approximately 15% and 30% in the 4th larval instar coelomic fluid, respectively. One of the most remarkable differences between component V and component VII is the replacement of the distal histidine by isoleucine, whereas component VII contains the usual distal His residue. Component V also differs from the known midge larval hemoglobin components in having an intra-molecular disulfide bridge bound by Cys 143 and Cys 148 (Fukuda et al., 1993).

Comparative studies of the tertiary structures of the two distinct types of *P. akamusi* Hb components can help understanding the structural differences among polymorphic Hb components found in invertebrate hemolymph, in particular, insect coelomic fluid (Shishikura et al., 2006). To clarify the tertiary structural differences among polymorphic Hb components of *P. akamusi* coelomic fluid, we performed X-ray crystallographic analyses of the two components. Here we report the crystallographic structures of components V and VII at resolutions of 1.64 Å and 1.50 Å, respectively.

2. Materials and methods

2.1. Crystallization

The two distinct types of Hbs, components V and VII, from the 4th-instar midge larvae of *P. akamusi* were purified as described by Fukuda et al. (1993), and crystallization was carried out separately by hanging drop vapor diffusion techniques at 20 °C. Crystals of component V were grown in a drop containing the protein solution (2.0 μl of 24 mg/ml) mixed with an equal amount of a reservoir solution (25% PEG 3350, 200 mM ammonium sulfate, and 100 mM sodium acetate, pH 4.6). The dimensions of the crystal of component V reached $0.7 \times 0.4 \times 0.05$ mm in about 1 week.

At the same time, crystals of component VII were grown as clusters of thin plates in the drop mixed with equal volumes of the protein solution (2.0 μl of 15 mg/ml) and a reservoir solution (30% PEG 3350, 200 mM $MgCl_2$, and 100 mM Tris–HCl, pH 8.5) within 1 week. Because these crystals were unsuitable for X-ray diffraction analysis, we employed a seeding experiment. The clusters of component VII thin plates were crushed and used as microseeds. Using a fine hair, the seeds of fine crystals were transferred into drops of a protein solution mixed with the reservoir solution containing 25% PEG 3350, 150 mM $MgCl_2$, and 100 mM Tris–HCl, pH 8.5. Finally, component VII crystals grew to about $0.2 \times 0.2 \times 0.05$ mm within a day.

2.2. X-ray diffraction analyses

X-ray diffraction data of Hb components were collected from flash-cooled crystals at 100 K using a Rigaku R-AXIS VII image plate detector with CuKα radiation from rotating anode generators: Rigaku FR-$E_{SuperBright}$ for component V and Rigaku RA-Macro7 for component VII (Rigaku Co., Tokyo, Japan). A crystal of component V was immersed in cryoprotectant solution containing 20% glycerol in the reservoir solution for 10–15 s, picked up with a loop, and rapidly transferred to a cold stream. Following that, diffraction data were collected at 1.64 Å resolution. A total of 140,766 observed intensities were measured, 20,177 of which were unique reflections resulting in an R_{merge} of 0.027 and completeness of 97.3%. One crystal of component VII was transferred to a reservoir solution for 24 h before flash-freezing and data collection. A total of 317,903 observed intensities were measured, 23,622 of which were the unique reflections. The resulting dataset was of 100.0% completeness with an R_{merge} of 0.041. Each set of diffraction data was processed using the software program CrystalClear (version 1.3.5; Rigaku Co., Tokyo, Japan). Crystal data and data collection statistics of components V and VII are summarized in Table 1.

2.3. Phasing and model building

The phases of component V were determined by the single-wavelength anomalous diffraction (SAD) method using the anomalous signal from heme irons. The SAD phases of component V were calculated and further refined by density modification using the CNX program suite (MSI Inc., CA, USA). The initial model of component V was built automatically using the program ARP/wARP (Morris et al., 2003) with the SAD phases calculated in CNX. The high quality electron density allowed tracing of the main- and side-chain atoms of most residues except for residues in one N-terminal and one C-terminal. The model was rebuilt manually using the X-BUILD

Table 1
Data collection statistics, crystal data and structure refinement statistics

	Component V	Component VII
Data collection statistics		
Resolution (Å)	49.14–1.64 (1.70–1.64)	33.77–1.50 (1.55–1.50)
No. of recorded observations	140,766	317,903
No. of unique reflections	20,177	23,622
R_{merge}	0.027 (0.068)	0.041 (0.244)
Completeness (%)	97.3 (94.2)	100.0 (100.0)
Multiplicity	6.98 (6.91)	13.46 (12.55)
$I/\sigma I$	48.6 (15.8)	31.0 (8.5)
Crystal parameters		
Space group	$P2_12_12$	$C222_1$
a (Å)	65.14	42.01
b (Å)	74.85	69.11
c (Å)	33.42	99.65
Model statistics		
R-factor/R_{free}	0.194 / 0.2140	0.194 / 0.227
No. of protein atoms	1213	1169
No. of heme atoms	43	43
No. of water atoms	226	262
rmsd bond lengths	0.006	0.015
rmsd bond angles	1.024	1.470
Ramachandram plot (%)		
Most favored	96.2	97.0
Additional allowed	3.8	3.0

Note: Values in parentheses are for the highest resolution shell.

program (QUANTA, MSI Inc.), guided by a simulated annealing omit map and/or 2Fo–Fc map. The model fitted to the electron density map was refined by CNX functions: it included simulated annealing, energy minimization, and *B*-factor refinements. Following several rounds of manual intervention and CNX refinement, the heme and water molecules were added to the model. The water molecules were added using the X-SOLVENT program (QUANTA, MSI Inc.), guided by Fo–Fc maps contoured at 3σ. The model including water molecules was refined by energy minimization and *B*-factor refinements. On the other hand, the SAD phases of component VII were determined using the anomalous signal from heme irons. With the SAD phases, the initial model of component VII was built using the program ARP/wARP (Morris et al., 2003), it included most residues except for one N-terminal and one C-terminal. The subsequent refinement of component VII model was conducted as described above.

The qualities of the final models were assessed using the program PROCHECK (Laskowski et al., 1993). The atomic coordinates and structure factors of *P. akamusi* Hb components

```
                    ———A———   ———B———
DmHb      ------MNSDEVQLIKKTWEIPVATPTDSGAAILTQFFNRF
GiHb1     ------MNSEEVNDIKRTWEVVAAKMTEAGVEMLKRYFKKY
PaHbV     --AFVGLSDSEEKLVRDAWAPIHGDLQGTANTVFYNYLKKY
PaHbVII   DPTWVDMEAGDIALVKSSWAQIHDKE----VDILYNFFKSY
CttHbIII  ------LSADQISTVQASFDKVKGDP----VGILYAVFKAD
                                  30
          —C—      —D—  ————E————
DmHb      PSNLEKFP-FRDVPLEELSGNARFRAHAGRIIRVFDESIQV
GiHb1     PHNLNHFPWFKEIPFDDLPENARFKTHGTRILRQVDEGVKA
PaHbV     PSNQDKFETLKGHPLDEVKDTANFKLIAGRIFTIFDNCVKN
PaHbVII   PASQAKFSAFAGKDLESLKDTAPFALHATRIVSVINEAIAL
CttHbIII  PSIMAKFTQFAGKDLESIKGTAPFEIHANRIVGFFSKIIGE
                    60
          ————F————   ———G———
DmHb      LGQDGDLEKLDEIWTKIAVSHIPRTVSKESYNQLKGVILDV
GiHb1     LSVDFGDKKFDDVWKKLAQTHHEKKVERRSYNELKDIIIEV
PaHbV     VG---NDKGFQKVIADMSGPHVARPITHGSYNDLRGVIYDS
PaHbVII   MGVAENRPALKNVLKQQGINHKGRGVTAAHFEEFETALEAF
CttHbIII  L------PNIEADVNTFVASHKPRGVTHDQLNNFRAGFVSY
               90                          120
          -G-  ————H————
DmHb      LTAASS-LDESQAATWAKLVDHVYAIIFKAIDDDGNAK-
GiHb1     VCSCVK-LNEKQVHAYHKFFDRAYDIAFAEMAKMG----
PaHbV     MHLD-----STHGAAWNKMMDNFFYVFYECLDGRCSQFS
PaHbVII   LESHASGYNAGTKKAWDSAFNNMYSVVFPEL--------
CttHbIII  MKAHTD--FAGAEAAWGATLDTFFGMIFSKM--------
                         150
```

Fig. 1. Sequence alignment of insect Hbs from *P. akamusi*, *D. melanogaster*, *G. intestinalis*, and *C. thummi thummi*. The helices of the typical globin fold are labeled A though H above the alignment. Amino acid residues in each helix are highlighted in gray. Data on the primary and the secondary structures for each Hb component were excerpted from the Macromolecular Structure Database (http://www.ebi.ac.uk/msd/). Sequence alignment was created with the help of the ClustalX program (Thompson et al., 1997). DmHb, *D. melanogaster* Hb (PDB code 2BK9); GiHb1, *G. intestinalis* Hb1 (2C0K); PaHbV, *P. akamusi* Hb component V; PaHbVII, *P. akamusi* Hb component VII; CttHbIII, *C. thummi thummi* HbIII (1ECA).

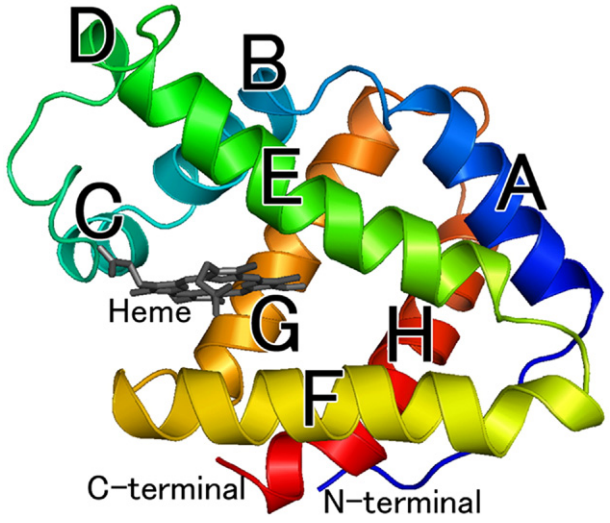


Fig. 2. The overall structure of Hb component VII from *P. akamusi*. The eight helices are labeled A–H. The structural figure was drawn with Pymol (http://pymol.sourceforge.net/).

have been deposited with the RCSB Protein Data Bank with the PDB codes of 1X3K for component V and 1X46 for component VII.

3. Results and discussion

3.1. Refined models of Hb components from P. akamusi

The crystal of component V (L-type Hb) from *P. akamusi* belongs to the orthorhombic space group $P2_12_12$ with unit cell dimensions of $a = 65.14$ Å, $b = 74.85$ Å, and $c = 33.42$ Å. The V_M value was calculated to 2.37 Å^3/Da, suggesting that there is one monomeric molecule in an asymmetric unit. The model of

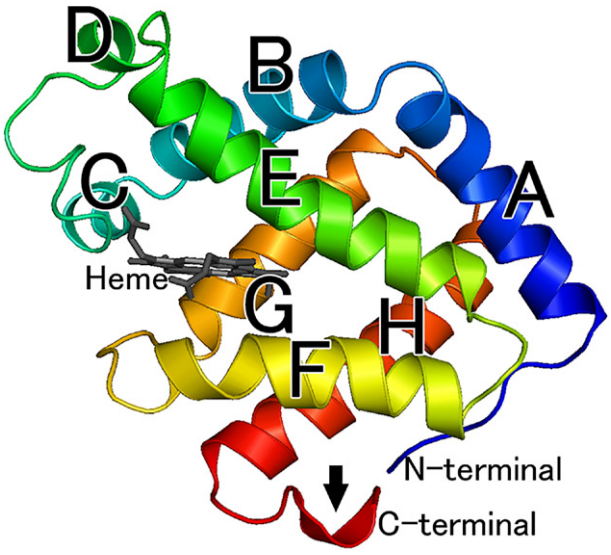


Fig. 3. The overall structure of Hb component V from *P. akamusi*. The eight helices are labeled A–H. An arrow indicates the C-terminal additional helix (post-H helix). The structural figure was drawn with Pymol (http://pymol.sourceforge.net/).

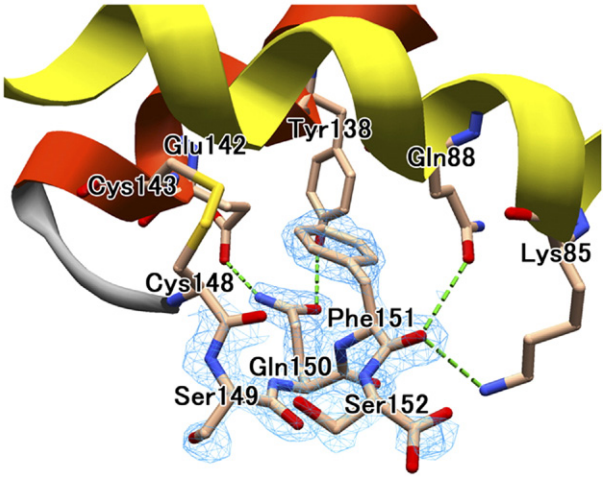


Fig. 4. The electron density around the additional helix at C-terminal region of component V. Amino acid residues at the C-terminal, the F helix (yellow), and the H helix (orange) are shown by sticks. Light green broken lines imply hydrogen bonds between the amino acid residues at the C-terminal and that of the F or H helices. The structural figure was drawn with DeepView (Guex and Peitsch, 1997). (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

component V was solved by the SAD method and refined at 1.64 Å resolution to an *R*-factor of 19.4% and an R_{free} of 21.4%. The final model of component V consisted of a complete polypeptide chain (152 amino acid residues), one heme group, and 226 water molecules.

The crystal of component VII (N-type Hb) belongs to the orthorhombic space group $C222_1$ with unit cell dimensions of a = 42.01 Å, b = 69.11 Å, and c = 99.65 Å. The asymmetric unit contains one monomeric molecule corresponding to a V_M of 2.19 Å^3/Da. The refined model of component VII solved by the SAD method contained all 150 amino acid residues of the native protein, one heme group, and 262 water molecules, yielding an *R*-factor of 19.4% and an R_{free} of 22.7% at 1.50 Å resolution. Detailed refinement statistics and geometric properties of the final models of *P. akamusi* Hb components are summarized in Table 1, showing that each was well refined with excellent stereochemistry.

3.2. Overall structures of *P. akamusi* Hb components

In general, hemoglobin components comprise eight helices (named A–H from N to C terminus) and corner regions connecting helices, and insect Hbs also possess this typical globin fold (Steingemann and Weber, 1979; Pesce et al., 2005; de Sanctis et al., 2005) (Fig. 1). Fig. 2 shows the crystal structure of component VII. It has the typical globin fold based on eight α helices except for residues 82–84 in the F helix and residues 147–149 in the H helix which display a 3_{10} helix conformation. Fukuda et al. (1993) reported that component VII showed higher sequence similarity to *Chironomus* Hb components than component V. When the secondary structures were compared, component VII resembled component III of *Chironomus* Hb in the A, B, and C helices. The three helices, A, B, and C in both component VII of *P. akamusi* and component III of *C. thummi thummi*, differ from Hbs in other insects in that they are 3, 4, and 2 residues shorter (Fig. 1).

Fig. 3 shows the crystal structure of component V. It contains the same α helices, A to H in which the only differences occur in the C helix and residues 96–98 in the F helix which display a 3_{10} helix conformation. The presence of an additional helix at the C-terminal of component V is particularly remarkable. The helix is a one-turn long 3_{10} helix containing three residues (Ser149–Phe151). Fig. 4 shows the electron density in the refined model around the C-terminal region of component V. An additional helix is next to the F and H helices dependent on several noncovalent bonds and a disulfide bond between Cys143 and Cys148; this additional helix does not provide crystal contacts between neighbor molecules. The additional helix located at the N-terminal region has been observed in dimeric Hb (Hb I) from the blood clam, *Scapharca inaequivalvis*, and was named the pre-A helix (Royer, 1994). We call an

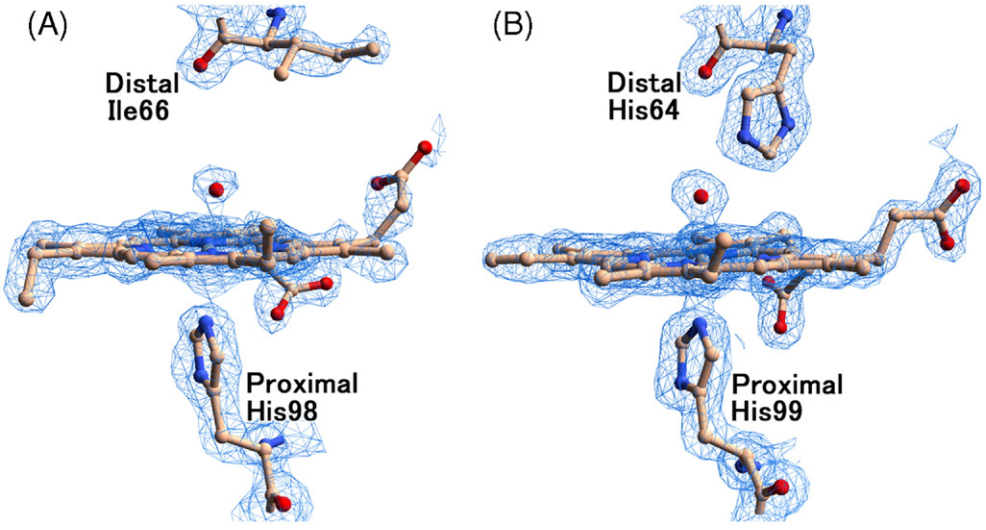


Fig. 5. Electron density maps (contoured at 1.5σ) corresponding to component V (A) and VII (B) heme regions. Distal His (Ile for component V), proximal His, and heme groups are shown by bolls and sticks. The structural figure was drawn with DeepView (Guex and Peitsch, 1997).

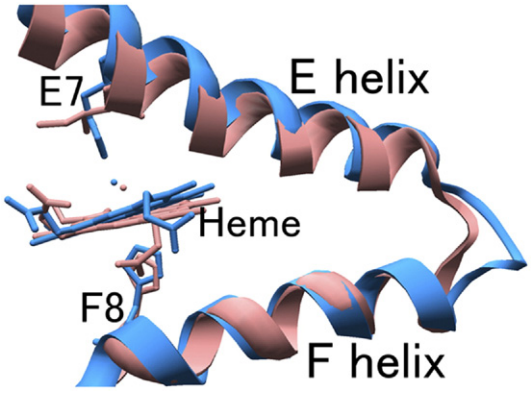


Fig. 6. Overlapping of component V (pink) and component VII (blue) based on the F helix. Distal His (Ile for component V) (E7), proximal His (F8), and heme groups are shown by sticks. Structure superimpositions and figure preparation were performed with DeepView (Guex and Peitsch, 1997). (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

additional helix of *P. akamusi* Hb the post-H helix, and we believe that it has no homologous relationships with any other known helix of Hb components. Its function cannot be confirmed at this stage, still the modification of the post-H helix associated with the heme state is remarkable because it contacts to the heme region.

3.3. Comparison of heme region

As shown in Fig. 5, electron density maps for the heme ligand are clearly evident in both components V and VII. Both ligands seem to be water molecules because of the shape and size of the electron density, and because there is no attention to ligand molecules on crystallization, which causes to form met-forms by autooxidation from oxy-forms (Ladner et al., 1977). Water molecules as ligand molecules are well fitted to electron density and display *B*-factors of 12.4 A^2 and 12.2 A^2 for components V and VII. The existence of these ligand molecules in *P. akamusi* Hb components suggests some interesting possibilities regarding the conformation of the heme region.

In vertebrate α2β2-type Hbs, the structural changes dependent on ligand binding are remarkable, thus the heme geometry of ligated Hb differs from that of unligated Hbs. Our findings indicate that both Hb components from *P. akamusi* are in the ligated form. However, overlapping the two components from *P. akamusi* reveals the differences between them. Fig. 6 indicates the narrow heme region of component V due to the close contact between the E and F helices: the distance between the Cα atom of E7 and the Cα atom of F8 is 13.58 Å, 0.69 Å shorter than the value in component VII (14.27 Å, nearly 14.5 Å (±0.2) in various Hb components (Riccio et al., 2002)); this is reminiscent of a scissors-like motion of the EF segment on hemichrome formation in the Antarctic fish Hb, where the distance between the Cα atoms of distal and proximal His is 12.5 Å (Riccio et al., 2002). Difference of the coordination geometry of the heme region between components V and VII is also significant.

The coordination geometry of the heme region in Table 2 is usual to illustrate the structural differences between ligated and unligated Hbs. Comparison of these parameters indicates that the heme geometry of component VII is similar to that of human ligated Hbs. In component V, the location of the iron atom in the heme plate is almost the same as in average ligated Hbs, and the distance between the heme iron and the ligand also is similar to those of ligated human Hbs. At the same time, component V closely resembles the unligated Hb components in the tilting of the imidazole of the proximal His. In many Hbs, the heme iron moves into the heme plate upon ligation, this brings about the movement of the imidazole of the proximal His. This structural feature of proximal His was not observed in ligated component V. It is possible that the structural changes in the heme on ligation do not trigger the associated movement of the imidazole of the proximal His in component V.

One of the most remarkable differences between component V and many other Hbs is related to the substitution of the usual distal His residue with Ile (Fukuda et al., 1993). Replacement of the distal His might cause unusual structural changes upon ligand binding. In fact, Hb components from a marine blood worm and a hagfish, in which each distal His is replaced, show small structural changes of heme region on ligation (Braden

Table 2
The heme geometry of *P. akamusi* and human Hbs

	Ligated Hb						Unligated (deoxy) Hb			
	Human				*P. akamusi*		Human			
	Oxy-α	Oxy-β	CO-α	CO-β	VII	V	α1	β1	α2	β2
Distance (Å)										
Fe-Ligand	1.82	1.78	1.74	1.71	2.08	1.98	–	–	–	–
Fe-Pn	0.05	0.04	−0.04	−0.07	0.01	0.07	0.39	0.35	0.31	0.32
Fe-F8Nε2	2.07	2.06	2.11	2.12	2.07	2.34	2.2	2.19	2.21	2.16
F8[Cδ-N3]-[Cε-N1]	0.14	−0.01	0.11	0.02	−0.11	0.49	0.55	0.53	0.46	0.48

Fe-Ligand, the distance between the heme iron and the ligand molecule. For calculation of this parameter in *P. akamusi* Hbs, a water molecule was treated as the ligand molecule. Fe-Pn, the distance between the heme iron and the center of the four porphyrin nitrogen atoms. This parameter represents the movement of the heme iron upon ligand binding. Fe-F8Nε2, the distance between the heme iron and the Nε atom of the proximal His. F8[Cδ-N3]-[Cε-N1], the difference between two distances representing a tilting of the imidazole of the proximal His. One is between Cδ of the proximal His and the nitrogen of pyrrole 3 of porphyrin, and the other is between Cε of the proximal His and the nitrogen of pyrrole 1. This tilting links the shift of the F helix toward the FG corner upon ligand binding (Mito et al., 2002). For comparison, the parameters from the structures of human Oxy-Hb (PDB code 2DN1), CO-Hb (2DN3), and Deoxy-Hb (2DN2) (Park et al., 2006) are also included.

et al., 1994; Mito et al., 2002). Our comparative studies of the crystal structure of *P. akamusi* Hb components revealed the characteristics of the tertiary structures of each component, and especially the interesting heme region of component V. It seems that the heme region of aquomet component V is in the intermediate state between a typical ligated (aquomet) and unligated or hemichrome form. Replacement of the distal His might play a part in the formation of the equivocal state of the heme region in component V. Crystallographic analyses of *P. akamusi* Hb in a variety of ligand state can prove helpful in understanding the details of the heme ligation and the associated structural changes.

Acknowledgements

This work was supported by the "Academic Frontier" Project for Private Universities: matching fund subsidy from MEXT (Ministry of Education, Culture, Sports, Science and Technology), Japan, 2000–2004 and 2005–2007 and in part by Nihon University Joint Research Grant (2005).

References

Braden, B.C., Arents, G., Padlan, E.A., Love, W.E., 1994. *Glycera dibranchiata* hemoglobin. X-ray structure of carbon monoxide hemoglobin at 1.5 Å resolution. J. Mol. Biol. 238, 42–53.

Bolognesi, M., Bordo, D., Rizzi, M., Tarricone, C., Ascenzi, P., 1997. Nonvertebrate hemoglobins: structural bases for reactivity. Prog. Biophys. Mol. Biol. 68, 29–68.

de Sanctis, D., et al., 2005. Bishistidyl heme hexacoordination, a key structural property in *Drosophila melanogaster* hemoglobin. J. Biol. Chem. 280, 27222–27229.

de Sanctis, D., et al., 2006. Cyanide binding and heme cavity conformational transitions in *Drosophila melanogaster* hexacoordinate hemoglobin. Biochemistry 45, 10054–10061.

Fukuda, M., Takagi, T., Shikama, K., 1993. Polymorphic hemoglobin from a midge larva (*Tokunagayusurika akamusi*) can be divided into two different types. Biochim. Biophys. Acta 1157, 185–191.

Guex, N., Peitsch, M.C., 1997. SWISS-MODEL and the Swiss-PdbViewer: an environment for comparative protein modeling. Electrophoresis 18, 2714–2723.

Ladner, R.C., Heidner, E.J., Perutz, M.F., 1977. The structure of horse methaemoglobin at 2.0 Å resolution. J. Mol. Biol. 114, 385–414.

Laskowski, R.A., MacArthur, M.W., Moss, D.S., Thornton, J.M., 1993. PROCHECK: a program to check the stereochemical quality of protein structure. J. Appl. Crystallogr. 26, 283–291.

Mito, M., et al., 2002. Crystal structures of deoxy- and carbonmonoxyhemoglobin F1 from the hagfish *Eptatretus burgeri*. J. Biol. Chem. 277, 21898–21905.

Morris, R.J., Perrakis, A., Lamzin, V.S., 2003. ARP/wARP and automatic interpretation of protein electron density maps. Methods Enzymol. 374, 229–244.

Park, S.Y., Yokoyama, T., Shibayama, N., Shiro, Y., Tame, T.J., 2006. 1.25 Å resolution crystal structures of human haemoglobin in the oxy, deoxy and carbonmonoxy forms. J. Mol. Biol. 360, 690–701.

Pesce, A., et al., 2005. Modulation of oxygen binding to insect hemoglobins: the structure of hemoglobin from the botfly *Gasterophilus intestinalis*. Protein Sci. 14, 3057–3063.

Riccio, A., Vitagliano, L., di Prisco, G., Zagari, A., Mazzarella, L., 2002. The crystal structure of a tetrameric hemoglobin in a partial hemichrome state. Proc. Natl. Acad. Sci. U. S. A. 99, 9801–9806.

Riggs, A.F., 1998. Self-association, cooperativity and supercooperativity of oxygen binding by hemoglobins. J. Exp. Biol. 201, 1073–1084.

Royer Jr., W.E., 1994. High-resolution crystallographic analysis of a co-operative dimeric hemoglobin. J. Mol. Biol. 235, 657–681.

Shishikura, F., et al., 2006. Crystal structures of insect hemoglobin components from *Propsilocerus akamusi* (Orthocladiinae, Deptera). XIV International Conference on Dioxygen Binding and Sensing Proteins, Stazione Zoologica Anton Dohr, Naples, September 3–7, p. 95.

Steingemann, W., Weber, E., 1979. Structure of erythrocruorin in different ligand states refined at 1.4 Å resolution. J. Mol. Biol. 127, 309–338.

Thompson, J.D., Gibson, T.J., Plewniak, F., Jeanmougin, F., Higgins, D.G., 1997. The ClustalX windows interface: flexible strategies for multiple sequence alignment aided by quality analysis tools. Nucleic Acids Res. 24, 4876–4882.

Vinogradov, S.N., 1985. The structure of invertebrate extracellular hemoglobins (erythrocruorins and chlorocruorins). Comp. Biochem. Physiol., B 82, 1–15.

Available online at www.sciencedirect.com
SCIENCE DIRECT
Comparative Biochemistry and Physiology, Part B 142 (2005) 258 – 268
CBP
www.elsevier.com/locate/cbpb

Axolotl hemoglobin: cDNA-derived amino acid sequences of two α globins and a β globin from an adult *Ambystoma mexicanum*

Fumio Shishikura [a,*], Hiro-aki Takeuchi [b], Takatoshi Nagai [c]

[a] *Department of Biology, Nihon University School of Medicine, Itabashi-ku, Tokyo 173-8610, Japan*
[b] *Department of Biology and Geosciences, Faculty of Science, Shizuoka University, Shizuoka 422-8529, Japan*
[c] *Department of Biology, Keio University School of Medicine, Hiyoshi, Yokohama 223-8521, Japan*

Received 4 April 2005; received in revised form 3 July 2005; accepted 5 July 2005
Available online 6 September 2005

Abstract

Erythrocytes of the adult axolotl, *Ambystoma mexicanum*, have multiple hemoglobins. We separated and purified two kinds of hemoglobin, termed major hemoglobin (Hb M) and minor hemoglobin (Hb m), from a five-year-old male by hydrophobic interaction column chromatography on Alkyl Superose. The hemoglobins have two distinct α type globin polypeptides (α^M and α^m) and a common β globin polypeptide, all of which were purified in FPLC on a reversed-phase column after *S*-pyridylethylation. The complete amino acid sequences of the three globin chains were determined separately using nucleotide sequencing with the assistance of protein sequencing. The mature globin molecules were composed of 141 amino acid residues for α^M globin, 143 for α^m globin and 146 for β globin. Comparing primary structures of the five kinds of axolotl globins, including two previously established α type globins from the same species, with other known globins of amphibians and representatives of other vertebrates, we constructed phylogenetic trees for amphibian hemoglobins and tetrapod hemoglobins. The molecular trees indicated that α^M, α^m, β and the previously known α major globin were adult types of globins and the other known α globin was a larval type. The existence of two to four more globins in the axolotl erythrocyte is predicted.

Keywords: Amphibia; Axolotl; *Ambystoma mexicanum*; Evolution; Globin; Hemoglobin; Nucleotide sequence; Primary structure

1. Introduction

Amphibian hemoglobins are of interest in the study of the evolution of vertebrate globin genes, gene expression and gene switching during metamorphosis, and functional properties of aquatic and terrestrial species. Although living amphibians comprise three major orders, Apoda, Urodela (Caudata) and Anura, most investigations have been carried out in anurans (frogs and toads): the bullfrog *Rana catesbeiana* (Hamada et al., 1966; Moss and Ingram, 1968; Aggarwal and Riggs, 1969; Chauvet and Acher, 1972; Baldwin and Riggs, 1974; Maruyama et al., 1980; Watt et al., 1980; Dorn and Broyles, 1982; Maples et al., 1983) and the South African clawed toad *Xenopus laevis* (Hentschel et al., 1979; Jeffreys et al., 1980; Hosbach et al., 1983; Kay et al., 1983; Banville and Williams, 1985; Knöchel et al., 1985; Weber et al., 1991). The Mexican axolotl *Ambystoma mexicanum*, on the other hand, belonging to order Urodela, is an important animal in the biological sciences because they become adult without metamorphosing, a phenomenon termed neotenous adulthood. Although axolotls receive much attention as a counterpart of metamorphosing species, little is known about hemoglobin components, population of globin subunits and primary structures of globins. In our references, there are only two reports on primary structures of axolotl globin polypeptides (Boissel et al., 1980;

* Corresponding author. Tel.: +81 3 3972 8111x2291; fax: +81 3 3972 0027.
E-mail address: fshishi@med.nihon-u.ac.jp (F. Shishikura).

1096-4959/$ - see front matter © 2005 Elsevier Inc. All rights reserved.
doi:10.1016/j.cbpc.2005.07.003

GenBank accession no. AF308869); both were assigned to an adult α type of globin and a larval α type of globin by computer analysis described later.

We describe here the separation of two kinds of hemoglobin components, Hb M (major) and Hb m (minor), from a neotenous adult axolotl *A. mexicanum* and establish the complete nucleotide sequences of the three kinds of globin chains, two of adult α type globins and adult β type globin common to both hemoglobin components.

2. Materials and methods

2.1. Blood cell collection

Blood of an adult axolotl *A. mexicanum* (five-year-old male about 20 cm in body length and 600 g in mass) was drawn from the second branchial artery into physiological saline (0.9% NaCl, 10 mM EDTA, 50 mM Tris–HCl, pH 8.0). The blood was then centrifuged at 1750 ×*g* for 15 min, and the packed cells were divided: one half of the packed cells was lysed with 50 mM Tris–HCl, pH 8.0, containing 10 mM EDTA for preparation of hemoglobin solution, and the other half was used for extraction of total RNA by Purescript RNA Isolation Kit (Gentra Systems, Minneapolis, MN, USA) according to the manufacturer's instructions. All procedures were done at 4 °C and the samples were stored at −80 °C.

2.2. Separation of hemoglobin components

A hemoglobin solution that had been saturated at 40% by adding 60% saturated ammonium sulfate was subjected to an Alkyl Superose column (Pharmacia Biotech, Uppsala, Sweden) equilibrated with 60% saturated ammonium sulfate (183 g/500 mL) in 50 mM ammonium bicarbonate, pH 8.0. The experimental procedures were conducted as previously described (Shishikura and Takami, 2001). Elution was carried out with a gradient of 60–0% saturated ammonium sulfate in the 50 mM ammonium bicarbonate buffer, pH 8.0. The flow rate was maintained at 0.5 mL/min, and fractions of the protein peaks were collected. The fractions were monitored at 415 and 280 nm by two spectrophotometers (Model Uvidec-340Q, Jasco, Tokyo, Japan, and Model 116, Gilson Medical Electronics, Inc., WI, USA), respectively.

2.3. Preparation of globin polypeptides

To separate α and β globin polypeptides, hemoglobin components were *S*-pyridylethylated separately by the method described previously (Friedman et al., 1970). Each of the *S*-pyridylethylated hemoglobin components was then subjected to a Resource column (Pharmacia Biotech) and eluted with a buffered gradient of 0.1% trifluoroacetic acid (TFA; Sigma-Aldrich, Inc., St. Louis, MO, USA) to 60% acetonitrile in 0.08% TFA. All fractions were monitored at 214 and 280 nm. For further purification, re-chromatography on the Resource column was conducted under shallower gradient conditions, as described in our previous report (Shishikura et al., 1987).

2.4. Protein sequencing

All three kinds of globin molecules modified by *S*-pyridylethylation were separately digested with lysyl endopeptidase (Achromobacter Protease I; Wako Pure Chemicals, Tokyo, Japan) at an enzyme/substrate ratio of 1:30 (mol:mol) for 4 h at 37 °C in 0.1 M ammonium bicarbonate, pH 8.2, containing 4 M urea. Peptide fragments derived from each of the parent molecules were separated using a reversed-phase column (Resource RPC, Pharmacia Biotech.) in a buffered gradient of 0.1% TFA to 60% acetonitrile in 0.08% TFA. Flow rates were maintained at 0.3 mL/min. All fractions were monitored at 214 and 280 nm. Re-chromatography of selected peptides was conducted as previously described (Shishikura et al., 1987). Sequencing analyses of these fragments as well as amino (N)-terminal amino acid residues of parent molecules were performed using a gas phase protein sequencer, PPSQ-10 (Shimadzu, Shiga, Japan), equipped with a class LC-10 amino acid analyzer (Shimadzu, Shiga, Japan). Phenylthiohydantoin (PTH)-derivatives from the sequencer were separated and quantified. The peptides were aligned with the assistance of sequence similarities toward the known globin structures of urodeles (Boissel et al., 1980; Kleinschmidt et al., 1988; AF308869).

2.5. Isolation of total RNA and purification of mRNAs

Total RNA was extracted from the blood cells by Purescript Total RNA Isolation Kits (Gentra Systems), and the mRNA fraction that included the three kinds of globin mRNAs was isolated with a Takara Oligotex™ -dT30<Super>mRNA Purification Kit (Takara Bio, Shiga, Japan).

2.6. Primer design

Degenerate primers were designed based on the amino acid sequences of lysyl endopeptidase-digested fragments of parent molecules as listed in Appendix 1. In order to sequence the PCR amplified fragments with a BigDye Terminator v. 1.1 Cycle sequencing Kit (Applied Biosystems, Foster City, CA, USA), the degenerate oligo-nucleotide primers were tailed with the pUC/M13 forward or pUC/M13 reverse sequencing primer tail (forward 17-mer: 5′-GTAAAACGACGGCCAGT-3′, Sigma-Aldrich Japan, Tokyo, Japan, and reverse 17-mer: 5′-CAGGAAACAGCTATGAC-3′, Promega, Tokyo, Japan).

2.7. Nucleotide sequencing

Single-strand cDNAs were synthesized with a Takara RNA PCR Kit (v. 2.1) using the Oligo dT-Adaptor Primer (M13 primer M4, 17-mer: 5′-GTTTCCCAGTCACGA-CT_{15}-3′), according to the manufacturer's instructions (Takara Bio).

For PCR amplification of the 3′ region of the cDNAs, the primers used were the adaptor and a redundant oligomer based on the N-terminal amino acid sequence of each globin (see Appendix 1). The second PCR-amplification was conducted with a nested PCR primer (a redundant oligomer) and the adaptor (Appendix 1). One major fragment was detected on agarose gel electrophoresis in each PCR. Then, the amplified fragment was extracted by a GenElute™ Agarose Spin Column (Sigma-Aldrich), and sequenced directly with a BigDye Terminator v. 1.1 Cycle Sequencing Kit. The rest of the unknown sequence of the 3′ end was afterwards confirmed by 3′ RACE (Frohman et al., 1988), with the adaptor and a gene-specific primer (Appendix 1).

For PCR amplification of the 5′ region of cDNAs, gene-specific primers with or without 5′ monophosphate as listed in Appendix 1, were designed in order to extend the sequences in the 5′ ends using a Takara 5′-Full RACE Core Set (Takara Bio) according to the manufacturer's instructions. All forward and reverse primers, except for the Oligo dT, listed in Appendix 1, were tagged with the pUC/M13 sequencing primers.

2.8. Computer analysis

A multiple alignment program, Clustal X v. 1.81 (Jeanmougin and Thompson, 1998), was used to align the three kinds of adult axolotl globins, previously established known sequences of amphibians and those from representative species of tetrapods. Two phylogenetic trees were constructed by the neighbor-joining method (Saitou and Nei, 1987), and the molecular trees were drawn by the NJ-prot program stored in Clustal X.

3. Results and discussion

3.1. Two hemoglobin components of adult axolotl

Separation of an intact component from multiple hemoglobins is a prerequisite to investigation of the properties of individual hemoglobin component from sources with such complicated life cycles as amphibians. Several investigators have succeeded: in adult anurans (toads and frogs), four hemoglobin components have been separated and characterized (Elli et al., 1970; MacLean and Jurd, 1971; Tam et al., 1993). In adult urodeles, four and eight intact hemoglobin components were separated from the newt *Triturus cristatus* (Salamandroidea) by anion-exchange column chromatography on DEAE-Sephadex (Grasso et al., 1979) and by analytical isoelectrofocusing (Koussoulakos et al., 1986), respectively. However, how many hemoglobin components are present in erythrocytes of the axolotl *A. mexicanum*, which belongs to the Ambystomoidea, in the embryo or larval stages, or even in the neotenous adult stage, has not been investigated.

Fig. 1 shows two well-delimited peaks corresponding to a major (Hb M) and a minor hemoglobin (Hb m) from erythrocytes of an adult axolotl. The two peaks, Hb M and Hb m, were detected at 280 and 415 nm. The peaks were separately pooled as shown by bars, and then, each pooled fraction was re-chromatographed on the same column under the same conditions (insets of Fig. 1). They exist in a ratio of about 5 (Hb M) : 2 (Hb m) based on chromatogram area calculation. This value varied from 5 : 1 to 5 : 3 depending on sample preparation. Several small distinct peaks (indicated by two arrows in Fig. 1) were still

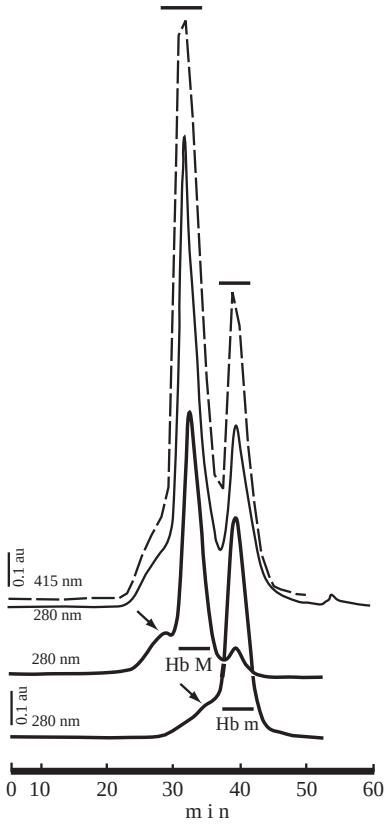


Fig. 1. Alkyl Superose HR5/5 column chromatography of intact hemoglobin of the axolotl, *A. mexicanum*. The hemoglobin solution was saturated at 40% by adding 60% saturated ammonium sulfate and applied to an Alkyl Superose HR5/5 column equilibrated with 60% saturated ammonium sulfate buffer. After washing with the buffer, the adsorbed proteins were eluted with a decreasing concentration of ammonium sulfate, pH 8.0. Concentration of ammonium sulfate and time: 60% (0 min) — 42% (10 min) — 30% (40 min) — 0% (60 min). Insets show the purification of Hb M and Hb m. Bars indicate the pooled (upper) and purified fractions (lower) of major and minor peaks, Hb M and Hb m, respectively.

detected in both purified and unpurified fractions, but no further purification was carried out because the other hemoglobin peaks obtained were very poor and it was impossible to characterize them further. Similar two peaks (Hb M and Hb m) were also detected in the case of hemoglobin solution prepared from a five-year-old female (about 18 cm in body length and 500 g in mass: data not shown here).

The Alkyl Superose HR 5/5 column that we used first in preparation of intact reptilian hemoglobin components (Shishikura, 2002; Kuwada et al., 2003) was used for preparation of hemoglobin components from axolotl blood. The separation pattern is superior to that obtained by isoelectric focusing (Hattingh and Bartels, 1973). The chemistry of the separation is based on hydrophilic interactions with the gel matrix and biological macromolecules. With the fraction, we are currently generating crystals of the intact hemoglobin components, Hb M and Hb m, of the axolotl hemoglobin.

3.2. Three globin chains, sequences, and alignment

Like birds, reptiles and mammals, amphibians produce different globin polypeptides during their life cycles (McCutcheon, 1936; McCutcheon and Hall, 1937; Gratzer and Allison, 1960; Bunn and Forget, 1986). Among them, the amphibians metamorphose dramatically, where the pattern of globin polypeptides of anurans (frogs and toads) is replaced completely by that of adult globin polypeptides (Hentschel et al., 1979; Maples et al., 1988). Some urodeles, on the other hand, coexpress larval and adult globin mRNAs in a single erythroid cell (Yamaguchi et al., 2000).

In contrast, Grasso et al. (1979) described at least 6–7 globin polypeptides in the adult urodele *T. cristatus*, of which Kleinschmidt et al. (1988) sequenced four chains including two α globin chains and two β globin chains. We purified one α and one β globin chain from the axolotl Hb M and Hb m, respectively, and then sequenced some of the peptide fragments derived from the parent molecules digested with lysyl endopeptidase (chromatograms of peptide maps by lysyl endopeptidase digestion are not shown here). Appendix 2 shows the results of amino acid sequence analyses, in which the two α globin chains were distinct, but the two β globin chains so far analyzed (only one sequence is shown here) had identical sequences. This reminds us that the two hemoglobin components, B and C, of the adult *R. catesbeiana* share a common β globin chain but have different α globin chains (Tam et al., 1986; Smith et al., 1993). The number of globin polypeptides we found in the adult axolotl erythrocytes confirmed the study of Ducibella (1974), who purified three globin subunits from the denatured hemoglobin solution of the neotenous adult, with apparent molecular weights of 15,000, 19,000 and 24,500 using sodium dodecyl sulfate-polyacrylamide gel electrophoresis.

Using the primers listed in Appendix 1, cDNA-fragments were amplified by PCR. Fig. 2 shows one major band and several very faint bands in each lane; each of the major fragments was extracted from the agarose gels and sequenced separately. Appendix 2 (A, B and C) shows the results of nucleotide sequencing of twelve cDNA-fragments providing enough information to determine complete nucleotide sequences of globin cDNAs. The entire coding regions of the three kinds of adult *A.*

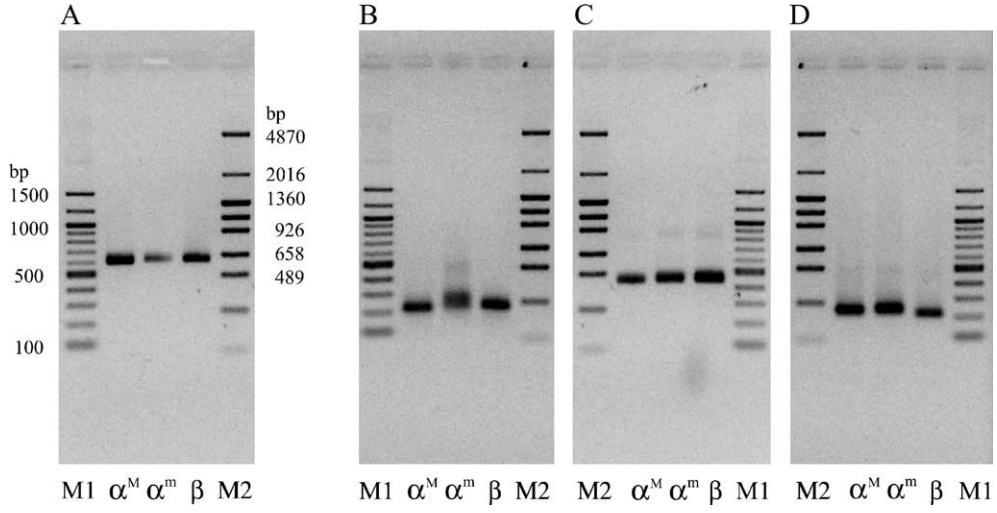


Fig. 2. Agarose gel electrophoresis (1.5% gels) of the PCR products amplified from three kinds of cDNAs using primers as listed in Appendix 1. A: RT-PCR, B: Nested PCR, C: 3′ RACE, and D: 5′ cRACE. The major fragment in each lane, except for the M1 and M2 lanes, was extracted from the gel and sequenced. PCR conditions: 30 cycles each consisting of 30 s at 94 °C for denaturation, 30 s at 60 °C for annealing, and 1 min at 72 °C for primer extension. M1 (100 bp ladder) and M2 (pHY) are DNA markers.

mexicanum globin genes have been deposited in the GenBank/DDBJ/EMBL (under DDBJ accession nos: AB185144 for α^M globin, AB185145 for α^m globin, AB185146 for β globin). The amino acid sequences deduced from the cDNA nucleotide sequences were identical with those determined by peptide sequencing analyses, whose complete primary structures are also shown in Appendix 2. Their sequences reinforce each other. The nucleotide sequences of mature proteins are composed of 141 amino acid residues for α^M globin, 143 for α^m globin, and 146 for β globin. Boissel et al. (1980) reported an extra residue for the α globin chain of the major component of *A. mexicanum* compared with that of ordinary α globin chains in tetrapods; however, this is not curious because one or two longer residues were recently found in newts *T. cristatus* α globin chain with 142 (Kleinschmidt et al., 1988) and *Taricha granulosa* α globin chain with 142 (Coates et al., 1977), and α^m of this study with 143 amino acid residues.

Fig. 3 shows an alignment in which two previously established globin sequences (Boissel et al., 1980; AF308869) and three other globin chains of the present study (α^M, α^m and β) are comparable to those of human α and β globin chains. Hereafter we would like to use the tentative nomenclature of globins of the axolotl: α-1 for the α globin chain named by Boissel et al. (1980) and α-2 for the α globin of AF308869. Compared with each other, four groups based on % homology can be detected: the first group is 61–70% homologous (α^m versus α-1; α^M versus α^m; α^M versus α-1), the second group is 51–57% homologous (α^m versus human α; α^M versus human α; α-1 versus human α), the third group is 47–49% homologous (α^m versus α-2; α^M versus α-2; α-1 versus α-2) and the fourth group is 28–38% homologous (α-2 versus human α;

```
Residue No.     1                          31
                  *       *            * *   * *          **         *
αM globin       --VLTAEDKANVKSAWDHIKGHEEAFGAEALCRMFDSEPVTKTYFGGK------DISEES
αm globin       MVPLSAGDKANVKAVWDHVKGHEEAFGADALHRCFKIDCQTQTYFPGK------DLNEGS
α-1             -FKLSGEDKANVKAVWDHVKGHEDAFGHEALGRMFTGIEQTHTYFPDK------DLNEGS
α-2             -MTLTAEDKALVVVLWGKIAGHTDALGGEALDRLFACFGQSRTYFSHF------DLSPGS

β globin        -VHLTAEERKDVGAILGKVN--VDALGGQCLARLMCVYPWSRRYFPDFGDMSTCDAICHN

                                                    ## ##
                                              o  ooo$$ $$$ $
Human α         --VLSPADKTNVKAAWGKVGAHAGEYDAEALERMFLSFPTTKTYFPHF------DLSHGS
                                                    ##
                                              o  ooo$$  $  $
Human β         -VHLTPEEKSAVTALWGKVN--VDEVGGEALGRLLVVYPWTQRFFESFGDLSTPDAVMGN
                                                   $$
Helix position   NH1               B1               C1     CD1     D1     E1

Residue No.     61                         91
                  ** **      *  * *        *   *   * ***  *
αM globin       SYLHSHGKKVMCALTNAVAHIDNIEACLDKLSDTHAHELMVDPTNFPRLGHNILLVIGIH
αm globin       AFLHSHGKKVMSALTNAVAHIDDLEAALSKLIDKHAHDLMVDPANFVLLNHHILAVLAMH
α-1             FALHSHGKKVMGALSNAVAHIDDLEATLVKLSDKHAHDLMVDPAEFPRLAEDILVVLGFH
α-2             ADVKRHGGKVLNAVGEAAKHIDDLDMGLSKLSDLHAYNLRVDPGNFQLLSHCIQTVLAAH

β globin        ARVLAHGAKVMRSVCEATKHLDNLQEYYADLSSTHCLKLFVDPQNFKLFGRIVVVCLAQT

                     #  ##                    #  ##   # #   ##  #
                                                $$ $$$$     o   o  oo
Human α         AQVKGHGKKVADALTNAVAHVDDMPNALSALSDLHAHKLRVDPVNFKLLSHCLLVTLAAH

                     #  ##  #              #  #  ##   # #   ##  #
                                                 $$$ $   $  o   o  oo
Human β         PKVKAHGKKVLGAFSDGLAHLDNLKGTFATLSELHCDKLHVDPENFRLLGNVLVCVLAHH
Helix position                   EF1      F1      FG1   G1

Residue No.     121                        149
                  *       *            *
αM globin       MPQLLTCAMHCSLDKFLCQVAEVLTSKYR  (This study, AB185144/DDBJ)
αm globin       LPQLFTPANHRSLDKFLHTVMRCLISKYR  (This study, AB185145/DDBJ)
α-1             LPAKFTYAVQCSIDKFLHVTMRLCISKYR  (P02015/Swiss-Prot)
α-2             FPADFTPQAQAAWDKFLAAVSGVLTSKYR  (Q90ZA5/Swiss-Prot, AF308869/DDBJ)

β globin        LQTEFTWHKQLAFEKLMRAVAHALSHSYH  (This study, AB185146/DDBJ))

                                  #
                 oo ooo  oo  o            $$
Human α         LPAEFTPAVHASLDKFLASVSTVLTSKYR  (Nagel, 1995)
                                  #
                 o  oooo oo  o            $$
Human β         FGKEFTPPVQAAYQKVVAGVANALAHKYH  (Nagel, 1995)
Helix position  GH1  H1                 HC1
```

Fig. 3. Alignment of the amino acid sequences of five kinds of globins from *A. mexicanum*. The 26 amino acid residues conserved in the five kinds of globins are shown by asterisks in the upper part of the alignment. The residue number is an arbitrary number with the numbering beginning from the N-terminal of the globins with the longest N-terminal extension. 'Helix position' refers to the helix position in human β globin (Nagel, 1995). The symbols refer to those residues functionally important in the human hemoglobin molecule (Fermi, 1975): #, residues involved in contact with the heme group; o, residues involved in contacts between subunits α_1 and β_1; $, residues involved in contacts between subunits α_1 and β_2.

α^m versus β; α^M versus β; α-1 versus β). Watt et al. (1980) reported that the tadpole β chain of *R. catesbeiana* differed greatly from that of the adult frog; only about 50% of the residues were identical. Based on the assumption of the overall rate of change in hemoglobin as about 1% per 3.3 million years (Wilson et al., 1977), they estimated a 50% difference corresponded to about 165 million years. Our findings, together with molecular relationships shown in Fig. 4A and B, suggest that the three α types of globins of the first group, α^M, α^m and α-1, which are homologous globin polypeptides in the same organism are caused by gene duplications of the α globin genes, and the second group, whose homology was estimated to be about 50%, may have occurred at least as long ago as the common ancestor to the axolotl and humans, dating about 165 million years ago.

Of considerable importance in hemoglobin function are the two contact regions between the subunits designated $\alpha_1\beta_1$ and $\alpha_1\beta_2$, and heme contact residues (Perutz et al., 1968; Fermi, 1975). In mammalian hemoglobin, oxygenation is accompanied by large changes at the $\alpha_1\beta_2$ contact and relatively small changes at the $\alpha_1\beta_1$ contact (Perutz, 1970). To gain information on axolotl hemoglobin, we

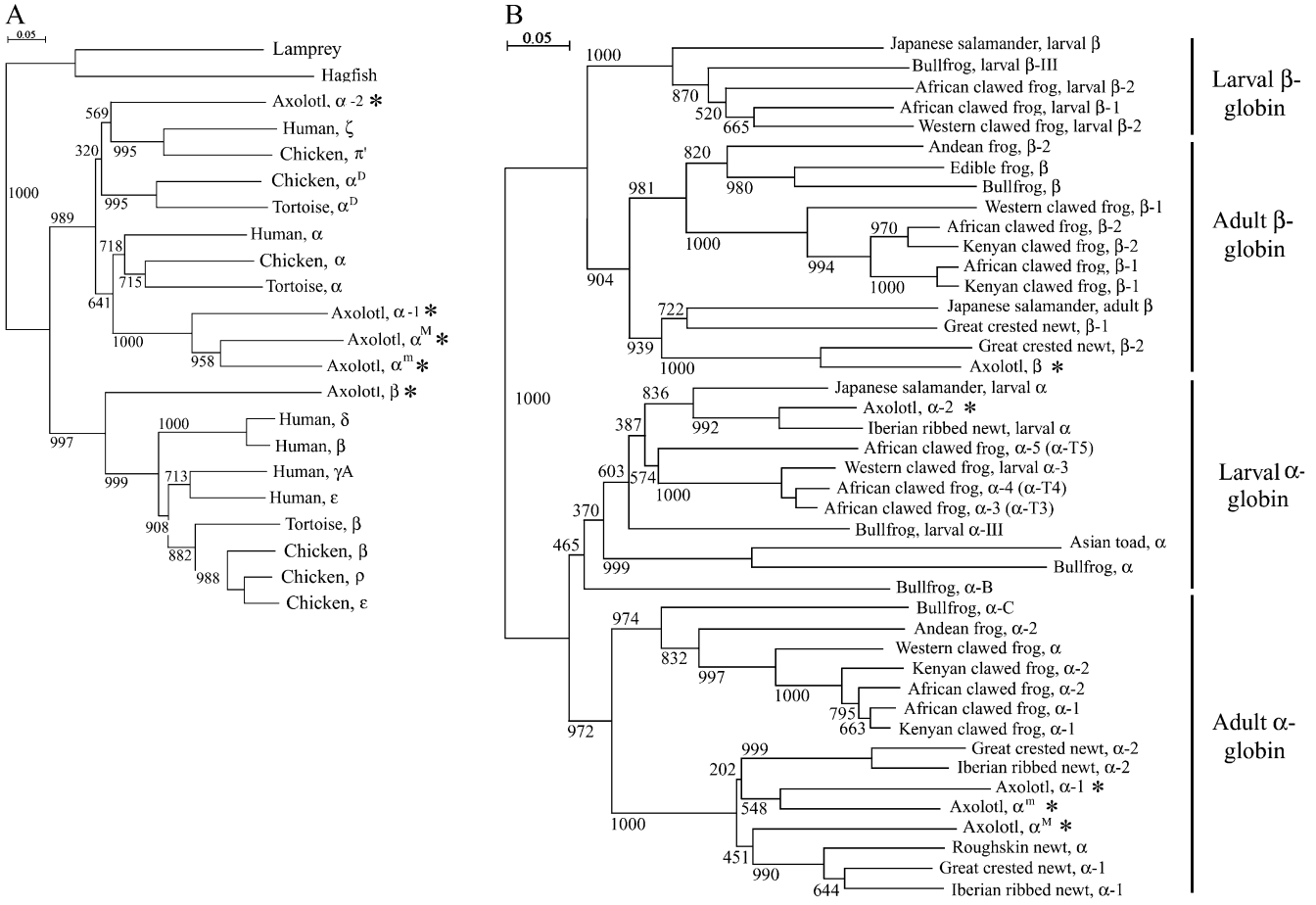


Fig. 4. Phylogenetic tree based on primary structures of 22 globins of tetrapods (A) and 43 globins of amphibians (B), constructed from 1000 bootstrap replications by the neighbor-joining method (Saitou and Nei, 1987). The scale shown in the upper left segment of the tree represents the evolutionary distances, given as the average number of substitutions per site. The scores shown at each node represent bootstrap values. The scientific names and accession numbers in data banks are as follows: African clawed frog: *Xenopus laevis* (larval β-1; P02137, larval β-2; P02133, α-1; P02012, α-2; P02013; α-T3; P06636, α-T4; P06637, α-T5; P06638, β-1; P02132, β-2 (fragment); Q91702); Andean frog: *Telmatobius peruvianus* (α-2 (fragment); P83113; β-2; P83114); Asian toad: *Bufo bufo gargarizans* (α; Q8AXX7); axolotl: *A. mexicanum* (α^M; AB185144, α^m; AB185145, α-1; P02015, α-2; Q90ZA5, β; AB185146); bullfrog: *Rana catesbeiana* (larval α-III; P02011, larval β-III; P02136, β; P02135, α-B; P51465, α-C; P55267, α (heart muscle); P02022); chicken: *Gallus gallus* (α; P01994, π′; P02007, α^D; P02001, β; P02112, ε; P02128, ρ; P02127; edible frog: *Rana esculenta* (β; P02134); great crested newt: *Triturus cristatus* (α-1; P10783, α-2; P10784, β-1; P10785, β-2; P10786); hagfish: *Myxine glutinosa* (P02209); human: *Homo sapiens* (α; P69905, ζ; P02008, β; P68871, δ; P02042, ε P02100, γ^A; P69891); Iberian ribbed newt: *Pleurodeles waltlii* (larval α; P11896, α-1; P06639, α-2 (fragment); P06640); Japanese salamander: *Hynobius retardatus* (larval α; Q9PVL9, larval β; Q9PVL8, adult β; Q9PVL7; Kenyan clawed frog: *Xenopus borealis* (α-1; P07430, α-2; P07431, β-1; P07432, β-2; P07433); lamprey: *Lampetra fluviatilis* (P02207); roughskin newt: *Taricha granulosa* (α P02014); tortoise: *Geochelone nigra* (α; P83135, α^D; P83124, β; P83123); Western clawed frog: *Xenopus tropicalis* (laval β-2; P08423, α; P07428, α-3; P08422, β-1; P07429). Axolotl globins are labeled with a symbol (*).

compared the appropriate amino acid residues with those involved in subunit interactions ($\alpha_1\beta_1$ and $\alpha_1\beta_2$) as well as in heme contacts as compared with amino acid residues responsible for subunit interactions and heme contacts of human deoxyhemoglobin described by Fermi (1975). As shown in Fig. 3, seven and eight out of 27 amino acid residues were variant at the $\alpha_1\beta_2$ contacts of Hb M and Hb m, respectively, whereas many more differences (18 and 17 residues) were apparent at the $\alpha_1\beta_1$ contacts of Hb M and Hb m. Such a marked difference in the $\alpha_1\beta_1$ contacts compared to that at the $\alpha_1\beta_2$ contacts has been also reported for *X. laevis* hemoglobin (Knöchel et al., 1985; Banville and Williams, 1985). The heme contact residues, on the contrary, were well conserved compared with those detected in human globins (Fermi, 1975): 13 residues (81.3%) and 12 residues (75.0%) were invariant for α^M and α^m or β globins, respectively.

In general, the Bohr effect in human hemoglobin has been postulated to involve the breakage of at least four salt bridges on oxygenation (Perutz, 1970). On this basis, the two salt linkages between the protonated NH_2-terminal group of the valine of one α chain and the COOH-terminal of the other α chain in the deoxyhemoglobin tetramer carboxyl group and between the guanidino group at HC3(141)α_1 and the aspartic acid at position 9 of the H helix of α_2 globin chain, designated Asp H9(126)α_2, are conserved. However, the other two salt linkages between the HC3(146)β imidazole group and the Asp (94)β in the same chain and between the carboxyl group of HC3(146)β and the ε-NH_3^+ of the lysine at C6(40)α are replaced with other amino acid residues. These findings may account in large part for the decreased Bohr effect on the axolotl hemoglobin compared with that of human hemoglobin (Amiconi et al., 1970; Hattingh and Bartels, 1973; Banville and Williams, 1985). It is worth noting that Lys C6(40) of the α^M chain is still conserved, compared with other α types of axolotl globins. Although Boissel et al. (1980) reported the replacement of histidine at H5(122)α by glutamine in axolotl, we detected that the amino acid residue at H5(122) of the two α chains, α^M and α^m, are unchanged, as reported in humans. Hence, it is advantageous to present multiple globins in situations where living organisms need to express a potential globin for adaptation to changes in surroundings or changes of life cycles, in particular, in amphibians (Hutchison et al., 1976; Perutz, 1983).

In addition, Boissel et al. (1980) found a proline at G6(107)α in the axolotl. Maruyama et al. (1980) noted that the proryl residue at this position causes distortion of the G helix of hemoglobin of *Glycera dibranchiata*, an annelid (Padlan and Love, 1974). We also detected a proline at G6 of α^M globin rather than α^m globin of the axolotl. To confirm the above estimation of changes in the tertiary structures, a crystallographic investigation is required.

3.3. Molecular Trees

The gene family of *X. laevis* globin consists of twelve genes in two clusters, each containing larval and adult α and β genes in a symmetrical arrangement (Hosbach et al., 1983). In the urodeles, we sequenced only three globin cDNAs from an adult axolotl, *A. mexicanum*, and analyzed their phylogenetic positions and relationships on molecular trees including representatives of tetrapodes (Fig. 4A) and, in particular, amphibians (Fig. 4B).

Fig. 4A shows that α types of axolotl globins (α^M, α^m and α-1, but not α-2) make a group, which is more closely related to the adult α type globins of amniotes (tortoises, birds and mammals) than the α^D type globins of some reptiles and birds or embryonic globins of birds and humans, π′ and ζ (Chapman et al., 1980; Aschauer et al., 1981). The α^D globin is one of the globin subunits of the hemoglobin D tetramer ($\alpha_2^D\beta_2$), which was first found in birds as a minor component of the embryonic and adult definitive erythrocytes (Hagopian and Ingram, 1971), and also found in some reptilians (Abbasi et al., 1988; Ruchnagel and Braunitzer, 1988; Gorr et al., 1998; Shishikura and Takami, 2001). On the contrary, the α-2 globin (AF308869) was assigned to an ancestral type of embryonic globins since human ζ and bird π′ globins split off.

Fig. 4B shows the molecular relationship of 43 globin sequences of amphibians, the members of which seem to be almost all the known sequences of amphibians stored in the Swiss-Prot data bank at the present time (release 45.5). It is not unreasonable to predict that the split of the adult types of globins from the larval types or vice versa might be significant for consideration of the evolution of vertebrate globin genes since each of the ancestor genes of α and β globins first duplicated itself into the four main clusters in total, and thereafter further divergences occurred, each branching into anuran species and urodele species.

Finally, to compare the distribution pattern of the five globin genes of the axolotl with that of *X. laevis*, *R. catesbeiana* or other urodele species (Fig 4B), we can predict that another two to four globin genes will be expressed in either larval stages or neotenous adult stages. However, it is necessary to investigate more currently unknown globin genes that will be equivalent to the unique gene arrangements reported in *Xenopus* and human globin gene families.

Acknowledgements

We thank Indiana University Axolotl Colony for the continuous supply of axolotl. This work was supported in part by Nihon University Joint Research Grant (2005) and the "High-Tech Research Center" Project for Private Universities and a matching fund subsidy from MEXT (Ministry of Education, Culture, Sports, Science and Technology), 2000–2004.

Appendix 1

Appendix 1A. Oligonucleotide primers used in the first PCR (RT-PCR) of globin cDNAs

Primer	Direction		Sequence
α^M–N		# of amino acid residue	1 2 3 4 5 6 7
	Forward	Amino acid sequence	V L T A E D K
		Nucleotide sequence	g t n y t n a c n g c n g a r g a y a a
α^m–N		# of amino acid residue	1 2 3 4 5 6 7
	Forward	Amino acid sequence	M V P L S A G
		Nucleotide sequence	a t g g t n c c n y t n w s n g c n g g
β–N		# of amino acid residue	6 7 8 9 10 11 12
	Forward	Amino acid sequence	E E R K D V G
		Nucleotide sequence	g a r g a r m g n a a r g a y g t n g g

Appendix 1B. Oligonucleotide primers used in 5' cRACE amplification of globin fragments

Primer	Direction		Sequence
Phosphated			A consensus sequence for three kinds of globins; α^M, α^m and β
		# of amino acid residue	65 64 63 62 61 60
	Reverse	Amino acid sequence	M V K K/A G H
		Nucleotide sequence	c a t c a c y t t s k y k c c r y g
SP–α^M		# of amino acid residue	41 40 39 38 37 36 35 34
	Reverse	Amino acid sequence	T K T V P E S D
		Nucleotide sequence	g t c t t g g t t a c a g g c t c c g a g t c
SP–α^m		# of amino acid residue	44 43 42 41 40 39 38 37 36
	Reverse	Amino acid sequence	Y T Q T Q C D I K
		Nucleotide sequence	a g g t c t g a g t c t g g c a g t c g a t c
SP–β		# of amino acid residue	40 39 38 37 36 35 34
	Reverse	Amino acid sequence	R R S W P Y V
		Nucleotide sequence	t c t c c g t g a c c a g g g a t a c a c a c

Appendix 1C. An oligonucleotide primer used in 3' RACE amplification of globin fragments

Primer	Direction		Sequence
3' RACE			A consensus sequence for three kinds of globins; α^M, α^m and β
		# of amino acid residue	60 61 62 63 64 65
	Forward	Amino acid sequence	H G K/A K V M
		Nucleotide sequence	c a y g g m r m s a a r g t g a t g

All primers were synthesized with a tag pUC/M13 sequencing primer (forward 17-mer: 5'-GTAAAACGACGGCCAGT-3', or reverse 17-mer: 5'-CAGGA-AACAGCTATGAC-3').

Appendix 2

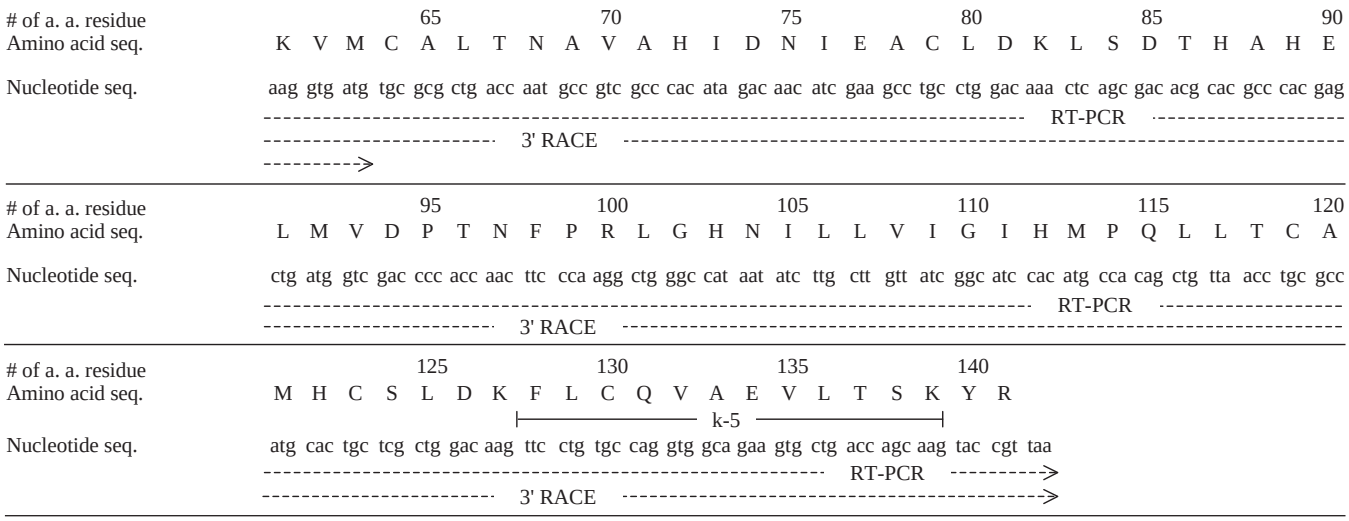
Appendix 2A. Primary structure and nucleotide sequence of α^M–globin of *Ambrystoma mexicanum*

# of a. a. residue	1 5 10 15 20 25 30
Amino acid seq.	V L T A E D K A N V K S A W D H I K G H E E A F G A E A L C
	N-terminal (1–24); k-1 (12–18); k-2 (19–)
Nucleotide seq.	gtg ctg act gct gag gac aag gcc aat gtg aag tcc gcg tgg gac cac atc aaa ggg cac gag gaa gcg ttt ggt gca gaa gcg ctt tgc
	RT-PCR; <– 5' cRACE; Nested PCR
# of a. a. residue	35 40 45 50 55 60
Amino acid seq.	R M F D S E P V T K T Y F G G K D I S E E S S Y L H S H G K
	k-2 (–40); k-3 (41–46); k-4 (47–60)
Nucleotide seq.	agg atg ttt gac tcg gag cct gta acc aag acc tac ttt ggc ggc aag gac atc tcc gaa gaa agc tct tac ctg cac agc cac ggc aag
	RT-PCR; 5' cRACE (–42); 58– ; Nested PCR

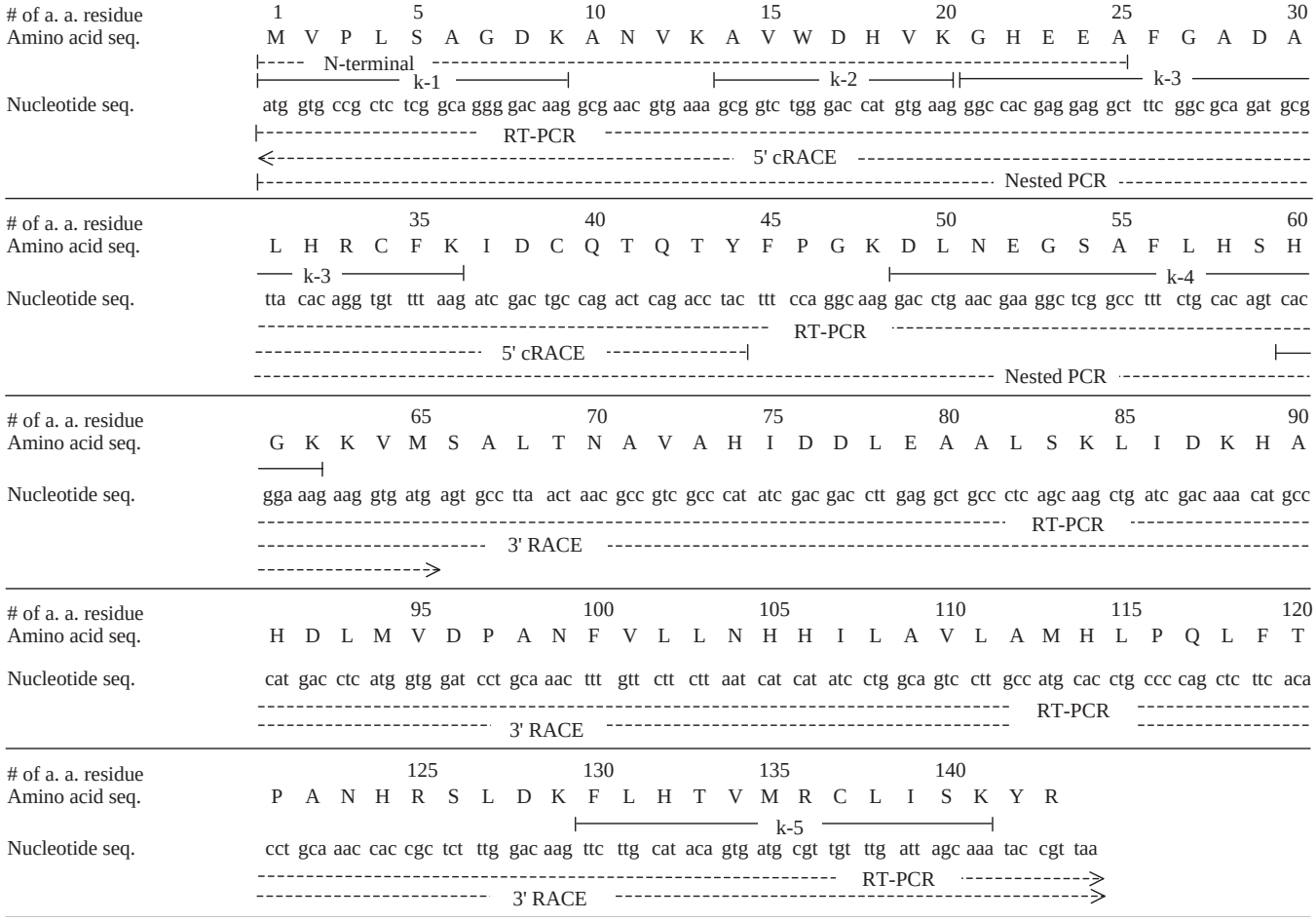
(continued on next page)

Appendix 2 (*continued*)

Appendix 2A

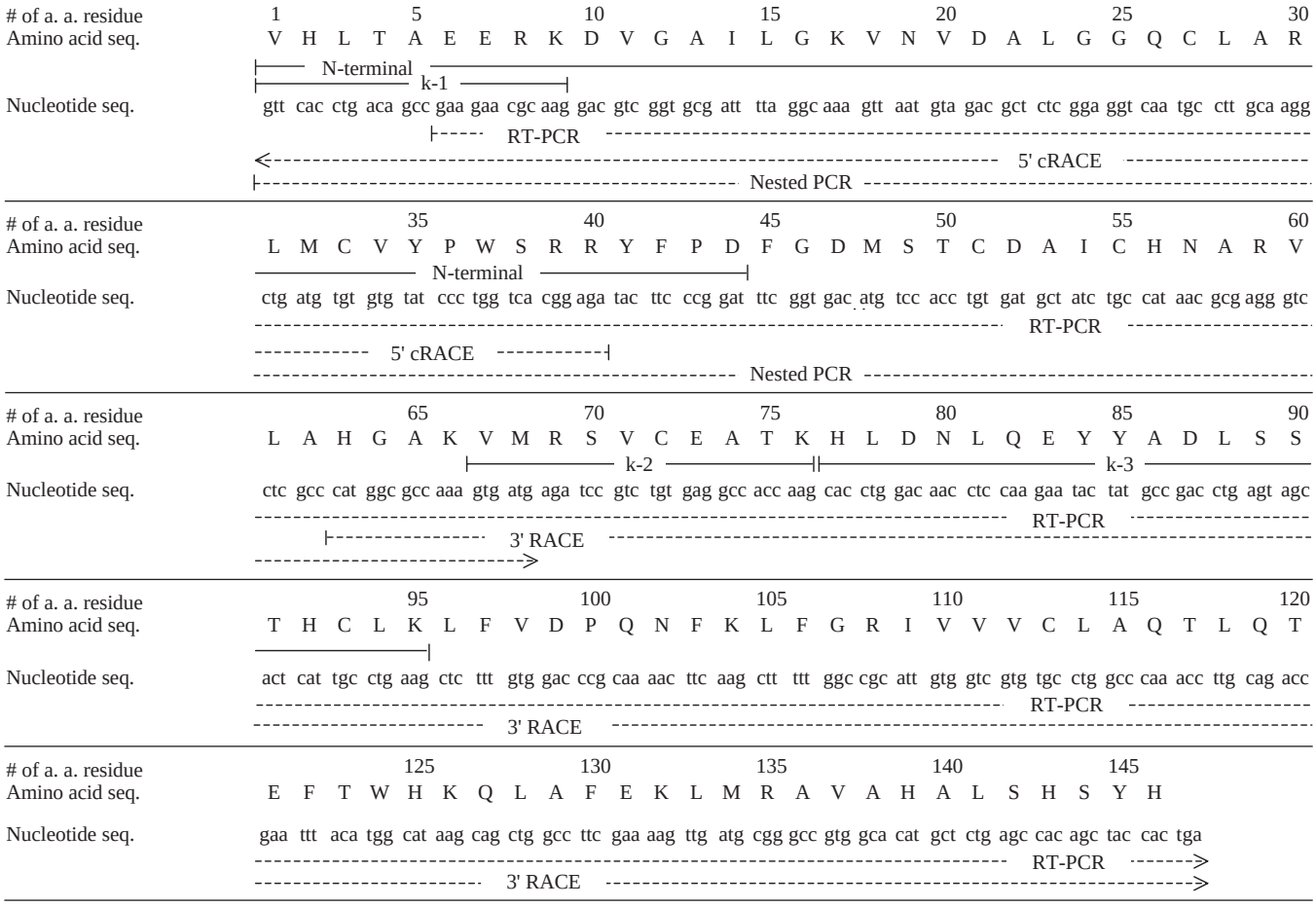


Appendix 2B. Primary structure and nucleotide sequence of α–globin of *Ambrystoma mexicanum*



Appendix 2 (*continued*)

Appendix 2C. Primary structure and nucleotide sequence of β–globin of *Ambrystoma mexicanum*



References

Abbasi, A., Wells, R.M.G., Brittain, T., Braunitzer, G., 1988. Primary structure of the hemoglobin from Sphenodon (*Sphenodon punctatus*, Tuatara, Rynchocephalia). Evidence for the expression of α^{D} -gene. Biol. Chem. Hoppe-Seyler 369, 755–764.

Aggarwal, S., Riggs, A., 1969. The hemoglobins of the bullfrog, *Rana catesbeiana*: I. Purification, amino acid composition, and oxygen equilibria. J. Biol. Chem. 244, 2372–2383.

Amiconi, G., Brunori, M., Antonini, E., Sorcini, M., Tentori, L., 1970. The haemoglobin of amphibia IX. Functional properties of haemoglobin from *Ambystoma tigrinum tigrinum* and Mexican axolotl. Int. J. Biochem. 1, 582–588.

Aschauer, H., Sanguansermsri, T., Braunitzer, G., 1981. Embryonale Hämoglobine des Menschen: Die Primärstruktur der ζ-Ketten. Hoppe-Seyler Z. Physiol. Chem. 362, 1159–1162.

Baldwin, T.O., Riggs, A., 1974. The hemoglobin of the bullfrog, *Rana catesbeiana*. Partial amino acid sequence of the β chain of the major adult component. J. Biol. Chem. 249, 6110–6118.

Banville, D., Williams, J.G., 1985. The pattern of expression of the *Xenopus laevis* tadpole α-globin genes and the amino acid sequence of the three major tadpole α-globin polypeptides. Nucleic Acids Res. 13, 5407–5421.

Boissel, J.-P., Wajcman, H., Labie, D., 1980. Hemoglobins of an amphibia, the neotenous *Ambystoma mexicanum*. Complete amino-acid sequence of the α chain of the major component using automatic solid-phase Edman degradation. Eur. J. Biochem. 103, 613–621.

Bunn, H.F., Forget, B.G., 1986. Hemoglobin: Molecular, Genetic and Clinical Aspects. Saunders, W.B., Co., Philadelphia.

Chapman, B.S., Tobin, A.J., Hood, L.E., 1980. Complete amino acid sequences of the major early embryonic α-like globins of the chicken. J. Biol. Chem. 255, 9051–9059.

Chauvet, J.-P., Acher, R., 1972. Phylogeny of hemoglobis. β chain of frog (*Rana esculenta*) hemoglobin. Biochemistry 11, 916–927.

Coates, M., Brimhall, B., Stenzel, P., Hermodson, M., Gibson, D., Jones, R.T., Vedvick, T., 1977. α-chain sequence of newt haemoglobin (*Taricha granulosa*). Aust. J. Biol. Sci. 30, 1–19.

Dorn, A.R., Broyles, R.H., 1982. Erythrocyte differentiation during the metamorphic hemoglobin switch of *Rana catesbeiana*. Proc. Natl. Acad. Sci. U. S. A. 79, 5592–5596.

Ducibella, T., 1974. The occurrence of biochemical metamorphic events without anatomical metamorphosis in the axolotl. Dev. Biol. 38, 175–186.

Elli, R., Giuliani, A., Tentori, L., Chiancone, E., Antonini, E., 1970. The hemoglobin of amphibia-X. Sedimentation behaviour of frog, triton and axolotl hemoglobins. Comp. Biochem. Physiol. 36, 163–171.

Fermi, G., 1975. Three-dimensional Fourier synthesis of human deoxy-haemoglobin at 2.5 Å resolution: refinement of the atomic model. J. Mol. Biol. 97, 237–256.

Friedman, M., Kruill, L.H., Cavins, J.F., 1970. The chromatographic determination of cystine and cysteine residues in proteins as *S*-β-(4-pyridylethyl)cysteine. J. Biol. Chem. 245, 3868–3871.

Frohman, M.A., Dush, M.K., Martin, G.R., 1988. Rapid production of full-length cDNAs from rare transcripts: amplification using a single gene-specific oligonucleotide primer. Proc. Natl. Acad. Sci. U. S. A. 85, 8998–9002.

Gorr, T.A., Mable, B.K., Kleinschmidt, T., 1998. Phylogenetic analysis of reptilian hemoglobins: trees, rates, and divergences. J. Mol. Evol. 47, 471–485.

Grasso, J.A., Casale, G.P., Chromey, N.C., 1979. Multiple hemoglobins in *Triturus cristatus*: their degradation by sulfhydryl compounds. Comp. Biochem. Physiol. B 63, 93–101.

Gratzer, W.B., Allison, A.C., 1960. Multiple hemoglobins. Biol. Rev. 35, 459–495.

Hagopian, H.K., Ingram, V.M., 1971. Developmental changes of erythropoiesis in cultured chick blastoderms. J. Cell Biol. 51, 440–451.

Hamada, K., Sakai, Y., Tsushima, K., Shukuya, R., 1966. Biochemical metamorphosis of hemoglobin in *Rana catesbeiana*: III. Molecular change of hemoglobin during spontaneous metamorphosis. J. Biochem. 60, 37–41.

Hattingh, J., Bartels, H., 1973. The oxygen affinity of axolotl blood and haemoglobin before and after metamorphosis. Respir. Physiol. 18, 1–13.

Hentschel, C.C., Kay, R.M., Williams, J.G., 1979. Analysis of *Xenopus laevis* globins during development and erythroid cell maturation and the construction of recombinant plasmids containing sequences derived from adult globin mRNA. Dev. Biol. 72, 350–363.

Hosbach, H.A., Wyler, T., Weber, R., 1983. The *Xenopus laevis* globin gene family: chromosomal arrangement and gene structure. Cell 32, 45–53.

Hutchison, V.H., Haines, H.B., Engbretson, G., 1976. Aquatic life at high altitude: respiratory adaptations in the Lake Titicaca frog, *Telmatobius culeus*. Respir. Physiol. 27, 115–129.

Jeanmougin, F., Thompson, J.D., 1998. Multiple sequence alignment with Clustal X. Trends Biochem. Sci. 23, 403–405.

Jeffreys, A.J., Wilson, V., Wood, D., Simons, J.P., 1980. Linkage of adult α- and β-globin genes in *X. laevis* and gene duplication by tetraploidizaiton. Cell 21, 555–564.

Kay, R.M., Harris, R., Patient, R.K., Williams, J.G., 1983. Complete nucleotide sequence of a cloned cDNA derived from the major adult α-globin mRNA of *X. laevis*. Nucleic Acids Res. 11, 1537–1542.

Kleinschmidt, T., Sgouros, J.G., Braunitzer, G., 1988. The primary structures of the major and minor hemoglobin components of the great crested newt (*Triturus cristatus*, Urodela, Amphibia). Biol. Chem. Hoppe-Seyler 369, 1343–1360.

Knöchel, W., Meyerhof, W., Stalder, J., Weber, R., 1985. Comparative nucleotide sequence analysis of two types of larval β-globin mRNAs of *Xenopus laevis*. Nucleic Acids Res. 13, 7899–7908.

Koussoulakos, S., Kaparos, G., Stathakos, D., 1986. Multiple hemoglobins in *Triturus cristatus*: their study by analytical isolelectrofocussing. Comp. Biochem. Physiol. B 83, 475–481.

Kuwada, T., Hasegawa, T., Satoh, I., Ishikawa, K., Shishikura, F., 2003. Crystallization and preliminary X-ray diffraction study of hemoglobin D from the Aldabra giant tortoise, *Geochelone gigantea*. Prot. Peptide Letters 10, 422–425.

MacLean, N., Jurd, R.D., 1971. The haemoglobins of healthy and anemic *Xenopus laevis*. J. Cell Sci. 9, 509–528.

Maples, P.B., Dorn, A.R., Broyles, R.H., 1983. Embryonic and larval hemoglobins during the early development of the bullfrog, *Rana catesbeiana*. Dev. Biol. 96, 515–519.

Maples, P.B., Palmer, J.C., Broyles, R.H., 1988. Determination of hemoglobin expression patterns in erythroid cells of *Rana catesbeiana* tadpoles. Comp. Biochem. Physiol. B 91, 755–762.

Maruyama, T., Watt, K.W.K., Riggs, A., 1980. Hemoglobins of the tadpole of the bullfrog, *Rana catesbeiana*. Amino acid sequence of the α chain of a major component. J. Biol. Chem. 255, 3285–3293.

McCutcheon, F.H., 1936. Hemoglobin function during the life history of the bullfrog. J. Cell. Comp. Physiol. 8, 63–81.

McCutcheon, F.H., Hall, F.G., 1937. Hemoglobin in amphibia. J. Cell. Comp. Physiol. 9, 191–197.

Moss, B., Ingram, V.M., 1968. Hemoglobin synthesis during amphibian metamorphosis: I. Chemical studies on the hemoglobins from larval and adult stages on *Rana catesbeiana*. J. Mol. Biol. 32, 481–492.

Nagel, R.L., 1995. Disorders of hemoglobin function and stability. In: Handin, R.I., Lux, S.E., Stossel, T.P. (Eds.), Blood: Principles and Practice of Hematology. J.B. Lippincott Co, Philadelphia, pp. 1591–1644.

Padlan, E.A., Love, W.E., 1974. Three-dimensional structure of hemoglobin from the polychaete annelid *Glycera dibranchiata*, at 2.5 A resolution. J. Biol. Chem. 249, 4067–4078.

Perutz, M.F., 1970. The Bohr effect and combination with organic phosphates. Nature 228, 734–739.

Perutz, M.F., 1983. Species adaptation in a protein molecule. Mol. Biol. Evol. 1, 1–28.

Perutz, M.F., Muirhead, H., Cox, J.M., Goaman, L.C.G., 1968. Three-dimensional Fourier synthesis of horse oxyhaemoglobin at 2.8 Å resolution: the atomic model. Nature 219, 131–139.

Rüchnagel, K.P., Braunitzer, G., 1988. The primary structure of the major and minor hemoglobin component of adult Western painted turtle (*Chrysemys picta bellii*). Biol. Chem. Hoppe-Seyler 369, 123–131.

Saitou, N., Nei, M., 1987. The neighbor-joining method: a new method for reconstructing phylogenetic trees. Mol. Biol. Evol. 4, 406–425.

Shishikura, F., 2002. The primary structure of hemoglobin D from the Aldabra giant tortoise, *Geochelone gigantea*. Zool. Sci. 19, 197–206.

Shishikura, F., Snow, J.W., Gotoh, T., Vinogradov, S.N., Walz, D.A., 1987. Amino acid sequence of the monomer subunit of the extracellular hemoglobin of *Lumbricus terrestris*. J. Biol. Chem. 262, 3123–3131.

Shishikura, F., Takami, K., 2001. The amino acid sequence of the α- and β-globin chains of hemoglobin from the Aldabra giant tortoises, *Geochelone gigantea*. Zool. Sci. 18, 515–526.

Smith, D.J., Zhu, H., Kolatkar, P.R., Tam, L.-T., Baldwin, T.O., Roe, B.A., Broyles, R.H., Riggs, A.F., 1993. The hemoglobin of the bullfrog, *Rana catesbeiana*. The cDNA-derived amino acid sequences of the α chains of adult hemoglobins B and C: their roles in deoxygenation-induced aggregation. J. Biol. Chem. 268, 26961–26971.

Tam, L.-T., Gray, G.P., Riggs, A.F., 1986. The hemoglobins of the bullfrog *Rana catesbeiana*. The structure of the β chain of component C and the role of the α chain in the formation of intermolecular disulfide bonds. J. Biol. Chem. 261, 8290–8294.

Tam, L.-T., Manning, D., Cox, D.J., Riggs, A.F., 1993. The hemoglobin of the bullfrog, *Rana catesbeiana*. Deoxygenation-linked association of tetrameric components B and C to form the trimer BC_2: sedimentation analysis and oxygen equilibria. J. Biol. Chem. 268, 26972–26977.

Watt, K.W.K., Maruyama, T., Riggs, A., 1980. Hemoglobins of the tadpole of the bullfrog, *Rana catesbeiana*. Amino acid sequence of the β chain of a major component. J. Biol. Chem. 255, 3294–3301.

Weber, R., Blum, B., Muller, P.R., 1991. The switch from larval to adult globin gene expression in *Xenopus laevis* is mediated by erythoroid cells from distinct compartments. Development 112, 1021–1029.

Wilson, A.C., Carlson, S.S., White, T.J., 1977. Biochemical evolution. Annu. Rev. Biochem. 46, 573–639.

Yamaguchi, M., Takahashi, H., Wakahara, M., 2000. Erythropoiesis and unexpected expression pattern of globin genes in the salamander *Hynobius retradatus*. Dev. Genes Evol. 210, 180–189.